



**Universität  
Zürich** UZH

Department of Geography

# **Quantifying the Impact of Transportation Diversity on Human Mobility Resilience During Extreme Disasters: A Case Study of NYC During Hurricane Ida**

GEO 511 Master's Thesis

**Author:** Mengqi Li, 23-741-986

**Supervised by:** Prof. Dr. Cheng Fu (cheng.fu@cau.edu.cn)

**Faculty representative:** Prof. Dr. Robert Weibel

23.04.2025



**Universität  
Zürich** <sup>UZH</sup>

Master's Thesis  
presented to the Department of Geography  
of the University of Zurich  
**Master of Science**

Quantifying the Impact of Transportation  
Diversity on Human Mobility Resilience  
During Extreme Disasters: A Case Study of  
NYC During Hurricane Ida

**Author: Mengqi LI**  
Student ID: 23-741-986

Supervised by: Prof. Dr. Cheng Fu (China Agricultural University)  
Faculty representative: Prof. Dr. Robert Weibel

Department of Geography

Submission Date: 22.04.2025

## Abstract

This study proposes an innovative multi-modal and spatially explicit framework to quantify urban mobility resilience during extreme weather events, demonstrated through the case study of Hurricane Ida's 2021 impact on New York City. The framework integrates dynamic baseline forecasting using the Prophet time-series model, enabling the precise characterization of expected mobility patterns under normal conditions. By comparing observed mobility with these dynamic baselines across subway, taxi, and bike-sharing systems, it effectively identifies the magnitude and duration of disruptions at a zone-level scale. Results revealed substantial differences by transport mode: subways experienced severe, prolonged disruptions due to fixed infrastructure, whereas taxis and bike-sharing systems recovered rapidly, acting as effective substitutes. Spatial analysis highlighted central areas as more resilient due to greater transportation diversity and infrastructure density, while peripheral neighborhoods faced longer disruptions. Transportation diversity was identified as a significant predictor of resilience, yet its impact varied according to local socio-economic and infrastructural contexts. Employing Multiscale Geographically Weighted Regression (MGWR), the study showed that diversity yielded the greatest resilience benefits in neighborhoods characterized by higher incomes, greater connectivity, and lower social vulnerability. In contrast, socio-economic and environmental constraints in other areas limited residents' ability to utilize available transportation options fully. The proposed framework offers practical guidance for urban planners and policymakers on identifying local vulnerabilities, prioritizing equitable infrastructure investments, and enhancing the adaptive capacity of transportation systems in the face of increasing climate-related risks.

**Keywords:** Urban Mobility Resilience, Transportation Diversity, Hurricane Ida, Dynamic Baseline Modeling

# Acknowledgment

I would like to express my sincere gratitude to Prof. Dr. Cheng Fu for his valuable guidance throughout this research. Thank you for those brainstorming sessions that always sparked new ideas, and for patiently answering my endless "What if..." questions. His insightful feedback, generous support, and patient mentorship greatly enhanced the quality of this work. My sincere thanks also go to Prof. Dr. Robert Weibel, my faculty representative, for his kindness and constructive feedback during my doctoral application process. I am incredibly fortunate to have had such supportive mentors at the University of Zurich, and I will always cherish their generosity and inspiration.

Lastly, I am deeply grateful to my parents for their endless love and quiet strength. To my boyfriend, thank you for standing by me through every late night and every moment of doubt. Your faith in me never wavered, and for that, I am truly thankful.

# Contents

<b>Acknowledgment</b>	<b>iii</b>
<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Acronyms</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Related Work</b>	<b>4</b>
2.1 Disaster-Induced Mobility Disruptions . . . . .	4
2.2 Urban Resilience . . . . .	6
2.3 Quantifying Resilience . . . . .	8
2.4 Transportation Diversity . . . . .	9
2.5 Data-Driven Approaches to Analyzing Mobility . . . . .	9
2.6 Research Gaps . . . . .	10
2.7 Objectives and Research Questions . . . . .	12
<b>3 Study Area and Data</b>	<b>14</b>
3.1 Study Area . . . . .	14
3.2 Data . . . . .	15
3.2.1 Mobility Data . . . . .	15
3.2.2 Road Network . . . . .	15
3.2.3 Census Data . . . . .	16
<b>4 Methodology</b>	<b>18</b>
4.1 Data Preprocessing . . . . .	18
4.1.1 Data Cleaning . . . . .	19
4.1.2 Data Aggregation . . . . .	20
4.1.3 Mapping with Geographic Zones . . . . .	21
4.1.4 Data Transformation . . . . .	22
4.2 Mobility Resilience Computation . . . . .	23
4.2.1 Conception of Mobility Resilience Model . . . . .	23

4.2.2	Improved Dynamic Baseline	27
4.2.3	Determine the Disaster Impact Period	30
4.2.4	Resilience Quantification	31
4.3	Transportation Diversity Index	34
4.3.1	Multi-modal Transportation Availability	34
4.3.2	Richness & Evenness	35
4.3.3	Transportation Diversity Index	36
4.4	Statistical and Spatial Analysis	37
4.4.1	Variables Selection	38
4.4.2	Statistical Analysis	40
4.4.3	Spatial Analysis	40
<b>5</b>	<b>Results</b>	<b>43</b>
5.1	Exploratory Data Analysis	43
5.1.1	Temporal Distribution of the Mobility Data	43
5.1.2	Spatial Distribution of the Mobility Data	47
5.1.3	Census Data	49
5.2	Descriptive Statistics	53
5.2.1	Transportation Diversity Calculation	53
5.2.2	Mobility Resilience Model Result	57
5.2.3	Variable Transformation	69
5.3	Variables Selection	71
5.3.1	Moran's I Result	71
5.3.2	Variance Inflation Factor Result	72
5.4	Statistical and Spatial Analysis	74
5.4.1	Correlation Between TDI and Resilience Metrics	74
5.4.2	Multivariate and Spatial Modeling	74
<b>6</b>	<b>Discussion</b>	<b>80</b>
6.1	Impacts of Hurricane Ida on Transportation Modes	80
6.2	Spatial Variation and the Role of Transportation Diversity	81
6.3	Socio-Spatial Interactions and Urban Vulnerability	82
6.4	Performance of the Proposed Framework	83
6.5	Limitations	84
<b>7</b>	<b>Conclusion</b>	<b>86</b>
7.1	General Conclusions	86
7.2	Future Work	87
	<b>References</b>	<b>88</b>

<b>8 Appendix</b>	<b>96</b>
<b>9 Personal Declaration</b>	<b>104</b>

# List of Figures

1.1	Number of international disasters . . . . .	1
1.2	Tracks of Hurricane Ida . . . . .	2
1.3	Modes of commuting to work in NYC . . . . .	3
3.1	Study area: New York City’s five boroughs . . . . .	14
3.2	Spatial distribution of transportation infrastructures. . . . .	16
4.1	Research framework. . . . .	18
4.2	Spatial overlay of taxizones and census tracts. . . . .	22
4.3	Resilience Triangle (adapted from Bruneau et al. (2003)). . . . .	25
4.4	Improved mobility resilience model . . . . .	26
4.5	Resilience model comparison . . . . .	32
4.6	Heatmaps of transportation infrastructure . . . . .	35
5.1	Temporal variations of transportation modes . . . . .	45
5.2	Citi Bike volume (daily and hourly) . . . . .	46
5.3	Subway volumes (daily and hourly) . . . . .	47
5.4	Operational coverage of transit systems . . . . .	48
5.5	Spatial distribution of mobility volume . . . . .	49
5.6	Histograms of socioeconomic and environmental variables . . . . .	50
5.7	Spatial distribution of socioeconomic and environmental variables. . . . .	52
5.8	Histograms of transportation’s richness and evenness. . . . .	54
5.9	Spatial distribution of transportation richness and evenness. . . . .	56
5.10	Histogram and spatial distribution of the TDI . . . . .	56
5.11	Mobility resilience model performance. . . . .	58
5.12	Mobility anomaly detection result . . . . .	60
5.13	Weighted aggregation of transportation modes . . . . .	61
5.14	Temporal patterns of disruptions and recoveries . . . . .	63
5.15	Correlation matrices of resilience metrics . . . . .	64
5.16	Violin plots of resilience metrics . . . . .	66
5.17	Spatial patterns of resilience metrics in NYC taxizones . . . . .	68
5.18	Comparison of variable histograms pre- and post-transformation . . . . .	70
5.19	Bar plot of Moran’s I values and p-values. . . . .	71

5.20 LISA cluster maps of Moran's I . . . . .	72
5.21 VIF values for candidate predictors. . . . .	73
5.22 Final selected variables for regression analysis. . . . .	73
5.23 Preliminary correlation analysis between the TDI and resilience metrics. . . . .	74
5.24 MGWR coefficient maps . . . . .	78
5.25 MGWR p-value maps . . . . .	79

# List of Tables

3.1	Description of mobility datasets . . . . .	15
3.2	Descriptions of socio-economic and environmental data . . . . .	17
4.1	Resilience model comparison . . . . .	27
4.2	Seasonality configurations in the Prophet forecasting model . . . . .	29
4.3	Metrics of traffic mode diversity. . . . .	36
4.4	Variable definitions and units for resilience modeling . . . . .	38
5.1	Descriptive statistics of provided model performance . . . . .	59
5.2	Variable transformation summary . . . . .	69
5.3	Regression results and statistical significance . . . . .	75
8.1	Prophet cross-validation summary by taxizone (Part 1) . . . . .	96
8.2	Prophet cross-validation summary by taxizone (Part 2) . . . . .	97
8.3	Prophet cross-validation summary by taxizone (Part 3) . . . . .	98
8.4	Prophet cross-validation summary by taxizone (Part 4) . . . . .	99
8.5	Prophet cross-validation summary by taxizone (Part 5) . . . . .	100
8.6	Prophet cross-validation summary by taxizone (Part 6) . . . . .	101
8.7	Prophet cross-validation summary by taxizone (Part 7) . . . . .	102
8.8	Prophet cross-validation summary by taxizone (Part 8) . . . . .	103

# List of Acronyms

ACS	American Community Survey
DCP	Department of City Planning
DOT	Department of Transportation
FHV	For-Hire Vehicles
GWR	Geographically Weighted Regression
HVFHV	High-Volume For-Hire Vehicles
LISA	Local Indicators of Spatial Association
MGWR	Multiscale Geographically Weighted Regression
MTA	Metropolitan Transportation Authority
OLS	Ordinary Least Squares
PLUTO	Primary Land Use Tax Lot Output
SAR	Spatial Autoregressive Model
SEM	Spatial Error Model
TDI	Transportation Diversity Index
TLC	Taxi and Limousine Commission
R	Resilience
RT	Recovery Time
RL	Resilience Loss
MD	Maximum Deviation
Ratio	Resilience Ratio

# 1 Introduction

In the 21st century, cities worldwide are increasingly challenged by the intensifying impacts of climate change. Extreme weather events, such as hurricanes, floods, wildfires, and heat waves, have grown in frequency and severity, posing significant threats to critical urban infrastructure. Systematically documented by the Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report<sup>1</sup>, since the 1970s, extreme weather events have intensified from hurricanes and heavy precipitation to extensive flooding. Concurrent evidence in EM-DAT's international disaster database<sup>2</sup> shows an exponential increase in disaster events from 1900 to 2024, further substantiating escalating climate hazards globally.

## Number of recorded natural disaster events, 1900 to 2023



The number of global reported natural disaster events in any given year. Note that this largely reflects increases in data reporting, and should not be used to assess the total number of events.



Data source: EM-DAT, CRED / UCLouvain (2024)

Note: Data includes disasters recorded up to April 2024.

OurWorldinData.org/natural-disasters | CC BY

Figure 1.1: Number of international disasters (from The Center for Research on the Epidemiology of Disasters)

<sup>1</sup> <https://www.ipcc.ch/>

<sup>2</sup> The Center for Research on the Epidemiology of Disasters (CRED): <https://www.emdat.be/>

Mobility resilience encompasses the capacity of a city to sustain and rapidly recover the movement of people despite disruptions, thereby securing both emergency response capabilities and everyday societal functions (Yao et al., 2024b). In the wake of Hurricane Ida in 2021, New York City experienced unprecedented challenges. The extreme rainfall and subsequent flash floods wrought widespread disruption, significantly impeding the flow of people and goods. The most intense period of impact observed between 6:00 p.m. on September 1st, 2021, and 2:00 a.m. on September 2, 2021 (Figure 1.2), revealed how even densely interconnected urban areas could see their mobility networks come to a near standstill (Yuan et al., 2024). Such events prompt us to reexamine resilience not solely in terms of infrastructure robustness but also in the adaptability of urban populations, how quickly movement patterns return to normal, and how diverse mobility options mitigate the risks of systemic failure.

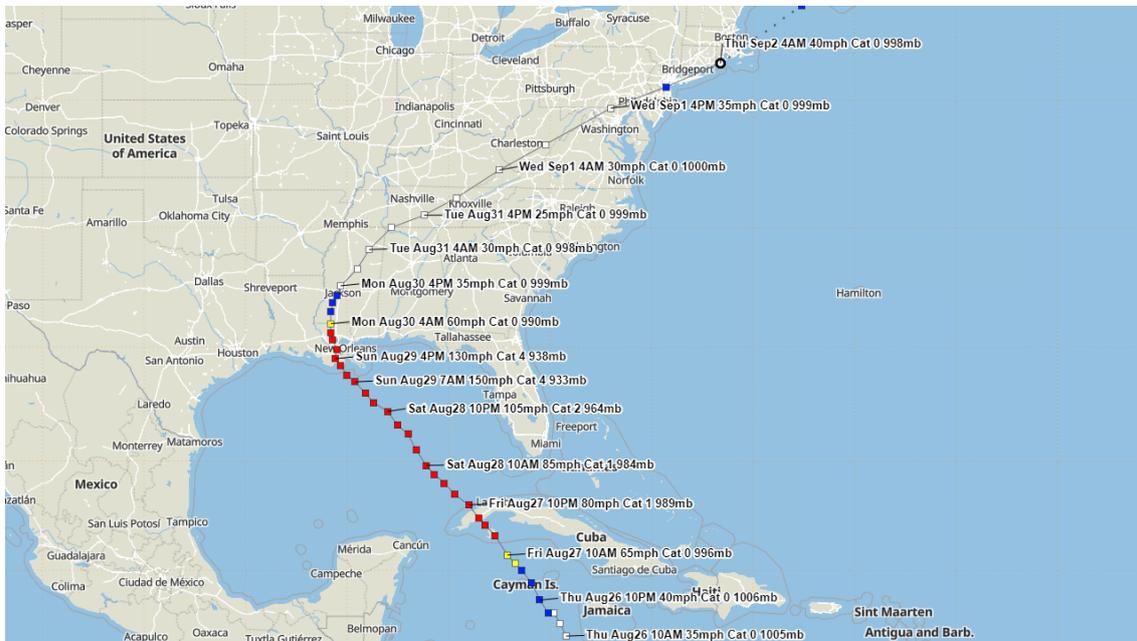


Figure 1.2: Tracks of Hurricane Ida (from Operational National Hurricane Center)

Transportation systems play a critical role in urban resilience by ensuring the mobility of residents and the continuity of essential services during and after crises. Human mobility, the movement of people within and between different areas of the city, is a vital indicator of how quickly and effectively a city can recover from such disasters (Hu et al., 2021; Xia et al., 2018). Public transportation data, encompassing various modes such as subways, buses, taxis, and shared bikes, provides a rich source of information that reflects the mobility patterns of diverse population groups across different urban zones (Cats, 2024; Hu et al., 2021; Tang et al., 2015; Veloso et al., 2011). As the subway system (32.4% of commuting) rep-

resents the backbone of daily travel for millions, disruptions to this system can significantly affect the city’s overall mobility. Meanwhile, modes such as driving alone (30.4%), buses (12.1%), and walking (10.1%) provide critical flexibility during emergencies, allowing different population groups to adapt and maintain mobility(Figure 1.3).

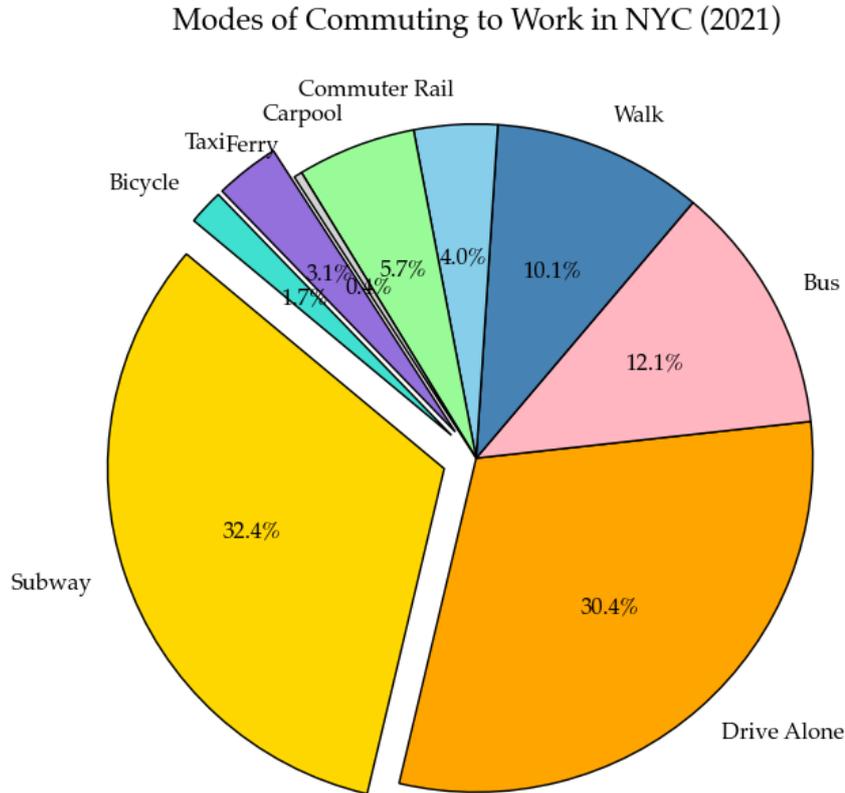


Figure 1.3: Modes of commuting to work in NYC (Source: U.S. Census Bureau, ACS 2021, Table B08406)

This study aims to explore how transportation diversity contributed to the city’s ability to withstand and recover from Hurricane Ida’s impacts. This research will offer insights into the importance of diverse, robust transportation infrastructure by analyzing overall traffic resilience loss and how different zones with varied transportation options were affected. Understanding these dynamics is crucial for developing more resilient urban environments and informing future disaster preparedness and recovery strategies.

## 2 Related Work

In urban areas facing extreme disasters, transportation systems' vulnerability and recovery capacity are pivotal for assessing urban resilience. This section reviews existing literature on disaster impacts on urban mobility patterns, urban resilience theories and their quantification methods, transportation diversity, and data-driven analytical approaches. It also identifies gaps in current research to establish a theoretical foundation for proposing a novel quantitative assessment framework.

### 2.1 Disaster-Induced Mobility Disruptions

Natural disasters such as hurricanes, floods, and pandemics can significantly disrupt urban transportation systems and travel patterns (Rodrigue, 2024; Wan et al., 2025; Yao et al., 2024a). Research shows that after large-scale disasters, urban travel demand typically plummets, and both spatial and temporal distributions of trips become irregular. For example, during Hurricane Sandy in 2012, all motorized transport in New York City was temporarily shut down (Kontou et al., 2017). In the immediate aftermath, travel volumes across different modes of transportation dropped sharply, and the suspension of public transit forced many passengers to seek alternatives or cancel their trips altogether (Kontou et al., 2016). A survey of 400 commuters revealed that a considerable proportion changed their mode of travel, routes, or departure times post-Sandy. Many former public transit users had to switch to other means or abandon commuting entirely (Kontou et al., 2016). Some groups mitigated travel needs through remote work, but this also prolonged the time required to return to normal commuting routines (Kontou et al., 2017). Overall, the recovery of commuting post-disaster is influenced by various factors: commuters with higher income and education levels tend to delay their return (often due to increased remote work), those with longer commuting distances face longer delays, and the longer the disruption in transportation services, the more prolonged the abnormal travel patterns persist (Zheng et al., 2024; Montazer and Young, 2024).

Disasters disrupt urban transportation systems, leading to significant variations in travel behavior across different regions and demographic groups. For instance, during Hurricane Harvey, analysis of GPS data from over 150,000 users in the Greater Houston Area revealed that disadvantaged minority populations were less likely to evacuate compared to wealthier white residents (Deng et al., 2021). Additionally, individuals exhibit distinct behavioral patterns during such events. Studies have categorized people into "returners" and "explorers" based on mobility behaviors. Research on Hurricane Ian indicated that "explorers" tend to travel longer distances, including evacuations, while "returners" stay closer to home but may expand their movement range due to evacuation needs (He et al., 2024).

Spatially, people generally moved away from hurricane paths and concentrated in nearby major cities for safety (He et al., 2024). Moreover, different types of disasters impact travel in different ways. Sudden-onset events such as earthquakes and hurricanes tend to damage transportation infrastructure, causing abrupt declines in travel immediately. At the same time, chronic disasters or public health crises induce long-term behavioral changes (Hunter et al., 2024). The COVID-19 pandemic transformed the way individuals got around in cities. Car use and public transportation declined significantly everywhere during the early 2020 lockdowns. Research indicates that car travel and walking recovered more quickly in less densely populated cities, while in cities that heavily rely on public transportation, recovery was slower (Hunter et al., 2024).

Additionally, the pandemic altered the spatial and temporal characteristics of travel: the average daily activity radius (spatial scale) shrank significantly, and commuting time regularity (temporal scale) was disrupted (Santana et al., 2023). After restrictions were lifted, travel distances recovered relatively quickly, but commuting timing patterns, such as peak-hour synchronization, recovered more slowly, indicating long-term behavioral adjustments (Santana et al., 2023). Some studies even compared the impacts of severe floods and large-scale social events on population mobility, finding similar patterns of abnormal travel distribution at national and local levels. This suggests that insights from familiar events may help us understand population movement during rare disasters (Loreti et al., 2025).

Overall, natural disasters disrupt normal urban travel patterns in the short term, resulting in pronounced spatiotemporal anomalies and modal shifts (Rodrigue, 2024). Travel recovery varies across regions and demographic groups, with vulnerable communities often more severely affected. Urban residents adopt various

adaptive strategies to mitigate the impact of disasters on transportation, such as altering travel routes and modes, shifting departure times, or substituting travel with remote options (Kontou et al., 2016). These studies provide empirical insights into traffic vulnerability and recovery pathways during disasters, laying a foundation for enhancing transportation system resilience.

## 2.2 Urban Resilience

Urban resilience is the ability of urban environments—infrastructures, institutions, and communities—to withstand, absorb, and recover from disruptive events, yet maintain vital functions (Meerow et al., 2016; Godschalk and Baxter). Resilience in the urban domain originally borrowed concepts from theories in ecology and systems (Walker et al., 2004) and was later developed in finer detail, especially in the fields of infrastructure and hazard management, using the “4Rs” model of robustness, redundancy, resourcefulness, and rapidity (Bruneau et al., 2003). Robustness is the ability of the system to resist disturbances without significant functional loss; it refers to the presence of alternative paths to maintain functions; resourcefulness describes the ability of the system to change; and rapidity refers to the rate of recovery (Bruneau et al., 2003; Henry and Emmanuel Ramirez-Marquez, 2012).

Urban resilience is defined by its dynamic and multidimensional facets, involving aspects of mid-term adaptation, short-term recovery, and long-term transformation (Chelleri et al., 2015). Empirical experience across diverse global settings proves the successful application of the concept of resilience in urban planning and government, as in the case of OneNYC and the Special Initiative for Rebuilding and Resiliency in New York City after Hurricane Sandy (Bill de Blasio (2015); Brown (2014)). These schemes reflect technical and social aspects, emphasizing strong infrastructure, backup, community solidarity, and fair access to recovery assets (Cutter et al., 2010; Roy and Kar, 2022).

Mobility resilience is a particular aspect of the overall concept of urban resilience that emphasizes the ability of transportation systems to withstand disruption. Having a choice of various redundant methods of transport—across various modes and routes—is crucial to enabling mobility in the event of crises (Rahimi-Golkhandan et al., 2021; Murray-Tuite, 2006). Empirical evidence of this comes from major cities. For example, in Hurricane Sandy in New York City, the presence of alternative forms of transportation, including buses, ferries, cycling paths, and walk-

ing paths, greatly lessened the negative effects felt when subway service ceased (Donovan and Work, 2015; Kaufman et al., 2012). Similar results were noted in other metropolises: Hong Kong's integrated multi-modal transit system showed high mobility resilience through its redundancy and flexibility (Haraguchi et al., 2022), while in London, the conditions showed the frailty of transit systems that, though highly efficient, are lacking in redundancy (Lu and Stead, 2013).

Measurement of mobility and urban resilience is critical in enabling comparative analysis and informing policy through the analysis of performance trends under disruption, traditionally known as the resilience curve, using the approach of the resilience triangle that calculates functionality loss and recovery (Bruneau et al., 2003; Roy et al., 2019). Network analysis methods also compare failures in nodes or links, hence measuring changes in connectivity, accessibility, and delay in travel (Chang and Nojima, 2001; Mattsson and Jenelius, 2015). Progress in this field resulted in the design of composite indicators that combine different aspects of resilience in enabling a comprehensive assessment (Rodrigue, 2024).

The application of data-driven methods has greatly improved the analysis of resilience using high-resolution data, for example, mobile phone location data, GPS traces, transit smartcard records, and social media data (Bengtsson et al., 2011; Ilbeigi, 2019; Roy et al., 2019). For instance, studies based on mobile phone use enabled swift assessments of displacement and recovery behaviors after catastrophic events, like the earthquake in Haiti (Bengtsson et al., 2011). Also, GPS records of taxi rides provided detailed knowledge of mobility disruption and recovery dynamics in significant events like Hurricane Sandy in the city of New York (Ilbeigi, 2019; Roy and Kar, 2022). Further, social media data complement these records by providing real-time analyses of behavioral response and social networks during crises (Roy et al., 2019). The academic literature on urban resilience continually emphasizes the need to include technical strength, redundancy, flexible management, and fair social determinants in urban systems. In mobility resilience, modal diversity in transport and data-driven analysis methods are critical for enhancing urban readiness and flexibility to deal with disruptions. Despite significant advancements, additional studies are needed to develop and ensure the successful incorporation of corresponding indicators in urban design methods.

## 2.3 Quantifying Resilience

Measuring urban resilience involves establishing how cities respond, resist, and recover after disasters. As initial models, like the triangle of resilience by [Bruneau et al. \(2003\)](#), measure resilience in terms of loss of performance and recovery over time and represented graphically as the area under curves of performance, this was later improved by [Cimellaro et al. \(2010\)](#); [Henry and Emmanuel Ramirez-Marquez \(2012\)](#) through the addition of time-dependent recovery measures that yield normalized indicators of the level and rate of system recovery.

Several studies employ composite indices to capture broader dimensions of resilience. [Cutter et al. \(2010\)](#) suggested the Baseline Resilience Indicators for Communities (BRIC), which integrates socioeconomic, institutional, and infrastructure variables into composite resilience scores. [Rose \(2004\)](#) also distinguished between static and dynamic economic resilience and quantified resilience as the difference between actual and potential economic losses following disruptions.

Mobility resilience measurements quantify the capacity of transportation networks to maintain or recover the level of service in a reasonable time interval after a disruption. Retaining connectivity, delay in travel time, network capacity loss, and service recovery are typical measurements. [Donovan and Work \(2015\)](#) used taxi GPS data to quantify New York City mobility resilience to Hurricane Sandy by estimating trends in delay and recovery. NYC's road network resilience was evaluated by [Mirjalili et al. \(2023\)](#) using graph-based connectivity measurements after snowstorms, highlighting the importance of fast recovery to enable mobility. Across the globe, researchers followed similar methods. [Chang and Nojima \(2001\)](#) examined the recovery of the Kobe transport system after the 1995 earthquake by calculating demand fulfilled as demand over time. [Cox et al. \(2011\)](#) estimated London transit recovery after the 2005 bombings by using retention in ridership relative to worst-case baselines. [Sun et al. \(2020b\)](#) also dealt with recovery indicators in transportation, classifying methods as topological, functional, and hybrid methods, noting that the choice of indicators depends on the context.

Overall, resilience measurement integrates various methods, ranging from performance curves and composite indexes to simulation models and network analyses, to support overall urban and mobility resilience assessments.

## 2.4 Transportation Diversity

Transportation diversity, or the richness and evenness of modes available in a city's transportation system (Rahimi-Golkhandan et al., 2019), is directly related to the latter's resilience. By drawing an analogy with ecology, researchers have proposed functional richness, the diversity and overall provision of modes of transport in a study area, and functional evenness, the evenness in the spatial distribution or use of those modes, which indicates equal access to transport for the population (Ren et al., 2022). Research has established transport diversity as an important factor in urban mobility resilience in the face of disasters (Shen et al., 2024). For example, following the 2017 Hurricane Harvey incident in Houston, researchers established that neighborhoods with greater transportation diversity suffered smaller mobility losses and also recovered faster, indicating better travel resilience (Wang et al., 2019). Transport diversity also sustains critical urban functions during emergencies (van der Gun et al., 2016). When Hurricane Sandy hit New York City, for instance, the temporary suspension of subway operations there spurred the collective use of alternative modes like shuttle buses and bicycles, highlighting the usefulness of diversified transport options as functional backups (Wang et al., 2019). In short, enhancing transport diversity by expanding multimodal transport infrastructure and making transport more inclusive and attractive across modes is central to strengthening the resilience of urban transport networks.

## 2.5 Data-Driven Approaches to Analyzing Mobility

Emerging technologies in the big data environment offer robust approaches to analyzing human mobility patterns before, during, and following catastrophic disasters. Historically, post-disaster mobility research was based on traffic surveys and evaluations. In recent years, the availability of large spatiotemporal trajectory data enables a more accurate representation of mobility patterns. Key data sources include GPS data (e.g., taxi or ride-hailing trajectories, mobile phone location data; (Wang et al., 2018; Kang et al., 2012; Barreras and Watts, 2024)), public transit smart card records (e.g., metro or bus entry-exit logs; (Cats, 2024; Zhao et al., 2016)), geotagged social media posts (e.g., Weibo, Twitter; (Gao et al., 2014; Zhu et al., 2024)), call detail records (CDRs; (Ayesha et al., 2021)), and integrated approaches that fuse multiple data sources (Rahimi-Golkhandan et al., 2021).

The synthesis of multiple data sources enhances travel analysis precision and completeness. For instance, blending smart card data and points of interest (POI) information allows researchers to infer travel intent and behavioral characteristics more easily (Yang et al., 2019). Additionally, integrating mobile signaling data, social media data, and traffic sensor data allows for exhaustive post-disaster traffic condition evaluations (Otsuka et al., 2016). Some researchers have put forward ecological information entropy approaches to monitor the disorder of cross-regional population migration and thereby evaluate the operational order of a city in disasters (Wang et al., 2019). Data-driven approaches tend to offer unprecedented fine-grained observations of advanced travel behavior in disaster situations. Through big data analysis, we can quantify objective changes in transportation system performance (e.g., road capacity reduction) (Hsieh and Feng, 2020; Aghababaei et al., 2021) and subjective changes in resident behavior (e.g., travel distance variation, activity scope, and mode choice variation) (Meister et al., 2022). This is crucial to formulate targeted emergency traffic management and recovery optimization. For example, if there is evidence that there is a broad activity range in some areas following the disaster, it would mean that varied modes of transportation enable people to move around (Rahimi-Golkhandan et al., 2021); however, if some segments take long to recover from journeys, more help would be needed (Deng et al., 2021). With improved IoT and cellular technology, future travel data will be more accurate and real-time, providing a reliable foundation to support urban resilience planning.

## 2.6 Research Gaps

Despite recent advances in understanding urban mobility resilience, several important gaps exist in the existing literature.

First, most existing research addresses resilience with single-mode transportation systems (Sun et al., 2020a; Yang et al., 2025). These analyses couldn't completely consider multi-modal interactions, overlooking the cascading effects of disturbance in one mode on other modes. Urban transportation networks, particularly in major cities such as New York City, are highly interconnected with complementary functions between subway systems, buses, taxis, bicycles, and pedestrian paths. This interdependency and complementarity are critical under disruption because users will change modes to provide the required mobility. Hence, current methods that neglect multi-modal integration present an incomplete picture of urban mobility resilience.

Second, most existing empirical research addresses resilience in single-mode transport networks (Donovan and Work, 2015; Haraguchi et al., 2022; Roy et al., 2019). Mobility is, however, dynamic in nature, with intricate temporal fluctuations during a disaster's unfolding and recovery stages. The static models fail to represent the dynamism of disruption and recovery and thus can't be employed as effective tools in real-time emergency response and urban planning. There is, therefore, a need for the formulation of dynamic, time-dependent models of resilience that effectively capture real-time variability in mobility and its effects during extreme events.

Third, transportation heterogeneity, that is, the variety and balance between transportation modes, has recently been recognized as an important factor for the resilience of urban mobility (Xu and Chopra, 2023; Rahimi-Golkhandan et al., 2021). While earlier studies have highlighted its importance, in-depth quantitative explorations explaining how exactly transportation diversity affects resilience outcomes remain limited. Rahimi-Golkhandan et al. (2021) demonstrated that diverse transportation infrastructure can enhance resilience, but the mechanisms through which diversity contributes to resilience, including specific pathways and interactions with other urban characteristics, remain under explored. Notably, limited research has specifically addressed how transportation diversity interacts with socioeconomic factors to affect mobility resilience outcomes. The complex interaction between transportation infrastructure attributes and socioeconomic disparities is probably an important determinant of a community's overall vulnerability and adaptive capacity during disruptions. However, the nature of these interactions and the extent to which they influence resilience remains poorly understood. Therefore, there is a significant knowledge gap relating to the moderating role of socioeconomic conditions on the effect of transportation diversity on urban resilience.

Furthermore, the impact of heterogeneity in transportation on resilience, and particularly on spatial heterogeneity, has been poorly investigated. Most current studies concentrate on overall city-level indicators, neglecting the localized differences arising from spatial variations in infrastructure provision, population density, and land use. Failing to capture spatial context adequately, current models risk making generalizations that mask important differences in resilience between neighborhoods or districts within cities.

In response to these research gaps, this study introduces a dynamic and geographically explicit framework that measures mobility resilience. It is founded on the integration of multimodal transport data, comprehensive metrics of trans-

portation diversity, and socioeconomic data. This study seeks to elaborate on not only the dynamic impact of transport diversity on resilience but also the extent to which socioeconomic characteristics influence such relationships across various geographic settings.

## 2.7 Objectives and Research Questions

The primary focus of this research lies in analyzing human mobility and the resilience of New York City's transportation system during and after Hurricane Ida. The event caused significant disruptions across different transportation modes, affecting millions of residents. Understanding how mobility was maintained or disrupted and how different areas and populations recovered from these disruptions is crucial for improving urban resilience. This leads to the following research questions:

**Research Question 1: How did Hurricane Ida impact different transportation modes in NYC, and which modes experienced the most significant disruptions?**

Answering this question will involve analyzing the hurricane's effect on each major mode of transport by measuring the drop in usage or service level for each. By identifying which transportation modes were hit hardest and which maintained service better, we can pinpoint the most vulnerable parts of the city's mobility system. This insight directly contributes to understanding urban mobility resilience by revealing the relative robustness of different transit options during an extreme weather event. In turn, it helps city planners prioritize which transportation infrastructures need reinforcement to withstand future disasters.

**Research Question 2: How did human mobility resilience and recovery time vary across different NYC neighborhoods or zones, and what role did transportation diversity play in mitigating these disruptions?**

Answering this question will examine the spatial variation in mobility resilience by comparing how much each area's overall transportation activity declined and how quickly it bounced back. This includes calculating combined resilience (across all modes) for each zone and the time taken to recover to normal levels. The analysis will also investigate whether areas with a more diverse mix of transportation options experienced smaller losses or faster recovery. Understanding these patterns shows where resilience was retained versus lost in the city and illustrates how transportation diversity can act as a buffer against disruptions. This knowl-

edge helps urban planners recognize which communities benefited from having multiple travel alternatives, reinforcing the idea that diverse transit networks can improve overall resilience to extreme events.

**Research Question 3: How do socio-economic, infrastructural, and environmental factors interact with transportation diversity to shape human mobility resilience in the face of Hurricane Ida?**

Answering this question will explore the interaction between transportation diversity and other key resilience factors in each city area. This means analyzing how attributes like population density, income levels, infrastructure quality, and environmental conditions combined with the availability of diverse transportation options influenced mobility outcomes during and after the hurricane. By examining these interactions, we can determine which combinations of factors had the greatest impact on a community's ability to maintain mobility and recover quickly. This contributes to a deeper understanding of urban mobility resilience by identifying critical drivers and barriers, for instance, if neighborhoods with strong infrastructure and high transit diversity fared much better or if socio-economic vulnerabilities amplified the impact despite multiple transit modes. Ultimately, the findings will highlight how holistic factors and a varied transportation network strengthen or weaken a city's resilience, guiding more effective strategies to protect human mobility during extreme weather events.

## 3 Study Area and Data

This chapter describes the study’s geographical coverage and data. The primary location of this study is NYC. NYC’s diverse transportation modes and exposure to extreme weather events, as evidenced by Hurricane Ida in 2021, make it a suitable case study. The data set contains travel records, transit system data, and population data. These resources enable the comprehension of how various urban regions respond to disruptions and how transport facilities differ in geographical settings.

### 3.1 Study Area

New York City provides an ideal case study for analyzing the relationship between transport diversity and resilience due to its dense urban environment, complex transport network, and past experiences with extreme weather events. The city consists of five boroughs: Manhattan, Brooklyn, Queens, Bronx, and Staten Island (Figure 3.1)<sup>3</sup>.



Figure 3.1: Study area: New York City’s five boroughs

<sup>3</sup> Source: [https://en.wikivoyage.org/wiki/New\\_York\\_City](https://en.wikivoyage.org/wiki/New_York_City), by PerryPlanet, under CC BY-SA 2.5.

## 3.2 Data

This work utilizes three primary datasets: (1) mobility data, (2) road network data, and (3) census data. These datasets facilitate end-to-end human mobility resilience analysis by incorporating transportation infrastructure and demographic characteristics.

### 3.2.1 Mobility Data

Mobility data captures real-world travel behavior before, during, and after disruptions. The details are listed in Table 3.1.

Table 3.1: Description of mobility datasets

Dataset	Description	Temporal Resolution	Spatial Resolution	Source
<b>Subway</b>	Provides counts of subway entries and exits at each station across NYC.	4-hour intervals	Station-based	MTA Open Data
<b>Taxi</b>	Trip records from NYC's Taxi and Limousine Commission (TLC), covering yellow taxis, green taxis, for-hire vehicles, and high-volume for-hire vehicles.	Minute-level	Zone-based	NYC TLC
<b>Citi Bike</b>	Shared bike trip records from NYC's Citi Bike program, capturing rental and return details.	Minute-level	Station-based	Citi Bike

### 3.2.2 Road Network

As shown in Figure 3.2, the road network data utilized for this research comprises the subway system (subway lines and stations), bike lanes, street network, and bus system (bus lines and stops) of New York City. The Metropolitan Transportation Authority (MTA), New York City Department of Transportation (DOT), and NYC Open Data provided the data. The data is mainly utilized to quantify transportation diversity or the presence of alternative modes of travel in case of service disruption.

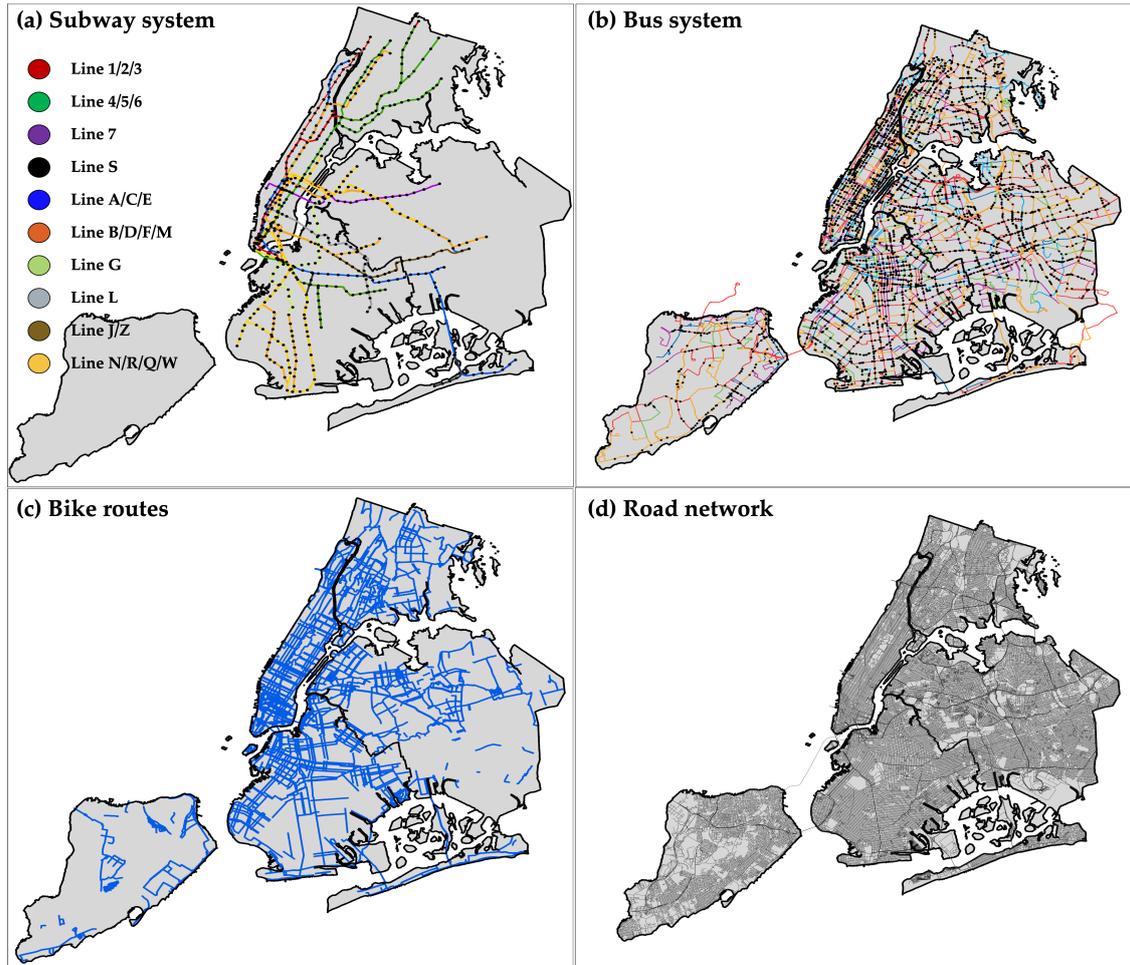


Figure 3.2: Spatial distribution of transportation infrastructures in New York City. (a) Subway system colored by official MTA line designations. (b) Bus system rendered using MTA bus map colors. (c) Bike routes. (d) Road network.

### 3.2.3 Census Data

As shown in Table 3.2, we utilized publicly available socio-spatial datasets, including the U.S. Census Bureau's American Community Survey (ACS) 5-Year Estimates (2022) at the census tract level and the PLUTO dataset from the New York City Department of City Planning (DCP). These datasets provide complementary demographic and land use information for spatial analysis.

Table 3.2: Descriptions of socio-economic and environmental data

Variable Code / Name	Description
<b>U.S. Census Bureau's American Community Survey</b>	
B19083_001E	Gini Index of Income Inequality
B01003_001E	Total Population
B15003_022E	Number of individuals aged 25 and over with a Bachelor's degree
B27010_017E	Number of individuals aged 19 to 64 without health insurance coverage
B23025_005E	Number of unemployed individuals aged 16 and over
B23025_002E	Total labor force aged 16 and over
B19013_001E	Median Household Income in the past 12 months
<b>Primary Land Use Tax Lot Output (PLUTO)</b>	
firecomp	Fire company identifier serving the tax lot's location.
healthcenterdistrict	Health center district number associated with the tax lot.
healtharea	Health area code related to the tax lot's location.
landuse	Numeric code representing the primary land use of the tax lot (e.g., residential, commercial, industrial).
policeprct	Police precinct number serving the tax lot's location.
resarea	Residential area in square feet within the tax lot.
comarea	Commercial area in square feet within the tax lot.
factoryarea	Factory (industrial) area in square feet within the tax lot.
otherarea	Other areas in square feet within the tax lot, not classified as residential, commercial, or industrial.

# 4 Methodology

As shown in Figure 4.1. This study adopts a streamlined methodological approach that integrates multiple data sources and analytical techniques. First, multi-modal mobility data (from taxis, subways, and bike-sharing), road network information, and socioeconomic indicators are collected and preprocessed to a common geographic scale. Mobility resilience is then quantified by comparing observed traffic volumes during Hurricane Ida against a forecasted baseline, with key metrics such as resilience and recovery time computed to capture the extent and duration of disruptions. In parallel, transportation diversity is assessed through an index that combines the availability and spatial distribution of various transit modes. Finally, regression and spatial analysis methods are applied to examine the relationships between transportation diversity, socioeconomic factors, and overall mobility resilience in New York City.

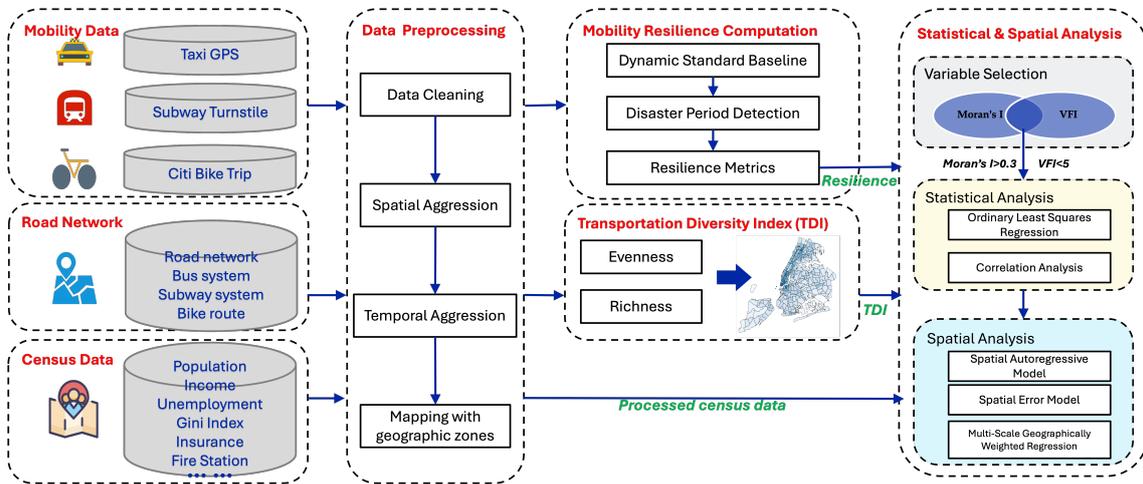


Figure 4.1: Research framework.

## 4.1 Data Preprocessing

Data preprocessing forms the foundation of this study and involves key steps: Data Cleaning, Data Aggregation, Spatial Integration, Data Transformation, and

Geographic Mapping. These steps ensure that multi-source data are standardized at a common geographic scale for calculating Mobility Resilience Metrics (including resilience, resilience loss, recovery time, start time, end time, max deviation, and the ratio of resilience-to-resilience loss), the Transportation Diversity Index (TDI), and processed census data for subsequent statistical analysis.

## **4.1.1 Data Cleaning**

### **4.1.1.1 Mobility Data Cleaning**

To ensure consistency and accuracy in mobility data, the following preprocessing steps are applied across all transportation modes:

- Extract key attributes: (start time, end time, start location, end location).
- Standardize time format for consistency.
- Remove duplicate records and entries with missing critical attributes.

### **Taxi Data Processing**

Taxi trip data includes multiple service types, including Yellow Taxi, Green Taxi, For-Hire Vehicles (FHV), and High-Volume For-Hire Vehicles (HVFHV) such as Uber and Lyft. To refine this dataset, outliers are removed by excluding trips with extremely short (<100m) or long distances (>50km) and those with unrealistically long durations exceeding three hours. Records missing start or end locations are also discarded. The final taxi trip volume is computed as the sum of Yellow Taxi, Green Taxi, FHV, and HVFHV trip counts, providing a consolidated metric of total taxi activity.

### **Subway Turnstile Data Processing**

Subway turnstile data captures public transit usage at different locations and is recorded at 4-hour intervals. Given its structure, preprocessing focuses on ensuring reliability and consistency. Outliers are identified by flagging entries where the number of recorded turnstile entries or exits exceeds five times the historical mean. Stations with persistently abnormal values or negative counts, likely due to system resets, are excluded. Missing values are addressed using interpolation: if a single period is missing, the average of adjacent periods is used; if data is absent for three or more consecutive periods, the station is excluded due to reliability concerns.

Timestamps are standardized to align with the official 4-hour recording schedule (00:00, 04:00, 08:00, 12:00, 16:00, 20:00), with adjustments made for daylight saving time shifts. Any misaligned timestamps (e.g., 03:45 or 07:30) are corrected to the nearest valid interval. Finally, net flow is calculated for each station by computing the difference between current and previous entry and exit values, yielding a measure of total flow per station.

### **Citi Bike Data Processing**

Citi Bike data, representing shared micromobility usage, requires preprocessing for spatial and temporal consistency. Outlier removal includes filtering out trips lasting **less than one minute or exceeding 24 hours**, as such cases are likely data errors. Additionally, instances where a bike is returned to the same station within seconds, suggesting unintended rentals or system malfunctions, are excluded.

#### **4.1.1.2 Census Data Cleaning**

Several derived variables were constructed from raw census and land use data to support the spatial analysis of mobility resilience. Population density (POP\_DENSIT) was calculated by dividing the total population (from ACS) by tract-level land area aggregated from PLUTO lots.

Socioeconomic indicators were also derived from the ACS. These include the percentage of residents with a bachelor's degree (BACH\_PCT), unemployment rate (UNEMP\_PCT), and percentage without health insurance (NO\_INS\_PCT), each calculated as ratios relative to total population or labor force.

Land use composition (RES\_PCT, COM\_PCT, FACT\_PCT, OTHER\_PCT) was computed from PLUTO by aggregating the area of different land use types within each tract and expressing them as percentages of total land area.

Lastly, counts of public service units, police precincts (POLICE), health-related stations (HEALTHAR), and fire companies (FIRE) were calculated per tract.

## **4.1.2 Data Aggregation**

### **4.1.2.1 Spatial Aggregation**

#### **(1) Assigning Geographic Coordinates**

Citi Bike and taxi trip datasets include precise latitude and longitude coordinates

for trip origins and destinations. These raw coordinate values were converted into spatial point data, enabling spatial joins with NYC taxi zones.

In contrast, subway turnstile data lacks direct latitude/longitude values and instead provides only station IDs. To assign geographic locations, a reference table containing station IDs and their corresponding coordinates was used to match each record to its respective spatial location.

## (2) Converting to a Unified Coordinate System

All spatial features were transformed into a common coordinate reference system (CRS), New York State Plane, to ensure spatial compatibility across datasets.

### 4.1.2.2 Temporal Aggregation

All transportation data were aggregated into standardized time intervals based on their respective recording structures to ensure comparability across datasets.

**Taxi and Citi Bike data** record trip details with precise timestamps for trip start and end times. These datasets were aggregated into **hourly intervals**, where the total number of trips per taxi zone per hour was computed. This hourly aggregation enabled fine-grained temporal analysis of mobility patterns.

**Subway turnstile data** follows a different temporal structure, as entries and exits are recorded at 4-hour intervals according to the NYC subway schedule. To maintain data accuracy and consistency with the original reporting structure, subway turnstile counts were retained in **4-hour intervals** rather than being interpolated into smaller time units.

## 4.1.3 Mapping with Geographic Zones

Taxi zones were used as the spatial unit of analysis. As predefined administrative units commonly used in ride-hailing data, they offer a balance between spatial granularity and computational efficiency, finer than boroughs but coarser than census tracts.

**Mobility data**, including Citi Bike trips, taxi rides, and subway entries, were mapped to taxi zones by spatially joining trip origin and destination coordinates to zone boundaries.

**Census data**, originally at the census tract level, were aggregated to taxi zones.

Each tract was assigned to the taxi zone containing its centroid. For count variables (like population), values were summed; for rate variables (like unemployment rate), averages were calculated to maintain consistency.

**PLUTO data** were also aggregated to taxi zones. The number of public facilities (fire, police, and health) in each zone was counted as a measure of service accessibility. Land use composition was calculated by summing the area of each land use type and expressing it as a percentage of the total zone area.

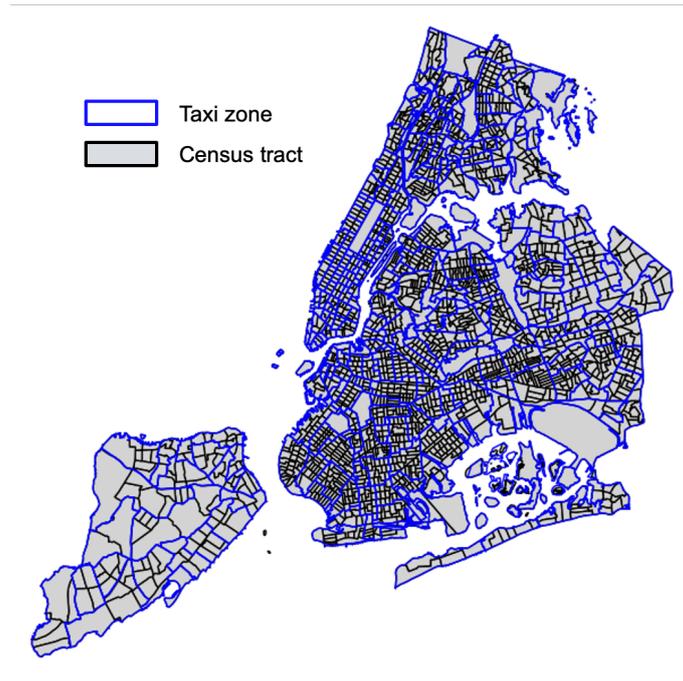


Figure 4.2: Spatial overlay of NYC Taxi Zones (blue) and Census Tracts (gray).

#### 4.1.4 Data Transformation

Data transformation is a critical preprocessing step in statistical and spatial analysis, intended to improve data quality and meet the assumptions of many modeling techniques. Real-world datasets often contain skewed distributions, heteroscedasticity, and extreme outliers, which can undermine the validity of regression and spatial models. To address these challenges, three transformation methods were applied in this study.

First, the Box-Cox transformation was used to normalize positively skewed variables with non-negative values, improving distributional symmetry and variance stability. Second, for variables that include zero or negative values, the Yeo-Johnson transformation was selected as a flexible alternative that accommodates a broader range of data while reducing skewness. Lastly, Winsorization was ap-

plied to cap extreme values at specified percentiles, reducing the influence of outliers without removing observations.

## 4.2 Mobility Resilience Computation

The process consists of three key steps: establishing a baseline mobility pattern, detecting disruptions as deviations from that dynamic baseline, and quantifying resilience through metrics that capture the impact and recovery duration for each mode. By following these steps for each transportation mode (taxi, subway, Citi Bike), we can assess how well each system withstands shocks and how quickly they rebound, thereby comparing their adaptive capacities.

### 4.2.1 Conception of Mobility Resilience Model

#### 4.2.1.1 Overview of the Proposed Model

The Mobility Resilience Model measures the degree to which urban transportation systems sustain and recover human mobility under disruption, rather than measuring physical infrastructure integrity. Rather than quantifying resilience in terms of structural damage or time out of service, it examines the operation of mobility i.e., how individuals persist in traveling within the city by road, transit, or otherwise, during an emergency. In this research, this model prioritizes that the provision of transportation service is crucial to facilitate emergency response and everyday necessities.

In this system, mobility resilience is understood as the ability of the system to recover, use redundancy, and adapt in the event of shocks. In practice:

**Adaptability** is the extent to which the transportation system or travelers can adapt to new conditions. For example, if a subway station or line is flooded, passengers and the transit system can reroute trips, reschedule, or shift to other transportation to keep people on the move. This adaptive capability is the capacity of the system to change or reshape operations during periods of stress.

**Redundancy** includes the existence of several transportation alternatives or standby systems. A strong mobility system possesses substitute routes or modes (roads, transit lines, cycling facilities, etc.) that can act as alternatives when part of the system is not functioning. Having different modes gives us valuable flexibility in

times of crisis by enabling different groups to continue mobility by converting to available modes. Effectively, redundancy means there is no failure point – when one mode or path is damaged, and others can meet the demand.

**Recovery** is the capacity of the mobility system to resume normal operation quickly after the disruption. That is, resuming transit service, restoring highway clearance, or otherwise having pre-disaster levels of travel capacity restored in a short time. High recoverability implies that mobility loss is temporary only, and the system quickly returns to functioning with minimal long-term effect on travelers.

Overall, the Mobility Resilience Model is concerned with service-level continuity (are destinations reachable by people?). It measures the resilience of the transportation system to endure the shock, remain operational (possibly in partial mode), and recover full mobility. This reflects the system's human-centered performance. Resilience, in terms of how well people can still move around, matters because mobility directly impacts how quickly a city's social and economic life can rebound.

#### 4.2.1.2 Transition from Fixed to Dynamic Standard Baseline

Conventional resilience models, like the Bruneau model (Figure 4.3), normally employ a fixed baseline of system performance to measure disruption and recovery. In Bruneau et al.'s prominent seismic resilience model, performance is graphed over time versus a fixed "fully functioning" benchmark level, creating the familiar "resilience triangle" that assesses loss and recovery via area under the curve. This is suitable for systems with comparatively fixed operational baselines, e.g., critical infrastructure.

But when extended to urban mobility systems, which are inherently dynamic in nature, the static baseline assumption can restrict its accuracy of representation. Traffic demand and transport use differ considerably by time of day, day of the week, or season. The Bruneau model does thus offer a useful conceptual model, but it must be modified to capture the temporal dynamics and behavioral heterogeneity of mobility systems. The addition of time-varying or flexible baselines can enhance the model's applicability to transportation resilience evaluation in cities.

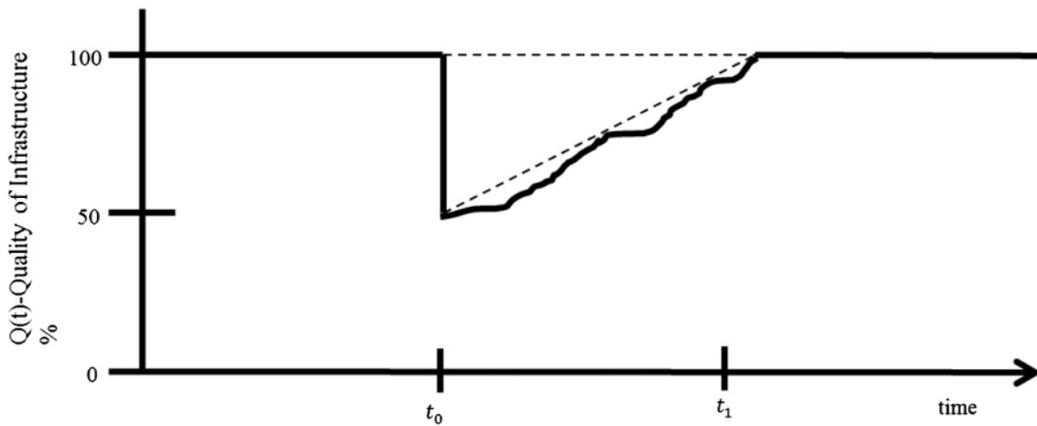


Figure 4.3: Resilience Triangle (adapted from Bruneau et al. (2003)).

The limitation of a fixed-baseline approach lies in its inability to capture the inherent temporal variability of urban mobility patterns. Urban traffic follows predictable rhythms, with pronounced peaks during rush hours and lower volumes at night or on weekends. Using a single static reference value, such as an average or maximum traffic flow, as the “normal” baseline may result in misinterpretation of normal fluctuations as disruptions or recovery. For instance, a baseline of 100,000 passengers per hour might misleadingly indicate a severe performance drop during a 2 AM disruption, even though near-zero ridership is typical at that time. Conversely, during peak demand periods (such as Monday at 8 AM), the same baseline may underestimate actual needs, making a significant reduction appear minor. In short, static baselines overlook the cyclical nature of traffic patterns.

As highlighted in previous studies, a single “standard” traffic volume is insufficient for capturing real-world mobility dynamics since passenger flows follow clear temporal patterns, like diurnal or weekly cycles. Relying on a fixed-value baseline may, therefore, obscure the true impact of disruption by conflating ordinary variation with actual system failure or recovery.

#### 4.2.1.3 Refinements in Mobility Resilience Model

To combat the limitations of a fixed baseline in contrast to the dynamic nature of urban areas, the improved **Mobility Resilience Model** (Figure 4.4) embraces a dynamic standard baseline that varies over time. In place of one set point of reference, the technique sets the baseline to the expected level of mobility at any moment in time with the assumption that there are no disruptions. This dy-

dynamic baseline is built using time-series prediction of historical mobility data, picking up on recurring patterns such as rush-hour peaks, late-night troughs, weekday-weekend fluctuation, and seasonal variation.

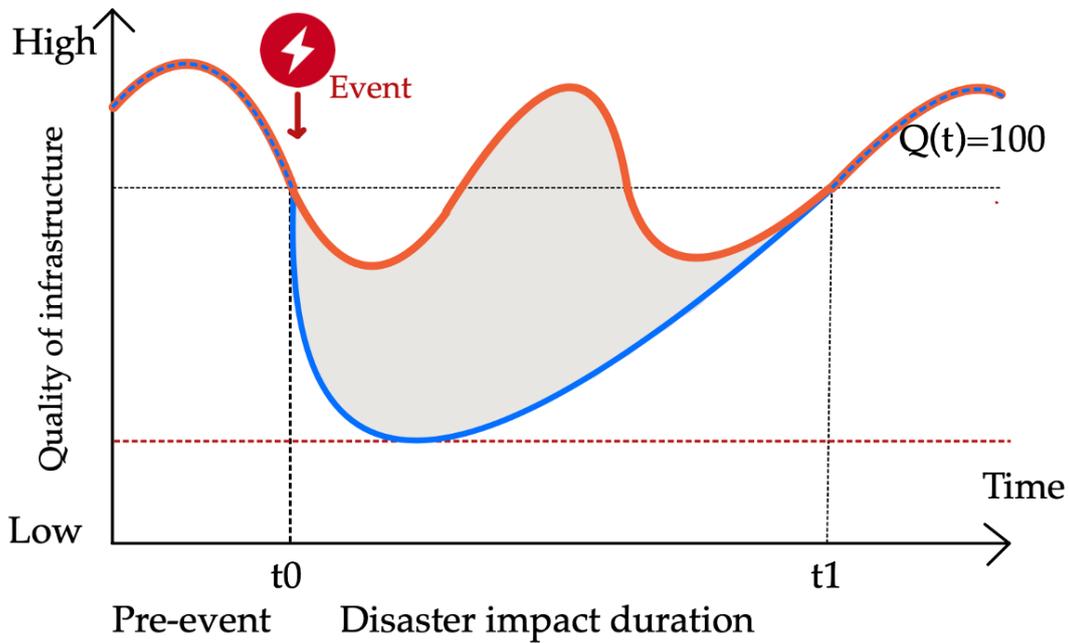


Figure 4.4: Improved mobility resilience model: Disaster impact period ( $t_0-t_1$ ) with observed mobility (blue) and predicted dynamic baseline (orange).

By comparing observed mobility during a disruption with the corresponding time-specific expected value, the model enables a more accurate assessment of resilience. Deviations from this dynamic baseline can be attributed to the disruption itself rather than being confounded by regular fluctuations. The updated model, therefore, strengthens classic models' theoretical underpinnings, like Bruneau's, and extends their applicability to temporal dynamics in urban mobility systems. Incorporating a time-variant reference point renders the model more efficient at precisely modeling the extent and timing of disruption effects in human-centered mobility contexts.

Table 4.1: Comparison between the Bruneau model and the improved resilience model

Aspect	Bruneau Model	Improved Model
Baseline Type	Fixed value	Dynamic baseline (time-series forecasting)
Adaptability to Variations	Low – assumes constant traffic patterns	High – captures fluctuations (rush hours, weekends, etc.)
Handling Seasonal Trends	Cannot account for daily variations	Handles seasonal and contextual variations
Response to Disruptions	Assumes resilience loss is proportional to physical damage	Measures resilience based on real-world mobility losses and recovery time
Outlier Handling	May misinterpret normal traffic variations as disruptions	High – requires data-driven modeling, historical analysis, and machine learning techniques for forecasting
Computational Complexity	Low – simple resilience calculations using predefined baseline	High – requires data-driven modeling, historical analysis, and machine learning techniques for forecasting
Data Requirements	Minimal – relies on basic infrastructure performance metrics	High – requires large datasets (historical mobility records, real-time transit data, traffic sensors)

## 4.2.2 Improved Dynamic Baseline

To assess human mobility resilience, a baseline mobility pattern that represents normal conditions in the absence of disruptions was established. This baseline serves as a reference against which deviations (caused by disruptions) can be detected. Given that mobility data exhibits strong temporal trends and seasonality, this study used Prophet, a robust time-series forecasting model, to construct an adaptive baseline.

### 4.2.2.1 Prophet Model Configuration

Prophet is an open-source forecasting tool developed by Facebook, designed to handle time-series data with strong seasonal effects, holidays, and trend changes. The model consists of three main components:

**Trend ( $g(t)$ ):** Captures long-term mobility changes, such as gradual increases or decreases in taxi or bike trips due to policy shifts, infrastructure development, or evolving travel behaviors.

**Seasonality ( $s(t)$ ):** Models periodic fluctuations, including daily commuting cycles and weekly variations.

**Holiday Effects ( $h(t)$ ):** Accounts for abrupt changes in mobility on special days, such as holidays, extreme weather events, or city-wide events.

The general formulation of the Prophet model is:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (4.1)$$

where  $\epsilon_t$  represents the error term accounting for noise in the data.

#### 4.2.2.2 Implement details

To accurately model mobility patterns, seasonality adjustments are applied within the time-series forecasting model, and the detailed implements are listed in Table 4.2. Daily seasonality ensures that recurring daily travel fluctuations are captured, while weekly seasonality accounts for differences between weekday and weekend travel behavior. Since the focus is on short-term mobility trends, yearly seasonality is disabled to avoid unnecessary complexity.

Table 4.2: Seasonality configurations in the Prophet forecasting model

Seasonality Component	Configuration & Purpose
<b>Daily</b>	<code>daily_seasonality=True</code> – Captures recurring daily travel patterns.
<b>Weekly</b>	<code>weekly_seasonality=True</code> – Accounts for differences between weekdays and weekends.
<b>Yearly</b>	<code>yearly_seasonality=False</code> – Annual cycles are disabled since the focus is on short-term mobility trends.
<b>Morning Peak</b>	<code>period=12, fourier_order=3</code> – Captures morning commute demand.
<b>Evening Peak</b>	<code>period=24, fourier_order=3</code> – Models evening rush-hour demand.
<b>Weekend Effects</b>	<code>add_regressor('is_weekend')</code> – Introduces a binary variable to differentiate weekend and weekday travel patterns.
<b>Uncertainty Interval</b>	<code>interval_width=0.95</code> – Sets a 95% confidence interval, allowing detection of significant deviations from normal patterns.

#### 4.2.2.3 Data Preparation & Model Training

The training period considered to train the model is from January 1st, 2021, to July 31st, 2021, thereby encompassing an entire year of previous mobility patterns. The specified time period enables the model to incorporate seasonal fluctuations, such as daily and weekly commute behavior, along with any exceptional events like holidays or anomalous activities.

At model training time, the taxi, subway, and Citi Bike ridership data are pre-processed to match their respective time intervals. Independent Prophet models are constructed for each transportation mode, with hyperparameter tuning to balance seasonality adjustment and changepoint sensitivity. The trained models thereby generated offer a dynamic baseline, forecasting anticipated ridership for every timestamp while accounting for long-term trend, seasonal, and contextually pertinent factors.

To find a feasible standard baseline for mobility at any time  $t$ , a predicted volume

is created along with a 95% confidence interval. In this way, it is guaranteed that any variation from normal trends has statistical significance. The dynamic baseline at the specific time  $t$  is defined as follows:

$$\text{standard volume} = \hat{y}(t) \pm CI_{95\%} \quad (4.2)$$

where  $\hat{y}(t)$  represents the predicted mobility at time  $t$ ,  $CI_{95\%}$  is the 95% confidence interval from the model.

### 4.2.3 Determine the Disaster Impact Period

After the dynamic baseline is established through the Prophet model, the second step is to flag episodes during which the observed mobility volume differs greatly from the expected volume. This is done by disruption detection based on prespecified threshold and duration parameters to guarantee that only large, disaster-specified reductions in mobility are indicated.

#### 4.2.3.1 Extreme Event Detection Criteria

To systematically identify major disruptions while filtering out minor fluctuations:

**95% Threshold:** An anomaly is flagged when actual mobility falls below the lower bound or rises above the upper bound of Prophet's 95% confidence interval. This ensures only significant drops are considered, as values outside this range indicate deviations beyond normal variability. Using this threshold balances sensitivity and specificity, excluding routine daily peaks while capturing major deviations caused by the disaster.

**Minimum Duration (2-Hour Rule):** A disruption must persist for at least two consecutive hours to be counted. This avoids flagging short-term variations, like brief rush-hour congestion or minor delays, that do not reflect sustained impacts. Requiring a 2-hour duration filters out transient noise, like brief accidents or hiccups, and focuses on meaningful, continuous disruptions.

#### 4.2.3.2 Defining Start and End Times

To ensure the analysis focuses solely on the impact of Hurricane Ida, specific temporal rules are applied:

**Start Time ( $t_0$ )** Detection is restricted to the hurricane's timeframe, between September 1st and September 3rd, 2021. In other words, anomalies are only considered if they occur during the known period of Hurricane Ida's impact. Any anomalous dips outside this window are treated as unrelated or minor fluctuations not caused by the hurricane.

**End Time (24-Hour Monitoring) ( $t_1$ )** The disruption period is considered over if no additional anomalies occur within 24 hours after the last detected drop. In practice, once a disruption is identified, the model continues to monitor for another day; if no new significant dips occur in that span, the event is closed out. This 24-hour rule captures the full duration of the hurricane's impact while excluding intermittent post-event wobbles, thereby focusing on the sustained disruption rather than sporadic aftershocks.

#### 4.2.4 Resilience Quantification

Once disruptions are identified, mobility resilience is quantified by measuring how quickly and effectively mobility recovers. The core method involves first calculating the Mobility Quality:

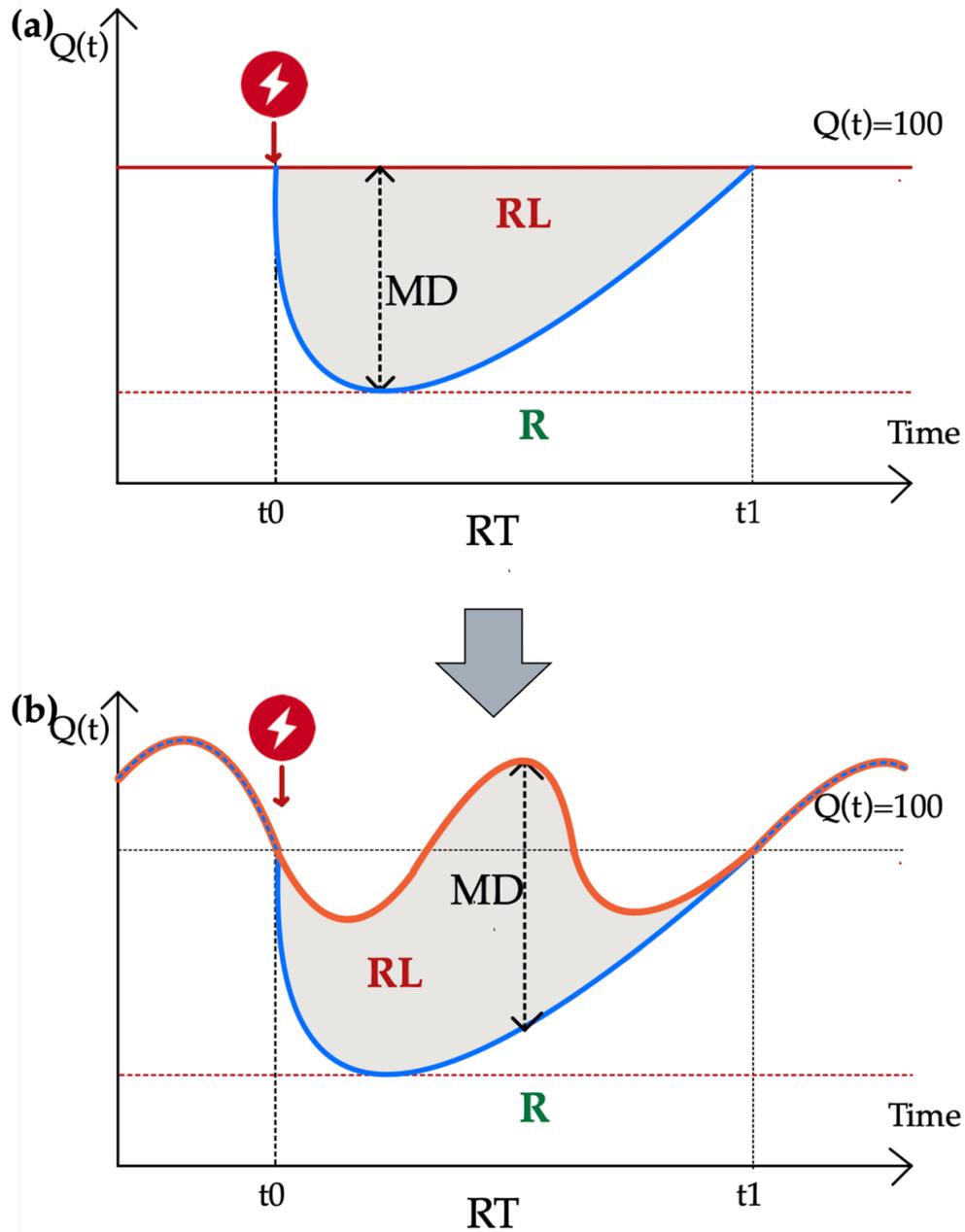


Figure 4.5: Comparison between (a) the traditional fixed-baseline model and (b) the proposed dynamic-baseline model.

Mobility resilience is assessed by comparing actual mobility against the expected standard mobility at each time step  $t$ . This is expressed through the Mobility Quality Function ( $Q_t$ ), which quantifies mobility performance as a percentage of normal conditions:

$$Q_t = \frac{\text{actual volume } (t)}{\text{standard volume } (t)} \times 100 \quad (4.3)$$

#### 4.2.4.1 Resilience Loss (*RL*)

Resilience loss measures the total reduction in mobility over the disruption period, representing the cumulative impact of the disturbance. A higher RL value means greater loss, either due to a longer disruption duration or a deeper mobility drop. This metric captures the cumulative impact of the disruption.

$$RL = \int_{t_0}^{t_1} [100 - Q_t] dt \quad (4.4)$$

where  $t_0$  is the start time of the disruption,  $t_1$  is the end time when mobility has recovered,  $100 - Q_t$  represents the percentage of lost mobility at each moment.

#### 4.2.4.2 Resilience (*R*)

Resilience represents the system's ability to maintain mobility relative to the lost mobility. A higher R means the system retained higher functionality throughout the disruption. A lower R suggests that mobility remained severely impacted for an extended period.

$$R = \int_{t_0}^{t_1} Q_t dt \quad (4.5)$$

#### 4.2.4.3 Resilience Ratio (*Ratio*)

To evaluate recovery efficiency, we compute the resilience ratio:

$$Ratio = \frac{RL}{R} \quad (4.6)$$

where:

- $Ratio < 1$ : The system recovered effectively, minimizing the impact.
- $Ratio \approx 1$ : The system had balanced resilience and loss.
- $Ratio > 1$ : The system struggled to recover, suffering prolonged disruption.

#### 4.2.4.4 Other Metrics

- Maximum Deviation (MD): The worst mobility drops relative to normal

$$MD = \max_{t \in [t_0, t_1]} (100 - Q_t) \quad (4.7)$$

- Recovery Time (RT): The time required for mobility to return to 95% of normal

$$RT = t_1 - t_0 \quad (4.8)$$

All in all, Resilience Loss (RL) quantifies the total mobility deficit over time, capturing the cumulative reduction in transportation availability throughout the disruption period. In contrast, Resilience (R) represents the total maintained mobility, reflecting the extent to which the transportation system remained functional despite adverse conditions. To assess recovery efficiency, the Resilience Ratio (Ratio) measures how effectively and quickly the system restores normal operations. Additionally, Maximum Deviation (MD) indicates the peak impact of the disruption, identifying the largest drop in mobility relative to the expected baseline. Finally, Recovery Time (RT) evaluates the speed of recovery, defining the duration required for the transportation network to return to pre-disruption performance levels. These metrics collectively provide a comprehensive understanding of mobility resilience, allowing for a nuanced assessment of both the severity of disruptions and the system's ability to recover efficiently.

### 4.3 Transportation Diversity Index

The Transportation Diversity Index (TDI) measures the diversity in mobility alternatives at various geographic places. A better TDI result implies a diversely balanced transport system that renders cities more mobile and resilient. It is calculated by examining the options for transport provision and their corresponding proportions according to two broad dimensions: Richness and Evenness.

#### 4.3.1 Multi-modal Transportation Availability

Before calculating the TDI, we assess the availability and spatial distribution of major transportation modes across geographic units. The modes considered include the bus system (lines and stations), subway system (lines and stations), bike

infrastructure (dedicated lanes), and the road network.

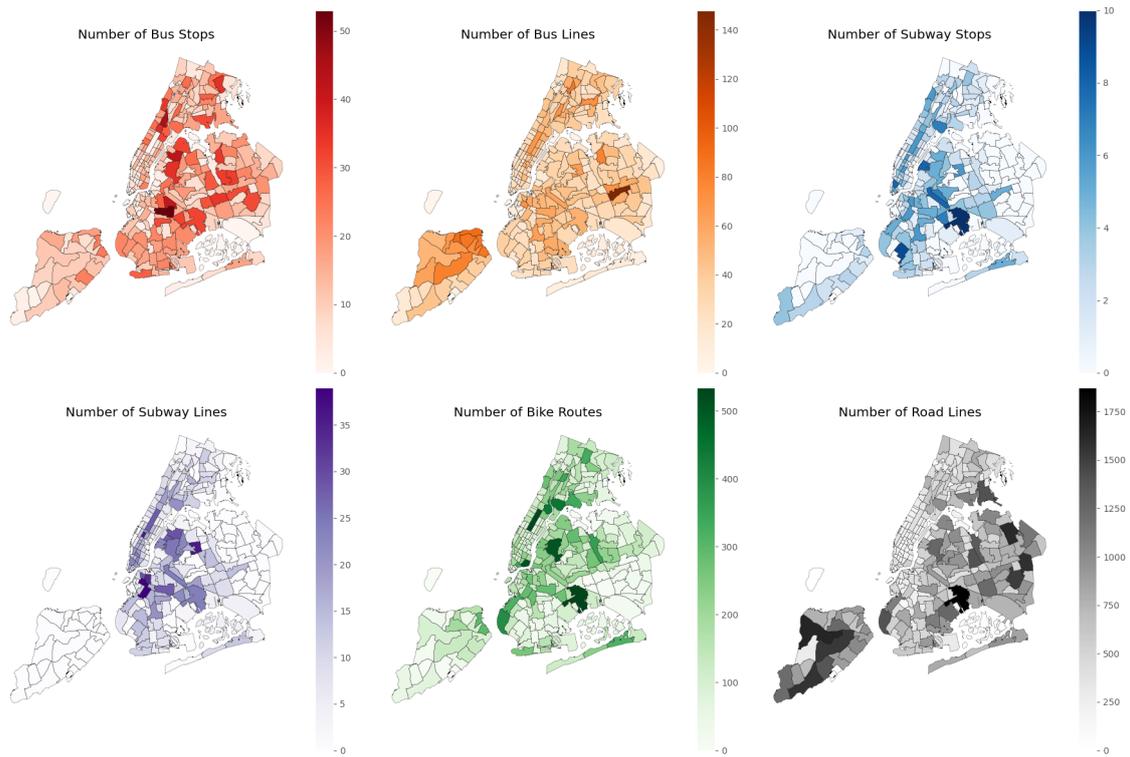


Figure 4.6: Heatmaps of bus stops, bus lines, subway stations, subway lines, bike routes, and road infrastructure in NYC.

### 4.3.2 Richness & Evenness

Transport diversity is quantified based on two important measures, i.e., functional richness and functional evenness (Rahimi-Golkhandan et al., 2019), and these are used for all modes of transport, like road networks, buses, subway, cycling routes, and pedestrian paths. Functional richness is an indicator of the presence of each mode of transport in a region. For example, road network richness is calculated as roadway mileage total divided by zone area, and bus and subway transit richness is calculated based on stops and lines per unit area. Bicycle route and walkway richness are calculated similarly by dividing the total route mileage by the zone area. Functional evenness, in contrast, quantifies the spatial distribution of the transport modes. It is estimated from the standard deviation (SD) of the minimum distance between research spatial units (taxizones) centroids and transport facilities, divided by the zone area. A lower SD indicates more evenly distributed transport services, i.e., greater functional evenness.

Table 4.3: Metrics of traffic mode diversity.

Mode	Richness	Evenness
Road Network (RN)	$R_{RN} = \frac{\text{total roadway mileage}}{A}$	$E_{RN} = \left( \frac{\sigma_{d_{RN}}}{\sqrt{A}} \right)$
Bus System (BS)	$R_{BS} = \left( \frac{n_s}{N_S^T \cdot A} \right) \left( \frac{n_l}{N_L^T} \right)$	$E_{BS} = \left( \frac{\sigma_{d_{BS}}}{\sqrt{A}} \right)$
Subway System (ST)	$R_{ST} = \left( \frac{n_s}{N_S^T \cdot A} \right) \left( \frac{n_l}{N_L^T} \right)$	$E_{ST} = \left( \frac{\sigma_{d_{ST}}}{\sqrt{A}} \right)$
Bicycle Routes (BR)	$R_{BR} = \frac{\text{total bicycle route mileage}}{A}$	$E_{BR} = \left( \frac{\sigma_{d_{BR}}}{\sqrt{A}} \right)$

Note:  $R$  is richness;  $E$  is evenness;  $n_s$ : number of stops/stations;  $N_S^T$ : total number of stops;  $n_l$ : number of lines;  $N_L^T$ : total number of lines;  $\sigma_d$ : standard deviation of the shortest distance from zone centers to mode facilities;  $A$ : area of the taxi zone.

### 4.3.3 Transportation Diversity Index

To quantify the resilience-enhancing role of transportation diversity, the Transportation Diversity Index (TDI)<sup>4</sup> is introduced as a holistic measure that integrates both the Richness and Evenness of available transportation modes. The motivation behind using TDI is to provide a comprehensive indicator of how well a transportation system can adapt to disruptions by offering multiple, balanced mobility options. A system with high TDI is expected to be more resilient, as users can switch between different modes when one is affected, reducing the severity of disruptions.

$$TDI = \sum_{modes} w_{mode} \cdot (R_{normalized} + E_{normalized}) \quad (4.9)$$

where  $w_{mode}$  are weights assigned to each transportation mod;  $R_{normalized}$  and  $E_{normalized}$  are the scaled values for Richness and Evenness, ensuring that each factor contributes equally to the final TDI score.

In this study, each transportation mode (roads, bikes, subways, and buses) is assigned an equal weight (0.25) to maintain fairness and independence in the Transportation Diversity Index (TDI). This approach avoids bias introduced by weighting modes based on usage volume, which would otherwise blur the line

<sup>4</sup> Note: Transportation diversity is calculated based on infrastructure data (bus, subway, bike routes, and roads), while resilience metrics are derived from observed mobility data (taxi, subway, and Citi Bike usage). The two serve different purposes—diversity reflects structural availability, whereas resilience captures functional performance under disruption.

between transportation diversity and resilience. By treating each mode as equally important, TDI provides a neutral, objective measure of transportation diversity without disproportionately emphasizing or diminishing any particular mobility option.

## 4.4 Statistical and Spatial Analysis

This section provides the step-by-step procedure of statistical and spatial analysis at the level of taxi zones in New York City to analyze the interactions between transportation diversity, socio-economic attributes, and mobility resilience. The analysis process is segmented into different steps. In the initial step, decisive variables are selected by analyzing spatial autocorrelation and multi-collinearity such that spatial relations are taken into account, and similar variables are not included. Subsequently, conventional statistical techniques are used: Ordinary Least Squares (OLS) regression describes global relations between independent variables, while correlation analysis investigates pair-wise relations between the variables. Lastly, the Spatial Autoregressive (SAR) model and the Spatial Error Model (SEM) are applied to account for spatial dependence and autocorrelation in the model residuals, thereby improving the accuracy and reliability of the regression results. Moreover, Geographically Weighted Regression (GWR) is utilized with a view to describing spatially varying relationships.

Table 4.4 presents a list of candidate variables, independent and dependent, that span multiple dimensions of transportation resilience.

Table 4.4: Variable definitions and units for resilience modeling

Variable	Description	Unit
<b>Potential Dependent Variables</b>		
<b>r</b>	Resilience	–
<b>rl</b>	Resilience loss	–
<b>mdev</b>	Max deviation	–
<b>rt</b>	Recovery time	Days
<b>ratio</b>	Ratio	–
<b>Potential Independent Variables</b>		
<b>tdi</b>	Transportation Diversity Index	–
<b>pop</b>	Total number of people	People
<b>pop_density</b>	Population density	People/km <sup>2</sup>
<b>gini</b>	Gini index (income inequality)	–
<b>bach_pct</b>	Percent with bachelor's degree or higher	%
<b>no_ins_pct</b>	Percent without health insurance	%
<b>unemp_pct</b>	Unemployment rate	%
<b>income</b>	Median household or per capita income	USD
<b>res_pct</b>	Percentage of land for residential use	%
<b>com_pct</b>	Percentage of land for commercial use	%
<b>fact_pct</b>	Percentage of land for industrial use	%
<b>other_pct</b>	Percentage of land for other uses	%
<b>police</b>	Number of police facilities	Count
<b>fire</b>	Number of fire stations	Count
<b>healthar</b>	Number of healthcare facilities	Count

#### 4.4.1 Variables Selection

Before developing regression models, two diagnostic procedures were used to guide variable selection: (1) a test for spatial autocorrelation in the data and (2) checking for multicollinearity of independent variables. These steps ensure that spatial structure is being detected and that chosen predictors are not highly collinear.

#### 4.4.1.1 Moran's I for Spatial Autocorrelation

Moran's I statistic was calculated to detect the presence of spatial autocorrelation in the input values across taxi zones. Moran's I is a global measure that evaluates whether similar values cluster spatially more than would be expected by chance (Moran, 1950; Li et al., 2007). It is defined as:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (4.10)$$

where  $N$  is the number of zones,  $x_i$  is the value of the variable (e.g. the resilience index) at zone  $i$ ,  $w_{ij}$  is the spatial weight between zones  $i$  and  $j$ , and  $\bar{x}$  is the mean value. A significantly positive Moran's I indicate spatial clustering, justifying the need for spatial regression models.

A positive Moran's I value (significantly above the expected value under randomness) indicates positive spatial autocorrelation, meaning high values are adjacent to high values (and low with low), forming clusters. Conversely, a negative Moran's I suggest spatial dispersion, where high values are near low values, and values close to zero imply no spatial autocorrelation (a random spatial pattern). Significance of Moran's I was assessed, and a statistically significant Moran's I confirm that spatial processes are at work, justifying the use of spatial modeling in subsequent steps.

#### 4.4.1.2 Variance Inflation Factor

Avoiding unstable estimates due to multicollinearity among the independent variables, we computed the Variance Inflation Factor (VIF) for each predictor (James et al., 2023). VIF quantifies the extent to which the variance of a regression coefficient is inflated by linear dependencies with other predictors. For a given independent variable  $X_i$ , VIF is defined as:

$$VIF_i = \frac{1}{1 - R_i^2} \quad (4.11)$$

where  $R_i^2$  is the  $R^2$  from regressing  $X_i$  on other predictors. A VIF of 1 means no correlation with other predictors, while larger values signal higher collinearity. We employed a VIF threshold value of 5 to remove highly collinear predictors to have stable estimates of coefficients. Predictors with VIF above this value were removed iteratively to make sure that the independent variables left were not too

intercorrelated. This process yielded a set of explanatory variables appropriate for regression modeling without causing multicollinearity problems.

## 4.4.2 Statistical Analysis

Following the variable selection, statistical tests were undertaken to provide baseline associations and examine linear associations in the data. This was accomplished by carrying out the correlation analysis and running an Ordinary Least Squares (OLS) regression.

### 4.4.2.1 Ordinary Least Squares (OLS) Regression

An OLS regression was first employed as a baseline model to quantify the relationship between the independent variables (transportation diversity index and socio-economic factors) and the resilience metrics.

$$Y = X\beta + \varepsilon \quad (4.12)$$

where  $\beta$  represents the regression coefficients, and  $\varepsilon$  is the error term. Model fit was assessed using  $R^2$ , adjusted  $R^2$ , and residual diagnostics, including Moran's I on residuals to check for spatial dependence. The OLS results establish which factors are significant predictors of resilience in a non-spatial context and how much variance in resilience is explained, setting the stage for subsequent spatial analysis.

## 4.4.3 Spatial Analysis

### 4.4.3.1 Spatial Autoregressive Model

The Spatial Autoregressive Model (SAR) accounts for spatial dependency in resilience by incorporating a spatially lagged dependent variable (Griffith, 2009). This model assumes that the resilience metric ( $Y$ ) of a given zone is influenced by the resilience metrics of neighboring zones. The SAR model is specified as:

$$Y = \rho WY + X\beta + \varepsilon \quad (4.13)$$

where  $Y$  is the  $n \times 1$  vector of resilience metrics (dependent variable).  $WY$  represents the spatially lagged resilience metrics computed as a weighted sum of

resilience values in neighboring zones.  $\rho$  is the spatial autoregressive coefficient, measuring the degree of spatial dependence; a significant positive  $\rho$  indicates that zones with high resilience tend to be surrounded by zones with similarly high resilience (spatial clustering).  $X$  is the  $n \times k$  matrix of independent variables (transportation diversity, socio-economic factors, etc.).  $\beta$  is the  $1 \times k$  vector of regression coefficients.  $\varepsilon$  is the error term, assumed to be independent and normally distributed.

#### 4.4.3.2 Spatial Error Model

The Spatial Error Model (SEM) accounts for spatial dependence in the error terms rather than in the dependent variable itself (Rey and Franklin, 2022). This model is appropriate when the resilience metrics in different zones are not directly influencing each other, but unobserved spatially structured factors, such as missing variables related to urban planning, environmental policies, or some socio-economic factors create correlated errors. The SEM is specified as:

$$Y = X\beta + \zeta, \zeta = \lambda W\zeta + \varepsilon \quad (4.14)$$

where  $\zeta$  is the spatially correlated error term, accounting for omitted spatial effects.  $W\zeta$  represents the spatially lagged errors, capturing spatial dependence in unobserved factors.  $\lambda$  is the spatial error coefficient, measuring the degree of spatial dependence in the error structure.  $\varepsilon$  is the independent and normally distributed error term.

The SEM model assumes that while the resilience of a given zone is not directly influenced by that of its neighbors, unobserved spatially correlated factors create residual dependence. A significant  $\lambda$  suggests the presence of spatially structured noise, indicating that missing spatial factors affect resilience in clusters of zones.

#### 4.4.3.3 Multi-Scale Geographically Weighted Regression

Though SAR and SEM control for spatial dependence, they assume that regression coefficients are constant across space, apart from the spatial effects introduced by lagged terms. Yet, the associations between resilience and its determinants can differ geographically because of localized socio-economic, infrastructural, and environmental variations. To accommodate this spatial non-stationarity, Multi-Scale Geographically Weighted Regression (MGWR) is used (Fotheringham et al., 2017), an advanced development of Geographically Weighted Regres-

sion (GWR) (Fotheringham et al., 2006). In contrast to the standard GWR, which utilizes a single spatial kernel (bandwidth) for all predictors, MGWR allows every variable to have its own adaptive bandwidth, investigating at various spatial scales. It is very handy in such a heterogeneous urban setting as New York City, where transportation and socio-economic variability could exert varying degrees of influence based on local circumstances.

$$Y_i = \sum_k \beta_k(u_i, v_i) X_{ik} + \varepsilon_i \quad (4.15)$$

where  $Y_i$  is the dependent variable for taxizone  $I$ ;  $(u_i, v_i)$  are the centroid of taxizone  $I$ ;  $X_{ik}$  represents the  $k$ -th independent variable in taxizone  $I$ ;  $\beta_k(u_i, v_i)$  is the spatially varying coefficient for predictor  $X_k$  at location  $(u_i, v_i)$ .

Unlike SAR and SEM, where the coefficients ( $\beta_k$ ) remain constant across all observations, MGWR allows  $\beta_k$  to vary spatially, meaning that the effect of each predictor can change from one zone to another.

MGWR applies a spatial weighting function to estimate regression coefficients locally. Instead of assuming that predictor effects are uniform across NYC, MGWR assigns a unique spatial bandwidth for each independent variable, allowing differentiation between global and local effects:

- Broad-scale effects (large bandwidths): Predictors whose coefficients remain stable across space.
- Localized effects (small bandwidths): Predictors with highly spatially variable impacts.

The MGWR estimation process produces spatially varying coefficient surfaces for each predictor. Mapping these coefficients allows for the identification of where and how the effects of transportation diversity and socio-economic factors on resilience vary across NYC. Unlike SAR and SEM, which assume uniform effects, MGWR allows each predictor's effect to vary geographically. MGWR assigns optimal bandwidths for each independent variable, distinguishing between large-scale and localized relationships. By identifying which factors drive resilience in specific areas, MGWR enables policymakers to design targeted interventions.

# 5 Results

## 5.1 Exploratory Data Analysis

This section presents an exploratory analysis of the mobility and socio-economic and environmental data used in this study, serving as a foundation for the resilience metrics and modeling to follow. The analysis is divided into three parts. Section [5.1.1](#) examines the temporal distribution of mobility data, highlighting patterns and fluctuations in subway, taxi, and bike-sharing usage before, during, and after Hurricane Ida. Section [5.1.2](#) highlights spatial differences in mobility coverage, showing which areas had better access to different transport modes. Section [5.1.3](#) introduces the census and land use data, describing key demographic, socioeconomic, and infrastructural features. Together, these analyses provide the first insights into urban transport dynamics and disparities in New York City.

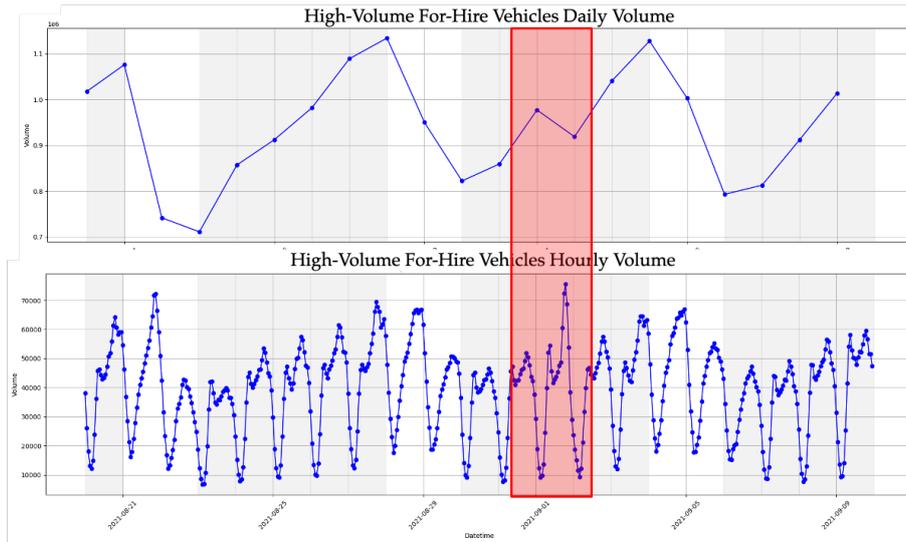
### 5.1.1 Temporal Distribution of the Mobility Data

This section is an exploratory data analysis of transportation volume data for taxis, CitiBike, and subway services in New York City from August 20th through September 10th, 2021. The main objective is to first determine the impact of Hurricane Ida, which struck the city on September 1st, 2021, and to examine temporal trends across these transportation modes.

Observation of the daily and hourly volumes of New York City taxis throughout the day reveals striking temporal patterns by service type. High-Volume For-Hire Vehicles (HVFHV, [Figure 5.1.a](#)) such as Uber and Lyft reveal pronounced peaks during morning and evening rush hours, with peak overall demand on Friday. Demand decreases precipitously over the weekend, indicating a strong connection to weekday commuting behavior. Standard For-Hire Vehicles (FHV, [Figure 5.1.b](#)), like traditional black cars and livery services, experience relatively stable demand during weekdays, with a moderate peak in the morning. However, their use also declines on weekends since this reflects their role in normal weekday

transport. Green and Yellow Taxis (Figure 5.1.c) have a consistent volume on weekdays but see a decrease on weekends. Unlike HVFHV and FHV, its peak usage primarily occurs in the evening, indicating a greater presence of recreational or convenience-oriented traveling during this period.

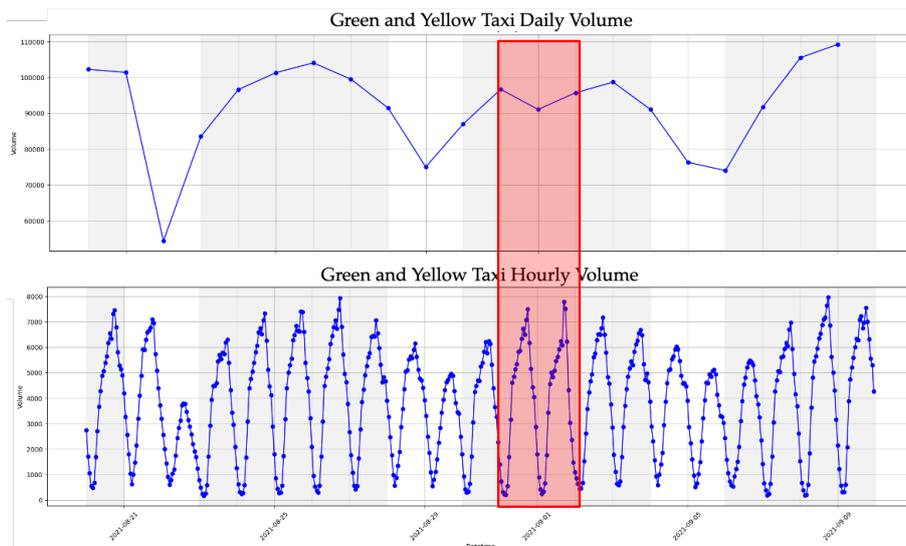
Most notably, on September 1st, 2021, all categories of taxis experienced a moderate but noticeable drop in usage, coinciding with the impact of Hurricane Ida. The hurricane brought record-breaking levels of rainfall to New York City, resulting in city-wide flooding and severe disruption to transport infrastructure.



(a) HVFHV volumes (daily and hourly).



(b) FHV volumes (daily and hourly).



(c) Green and yellow taxi volumes (daily and hourly).

Figure 5.1: Temporal variations of different transportation modes during Hurricane Ida, highlighting the disaster impact period (red band).

Citi Bike usage exhibits moderate regularity, with clear hourly peaks during typical commuting times—morning and evening rush hours—highlighting its role in daily work-related travel (Figure 5.2). While daily volume patterns are more variable than those of taxis, weekday usage generally remains higher, consistent with commuting behavior. A temporary dip appears around September 1st, 2021, likely reflecting short-term disruption from Hurricane Ida. However, usage rebounded rapidly, and Citi Bike even set a single-day ridership record on September 2nd, as it became a vital alternative when the subway system was shut down.<sup>5</sup>

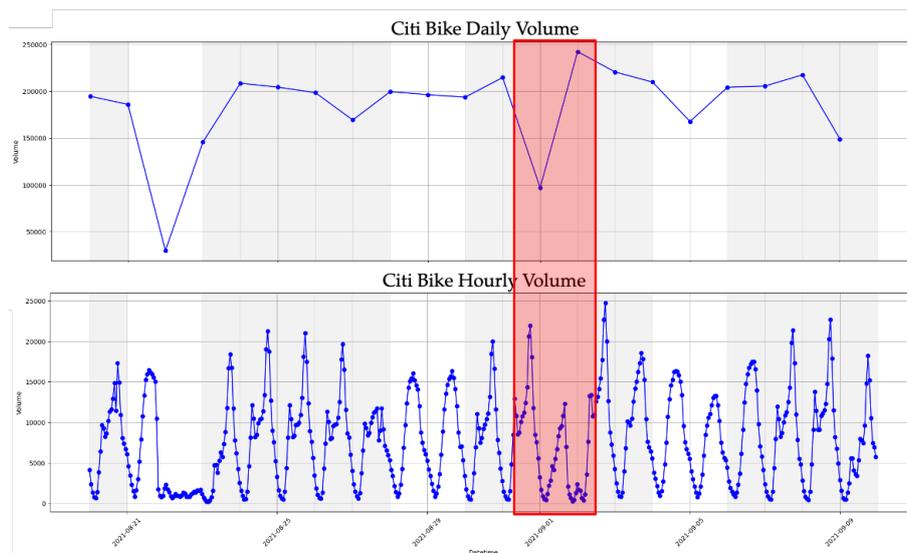


Figure 5.2: Citi Bike volume (daily and hourly)

As Figure 5.3 shows, subway patterns are characteristically commuter-oriented, with steep bimodal peaks reflecting morning and evening rush-hour usage and overnight low utilization. Subway volume metrics record an abrupt drop reflecting exactly the impact of Hurricane Ida, showing extreme disruption due to flooding and associated infrastructural damage.

<sup>5</sup> Citi Bike set single-day record after Hurricane Ida closed subways, New York Post, Sept 8, 2021. <https://nypost.com/2021/09/08/citi-bike-sets-record-after-hurricane-ida-shuts-down-nyc-subway>

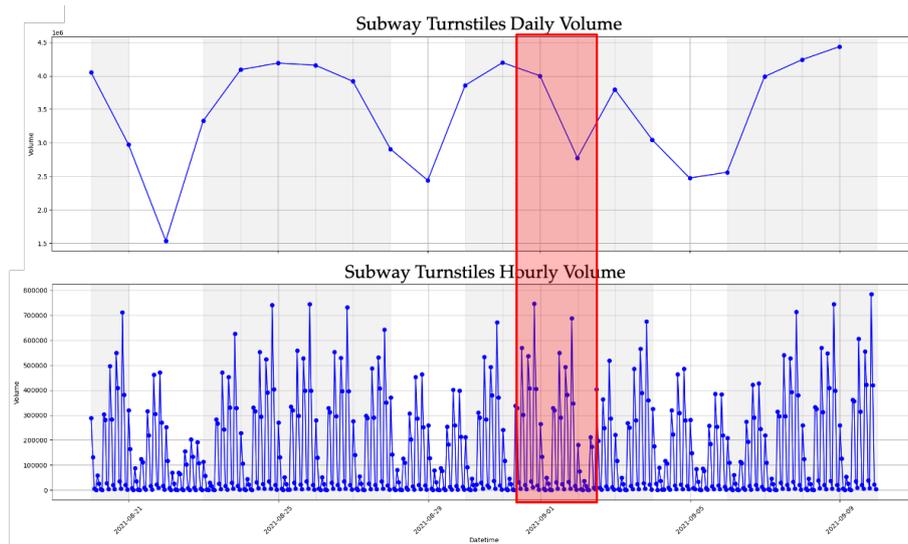


Figure 5.3: Subway volumes (daily and hourly)

In conclusion, initial evaluations show that all three transportation modes were disrupted during Hurricane Ida’s occurrence. The subway mode faced serious disruptions, including extensive flooding and halting of services. On 1st September 2021, CitiBike usage decreased considerably, possibly due to safety reasons during the storm. Interestingly, we observed record-breaking rides the next day, with residents resorting to bike-sharing without subway operations. Conversely, taxi services were more resilient, keeping relatively steady operations throughout this time. These observations highlight the importance of examining mode-specific vulnerabilities in evaluating the resilience of transportation systems and anticipating future extreme weather conditions.

### 5.1.2 Spatial Distribution of the Mobility Data

New York City has diverse modes of transport catering to diverse spatial demands. HVFHV, For-Hire Vehicles, and Green/Yellow Taxis provide almost the entire city with extensive and flexible mobility. In contrast to taxis, as shown in Figure 5.4, the subway system is larger in scope, serving prominent commercial and residential hubs in Brooklyn, Queens, Manhattan, and sections of the Bronx. It is lacking, though, in reaching further outlying sections. CitiBike is limited in the scope it can reach, yet it is concentrated in central Brooklyn and Queens’ inner communities and provides last-mile service in high-population urban centers.

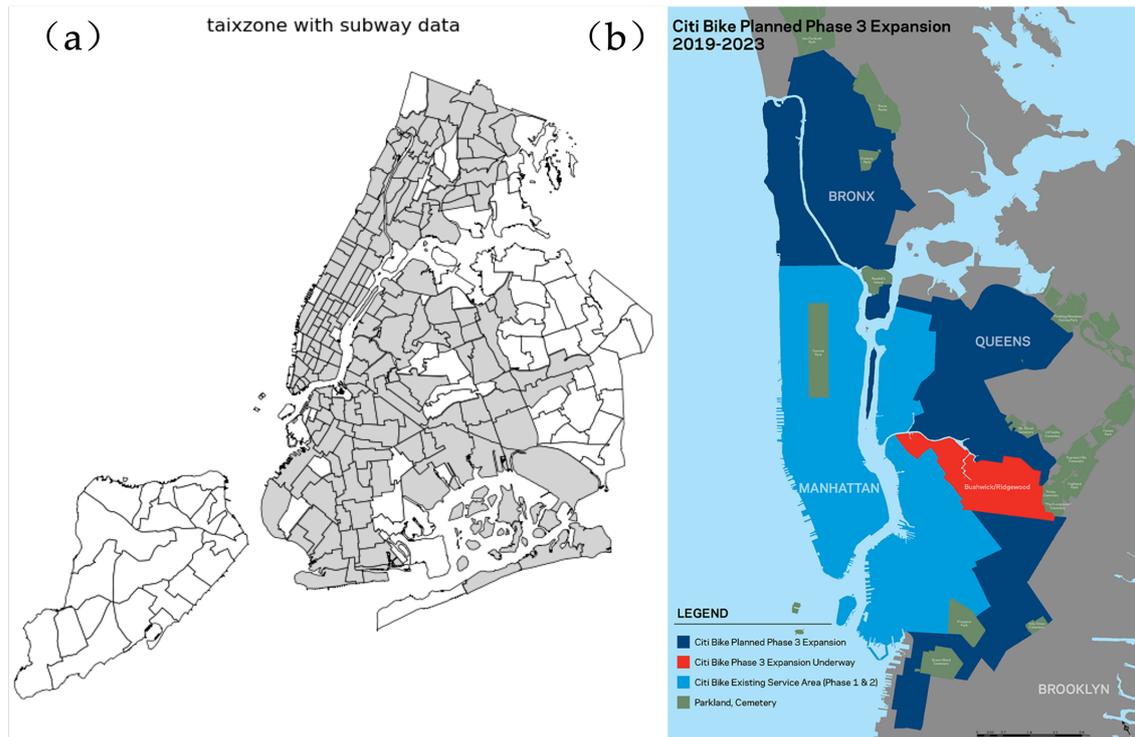


Figure 5.4: Operational coverage of the subway system and Citi Bike across NYC taxi zones. (a) Subway coverage (Source: NYC Open Data). (b) Citi Bike expansion (Source: NYC Office of the Mayor).

On September 1st, 2021, the spatial distribution of transportation volumes across New York City highlighted distinct usage patterns among different modes (Figure 5.5). Taxi services, particularly HVFHV, showed intense activity in central and southern Manhattan and pockets of high demand in southern Brooklyn and parts of Queens. Citi Bike usage remained largely confined to Manhattan, particularly in Midtown and Downtown, reflecting its limited operational coverage. Subway ridership was heavily concentrated along major transit corridors, especially in Midtown Manhattan and parts of northern Brooklyn, emphasizing its central role in commuter flows. These spatial patterns highlight each mode's unique functional role and geographic footprint in response to urban disruptions.

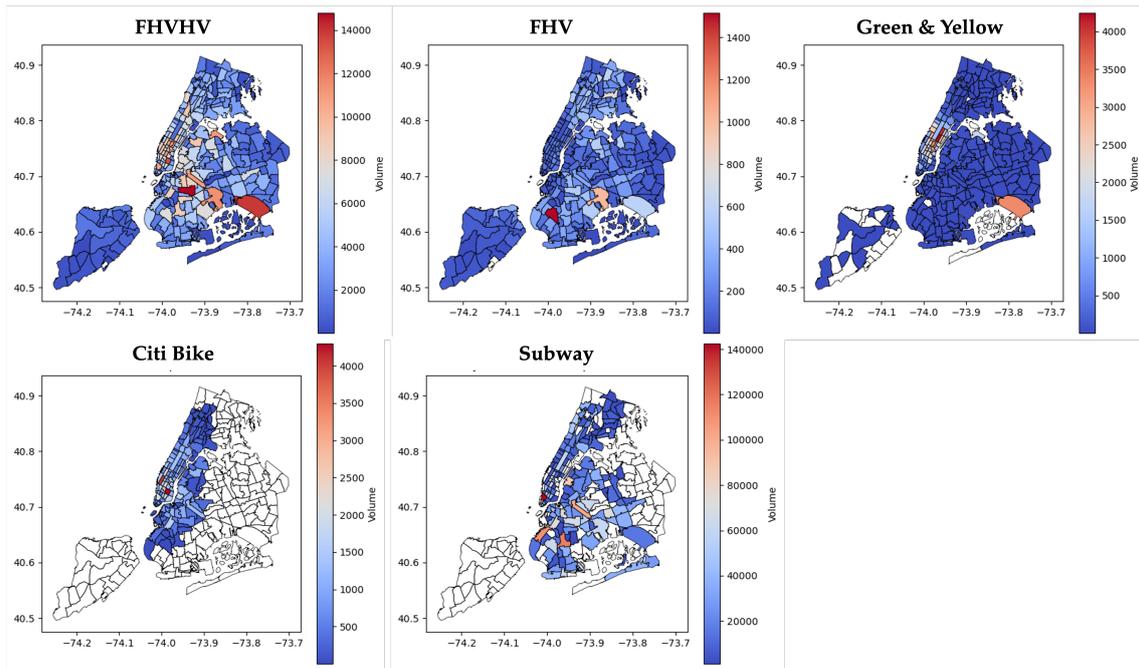


Figure 5.5: Spatial distribution of mobility volume by mode on September 1, 2021

### 5.1.3 Census Data

This exploratory data analysis, as shown in Figure 5.6, investigates the distributional properties of the selected demographic, socioeconomic, land use, and public service variables such as population density, income inequality measured by the Gini index, educational level measured by the percentage of bachelor's degree holders, health insurance coverage, unemployment rate, median income, land use (including residential, commercial, industrial, and other types), and access to emergency services (such as police, fire, and healthcare facilities).

The population is strongly right-skewed, so most areas have relatively low population numbers, and only a handful, primarily in inner urban areas, have very large populations. Gini Index indicates an approximately symmetric distribution with a mean of approximately 0.46, indicating moderate inequality for most areas but with both reasonably equal and reasonably unequal pockets.

The bachelor's Degree percentage is bimodal, reflecting polarization between well-educated and poorly educated neighborhoods. The no insurance percentage and Unemployment rate are extremely right-skewed, with most neighborhoods having low values and a small number having disproportionately high values. Median Income is right-skewed, as one would anticipate with numerous modest-income neighborhoods and some high-income enclaves.

Residential uses dominate most tracts of the city most tracts of the city are dominated by residential uses. Commercial and factorial land use distributions are skewed, indicating fewer areas are dedicated to these specialized functions. Emergency service counts (Police, Fire, Health Facilities) also take on a right-skewed shape, with few zones containing many facilities and most containing fewer or none.

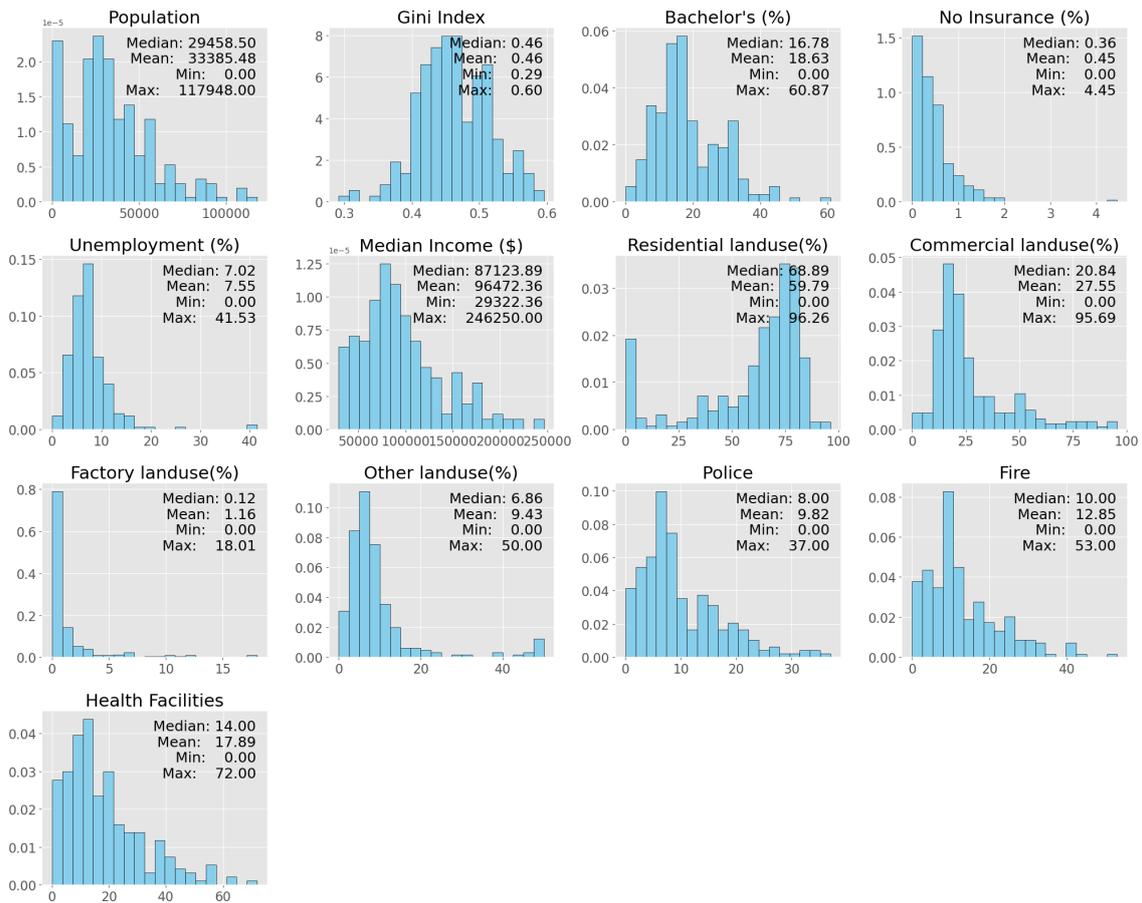


Figure 5.6: Histograms of socioeconomic and environmental variables

The spatial pattern in Figure 5.7 describes and explains the observed statistical distributions by indicating where and why patterns are found:

Population densities in Manhattan, middle Brooklyn, and western Queens coincide with mixed-use zoning and transit accessibility, explaining the right-skewed distribution. Few highly populated areas exist, and densities in most peripheral areas are low.

Distributions of Income and Gini Index reveal spatial inequality. The relatively few very high-income tracts of Midtown and Upper Manhattan generate the long upper tail of the income histogram. Meanwhile, low-income neighborhoods in

the Bronx, southern Brooklyn, and parts of Queens are responsible for the range of values seen in Gini values.

The attainment of bachelor's degrees follows the same pattern: successful areas, like Manhattan's west side and Park Slope in Brooklyn, show high levels of education (describing the second mode in the bimodal distribution), while lower levels in the Bronx and southeastern Queens areas are discovered matching the first mode.

The inequitable distributions of the uninsured and unemployment rates are geographically realized through their concentration in deprived neighborhoods, most notably the Bronx, as well as parts of Brooklyn, where access to healthcare services and secure employment opportunities has long been limited. Right-skew of median income is due to the presence of very few high-income places, while the remaining NYC belongs to a middle-income group—this maps geographically to concentrated income gradients among nearby tracts.

Land use distributions also coincide quite closely with zoning. Commercial and factory uses are mostly confined to business districts and industrial corridors (e.g., Long Island City, Sunset Park), leading to long-tailed distributions. The majority of tracts are overwhelmingly residential (explaining the skewed histogram). Emergency services are unevenly concentrated in central areas, namely Manhattan and downtown Brooklyn, contributing to skewed distributions. These patterns are due to planning priorities and the imperative of rapid coverage in the densest areas, but also reveal the gaps in peripheral or low-density areas.

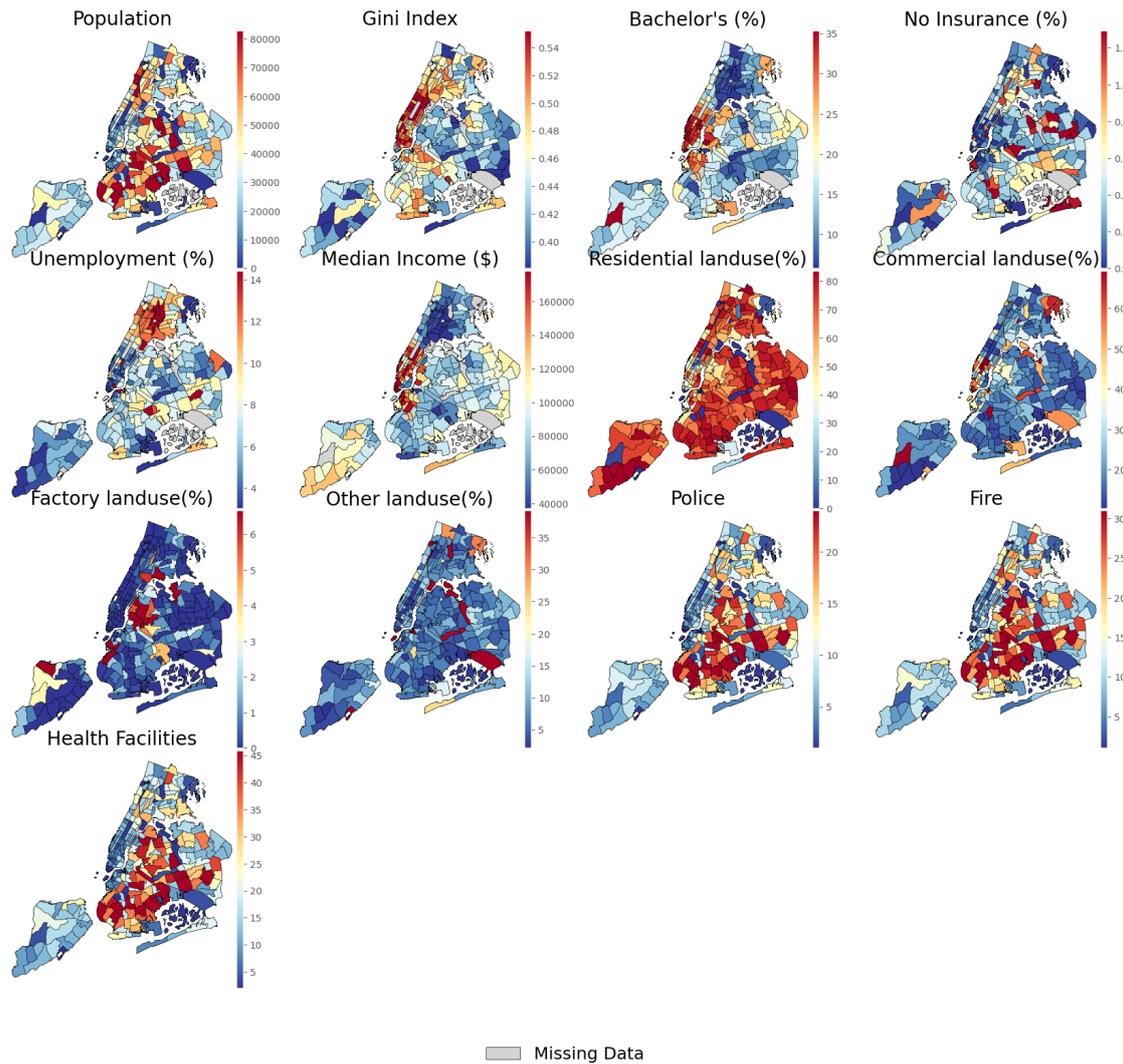


Figure 5.7: Spatial distribution of socioeconomic and environmental variables.

Collectively, the spatial patterns provide the spatial logic behind the statistical forms. Centralization of urban areas, land use concentration, and systematic inequality produce right-skewed distributions. Large values are found only in a few zones, and most areas are on the low side. Bimodal or symmetrical distributions suggest spatial segregation in the sense that high and low values fall into discrete pockets instead of being spread out in space uniformly. These intertwined patterns are necessary to understand the social context within which transportation and mobility resilience unfold.

## 5.2 Descriptive Statistics

This section provides a comprehensive overview of the key variables and indicators used in the analysis. It begins with the calculation and spatial assessment of transportation diversity based on metrics of richness and evenness (Section 5.2.1). It then presents the results of the mobility resilience model, including model performance evaluation, disaster period detection based on dynamic forecasting, and the quantification of resilience metrics such as resilience loss, recovery time, and maximum deviation (Section 5.2.2). Finally, Section 5.2.3 describes the data transformation steps applied prior to regression modeling.

### 5.2.1 Transportation Diversity Calculation

Here, for this transportation diversity calculation (Figure 5.8), publicly available NYC transportation data were taken into account, such as bus, subway, bike route infrastructure, and road networks for driving and walking access. These modes of transport reflect urban multi-modal mobility quite well: buses and subways reflect the primary public transit services, bike routes reflect environmentally friendly active modes of transport, and road networks reflect overall access for cars and pedestrians. The richness and evenness distributions of these infrastructures all exhibited unique spatial patterns. Bus, subway, and bike route measures all exhibited highly right-skewed distributions, revealing extreme spatial disparities with infrastructures concentrated within very small spaces. Road network measures exhibited more centrally located distributions with lower skewness, reflecting relatively even spatial coverage. The results reveal unequivocal differences in the distribution of transportation infrastructures in NYC.

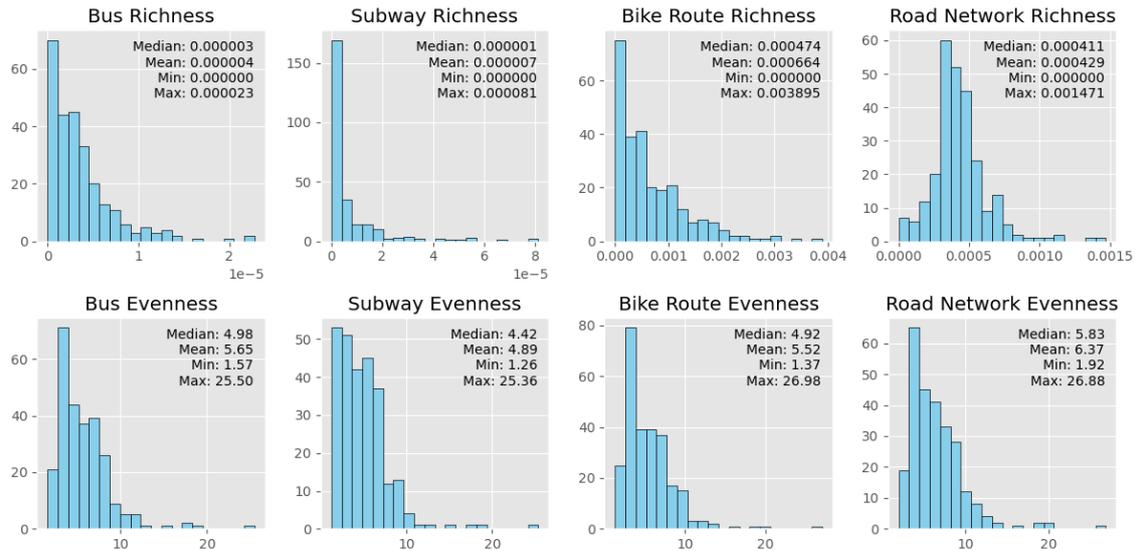
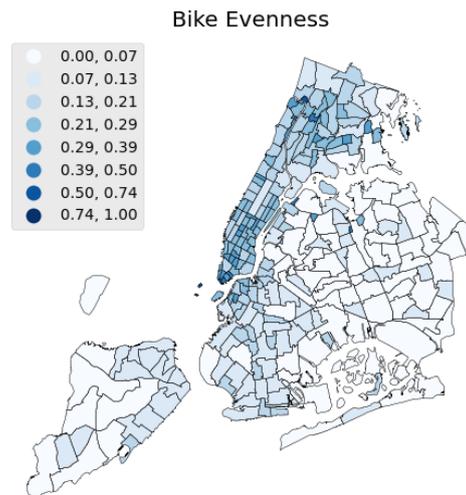
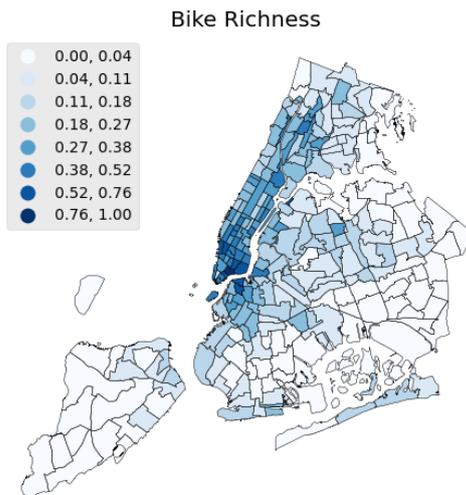
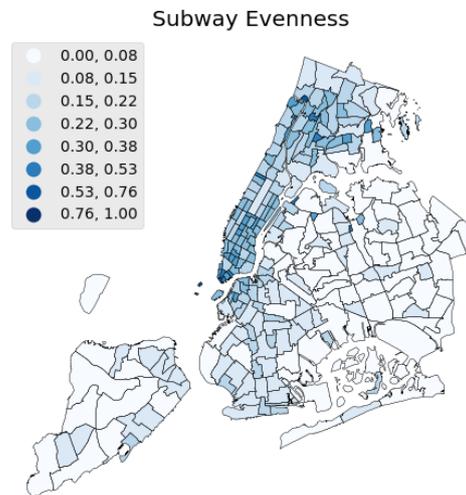
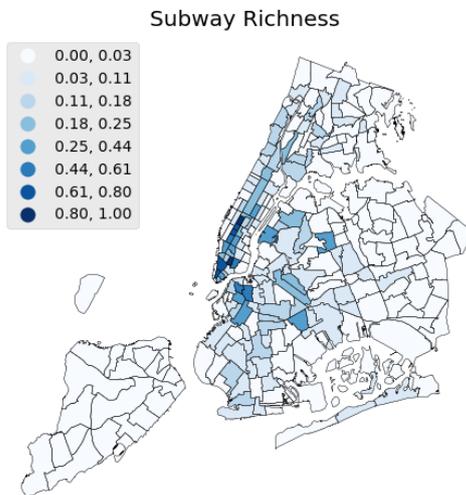
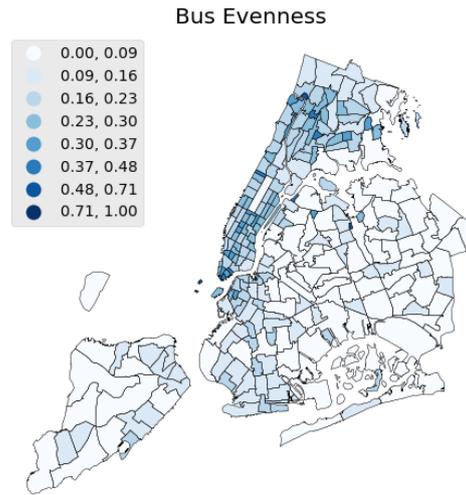
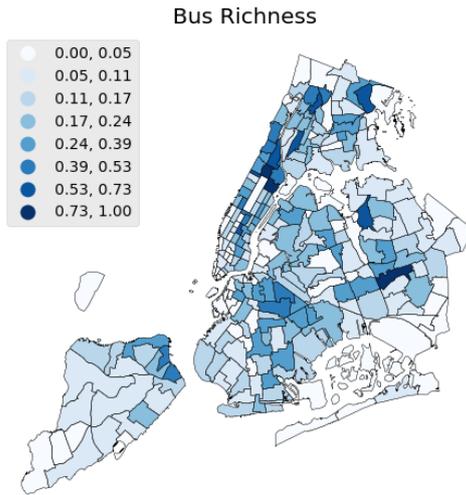


Figure 5.8: Histograms of transportation's richness and evenness.

To enhance the calculation of the transportation diversity index, all of the corresponding indicators were normalized. Figure 5.9 illustrates the geographic distributions of these indicators, showing a strong core-periphery pattern. Highly connected and evenly distributed bus, subway, and bicycle routes are mainly concentrated in Manhattan and the central parts of Brooklyn and Queens. On the other hand, outlying areas, such as Staten Island and eastern Queens, exhibit comparatively lower values, implying limited availability of multi-modal transport options. However, the complexity of the road network is more evenly spread across the city, although central areas have slightly higher complexity. These patterns point to the unequal distribution of multi-modal infrastructure in New York City.

A spatial pattern analysis of transport mode richness and evenness reveals evident complementarities among infrastructures. Given their wide geographical reach, bus routes complement subway operations by penetrating areas with poor subway coverage, particularly in the outlying areas of Brooklyn, Queens, and the Bronx. In contrast, subway lines are most dense in Manhattan and surrounding highly populated areas and serve high-capacity transportation needs in the city center. Bike routes largely coincide with areas serviced by subway and bus systems, particularly in Manhattan and western Brooklyn, facilitating green transit alternatives for short-range commutes. Yet, the absence of bike infrastructure in outer neighborhoods highlights a deficiency of varied transit options.



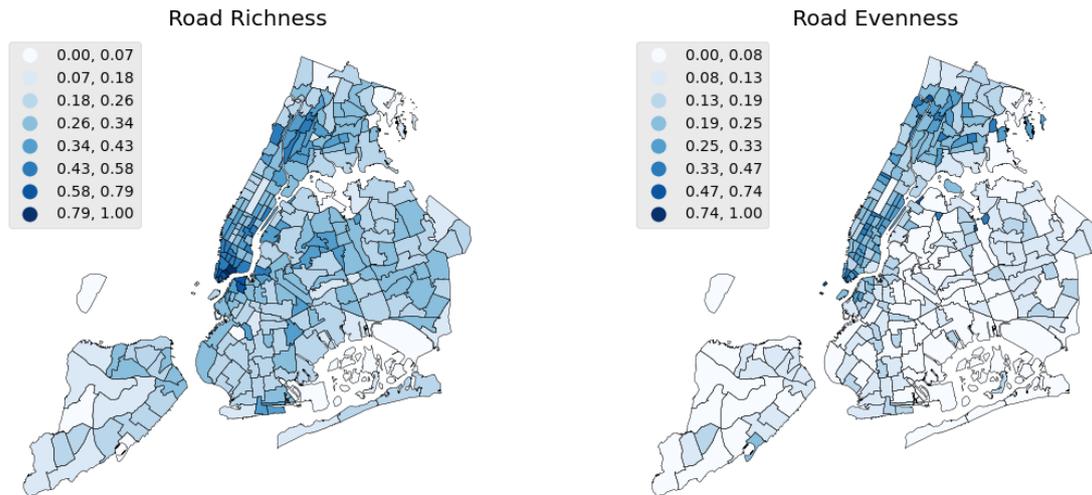


Figure 5.9: Spatial distribution of transportation richness and evenness.

The resulting Transportation Diversity Index (TDI), which integrates these complementary strengths by the equal weighting of equally normalized evenness and richness values between modes of transportation, emphasizes extreme spatial differences. As shown in Figure 5.10, the highest TDI rankings, in particular, are concentrated in central and lower Manhattan and adjacent areas of Brooklyn and Queens, while considerably lower rankings are widespread in peripheral neighborhoods such as Staten Island and outer Queens. The center-periphery disparity emphasizes the need for strategically directed transportation infrastructure investments with a focus on enhancing accessibility and equity in less transportation-diverse communities.

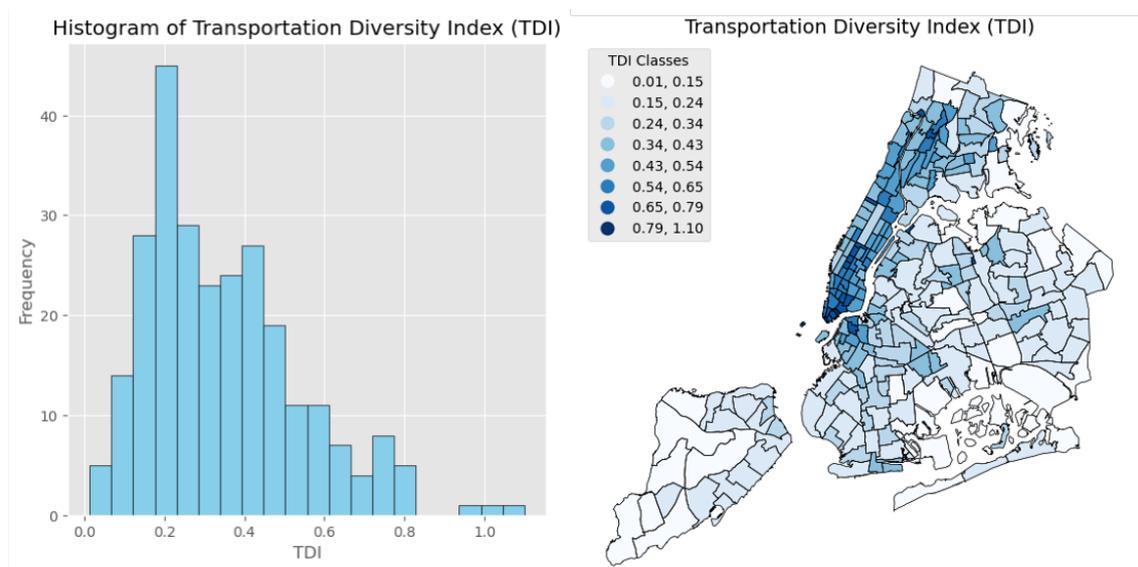


Figure 5.10: Histogram and spatial distribution of the TDI

## 5.2.2 Mobility Resilience Model Result

This study adopts a dynamic baseline approach using the Prophet time-series forecasting model to accurately capture human mobility resilience under the impact of extreme disasters. Unlike traditional static baselines that rely on fixed average values, the dynamic baseline reflects time-specific expectations of mobility volumes, accounting for natural daily and weekly fluctuations.

### 5.2.2.1 Model Performance

Figure [5.11](#) illustrates the performance of the Prophet model according to the method described in Section [4.2.2.1](#). It was trained using historical data from January 1st to July 31st, 2021, with August 9th to August 15th as the test dataset.

The test window in August was chosen for two primary reasons. First, it comes right before the anticipated disaster window of Hurricane Ida in early September, so the mobility patterns are under the influence of the same seasonal, meteorological, and policy conditions. Such a close temporal relationship is required for creating a dynamic, realistic baseline. Second, the absence of significant disruptions or anomalies during this interval qualifies it as an untainted benchmark period suitable for model validation under normal, steady-state mobility conditions.

There are several observations supporting the model validity: (1) Predicted values accurately capture true trends, with the ability to satisfactorily portray both weekly and diurnal mobility patterns; (2) The majority of actual values are contained within the 95% confidence interval, indicating very good model stability and reliability; (3) Minor exceedances of the upper bound during peak intervals (e.g., following September 9th) portray the flexibility of the model in responding to unforeseen demand spikes. The results confirm that the dynamic baseline derived via the Prophet methodology provides a reliable estimate of normal, undisturbed mobility conditions and a strong baseline for identifying disaster-related disruptions.

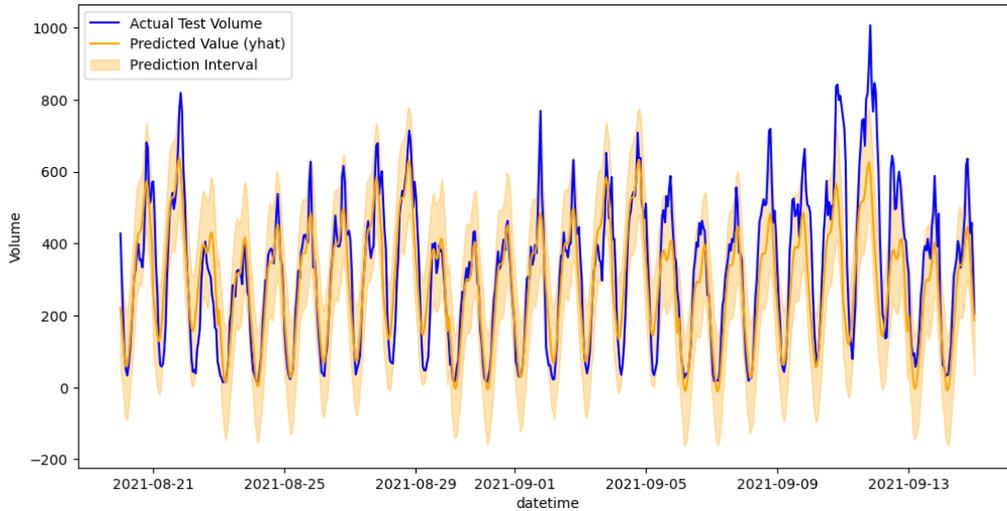


Figure 5.11: Mobility resilience model performance.

To assess the model’s performance in all taxi zones, we calculated important evaluation measures such as RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), and percentage of test observations in the 95% prediction interval. Summary statistics of the measures by the three modes of transportation are presented in Table 5.1. The ‘Accuracy\_95CI’ metric gives the proportion of observations within the 95% confidence interval, averaging at or above 0.95 in every mode, reflecting a very stable performance.

RMSE and MAE values also differ by mode because of variations in volume scale. Subway volumes are the highest in magnitude and inevitably have higher RMSE and MAE, whereas bike and taxi volumes have lower absolute errors. However, across all modes, the median errors are within acceptable limits, indicating that the model generalizes well to a variety of mobility settings and scales.

### 5.2.2.2 Disaster Period Detection

After establishing a dynamic mobility baseline with the Prophet model, the Hurricane Ida period was determined by following an orderly anomaly detection and temporal window estimation procedure. This procedure consisted of quantifying deviations from the expected baseline to determine statistically significant travel pattern disruptions. This approach was applied to the period from September 1 to 14, 2021, to forecast mobility and detect anomalies indicative of Hurricane Ida’s impact. Furthermore, the same process was exemplified using Taxizone 3 taxi data as an example.

To begin with, the dynamic baseline was established through the comparison

Table 5.1: Descriptive statistics of provided model performance

Mode	Metric	Mean	Std	Min	50%	Max
Taxi	Accuracy_95CI	0.950	0.033	0.793	0.959	1.000
	RMSE	48.874	49.373	0.000	30.264	298.339
	MAE	37.561	37.805	0.000	23.392	212.210
Subway	Accuracy_95CI	0.957	0.041	0.798	0.967	1.000
	RMSE	903.479	897.545	38.264	558.548	4697.932
	MAE	765.028	778.867	30.256	467.158	4227.533
Bike	Accuracy_95CI	0.948	0.023	0.890	0.951	1.000
	RMSE	32.784	27.162	2.673	28.317	146.132
	MAE	25.913	21.826	2.031	22.303	120.089

of observed volumes of mobility with what was expected and their respective 95% confidence intervals (Figure 5.12 (a)). The baseline rightly captures day-to-day changes and weekly cycles as the gold standard for identifying abnormal mobility patterns.

Subsequently, following Section 4.2.3.1, anomalies were detected where observed values were less than the lower bound of the confidence interval (Figure 5.12. b). A large cluster of such anomalies occurred from September 1st to September 5th, 2021, revealing major deviations from normal trends.

Lastly, the disaster window was established following the method described in Section 4.2.3.2. As illustrated by Figure 5.12 (c), the disruption period starts in the evening of September 1. It concludes on September 6, matching the extent of the most extreme and prolonged deviation from regular mobility activity. Following this period, mobility steadily increased to normal levels.

The given time window is the time boundary for the computation of resilience multi-modal, including resilience, resilience loss, recovery time, and maximum deviation (Section 4.2.4). The findings validate that the joint application of a dynamic baseline and anomaly detection in an educated time-series model efficiently identifies disruptions caused by disasters and delineates time periods of system vulnerability.

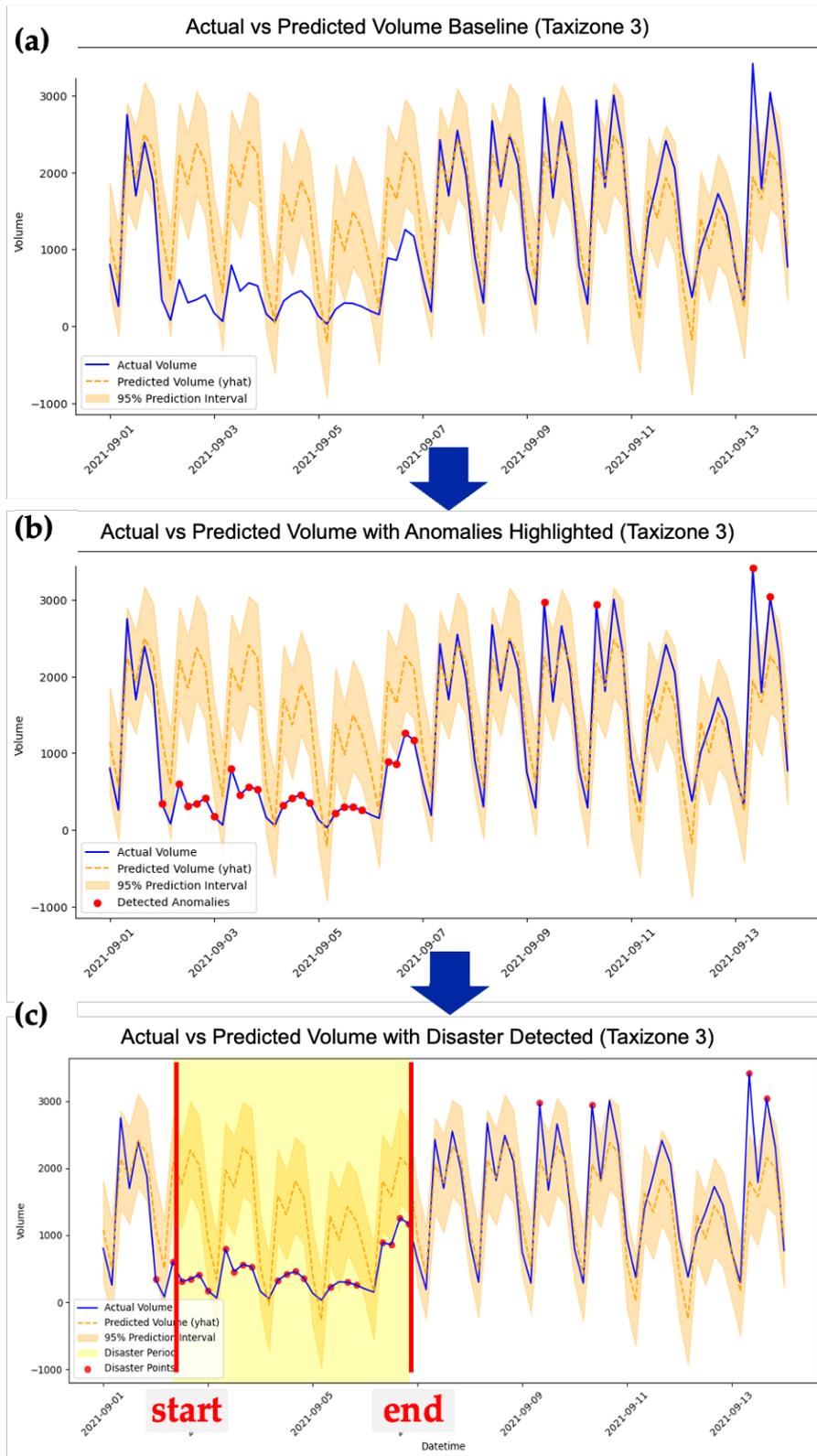


Figure 5.12: Mobility anomaly detection for subway in Taxizone 3. (a) Dynamic baseline prediction; (b) Anomaly identification; (c) Disaster period delineation.

### 5.2.2.3 Weight Assignment for Different Transport Modes

Once the disaster duration for every mode for every taxizone was ascertained, the corresponding resilience multi-modal for every transport mode was computed based on the equations provided in Section 4.3.3. Having determined these values, the task was to create an overall resilience indicator for every geographic unit. This was done by consolidating the computed resilience multi-modal for Taxi, Subway, and CitiBike trips into a combined predictor of overall resilience.

This integration was done by applying weights to each mode in terms of its relative passenger volume in baseline conditions. For every taxi zone, the total baseline mobility volume was calculated as the sum of the average trip counts of all three modes during the non-disruption period. The weight given to a specific mode was then calculated by dividing its average baseline volume by the total, thus representing the mode's influence on general travel patterns in that area.

Figure 5.13 illustrates the structure of this weighted aggregation method. Each mode's respective multi-modal resilience was weighted by its respective weight and aggregated to yield comprehensive metrics. This weighted aggregation method captures both the size of disruptions and the relative significance of each transport mode in a localized context, thereby guaranteeing that the ensuing resilience indicators are proportionally correct and contextually meaningful representations of multi-modal human mobility.

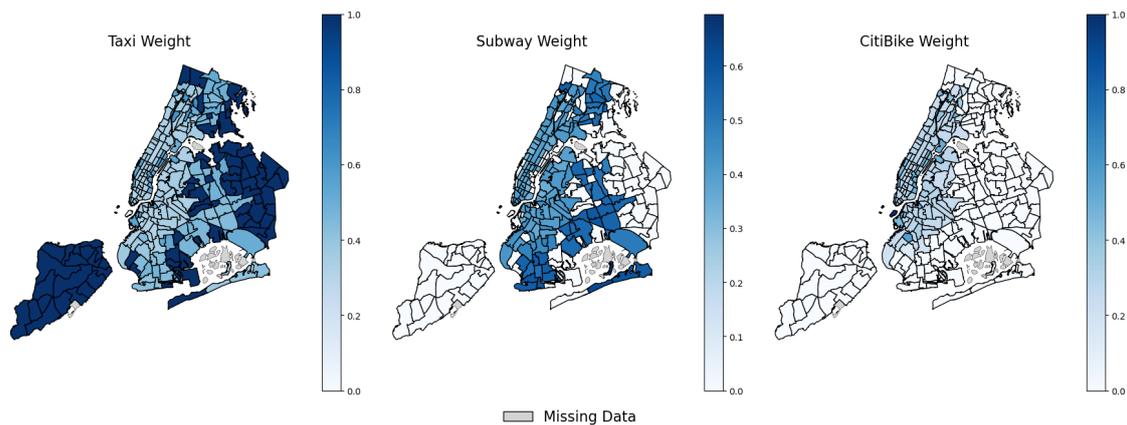


Figure 5.13: Weighted aggregation of resilience metrics based on baseline mobility volumes based on different transportation modes.

#### 5.2.2.4 Distributional Patterns of Multi-modal Resilience Metrics

In this study, we calculated five key resilience metrics: Resilience (R), Recovery Time (RT), Resilience Loss (RL), Max Deviation (MD), and Resilience Ratio (RA), for each transportation mode (Taxi, Subway, and Citi Bike) and subsequently aggregated these into overall resilience metrics using the weighted approach described in Section [5.2.2.3](#).

**Temporal Patterns of Disruptions and Recoveries** The heatmaps shown in Figure [5.14](#) illustrate the temporal pattern of observed disruptions (start times) and recoveries (end times) across taxi zones. Each cell represents the number of zones experiencing either the start or end of mobility disruptions at a given hour, with higher values indicating simultaneous events.

Disruptions peak in the late hours of September 1st, shortly after Hurricane Ida's landfall. Recovery events increase notably over the next 24–48 hours, particularly between September 2nd and 3rd. According to the MTA report, full subway service was restored within 32 hours after the storm<sup>6</sup>, which aligns with the temporal pattern captured by the model.

---

<sup>6</sup> <https://www.mta.info/press-release/new-york-city-transit-announces-restoration-of-all-subway-lines-morning-commute>

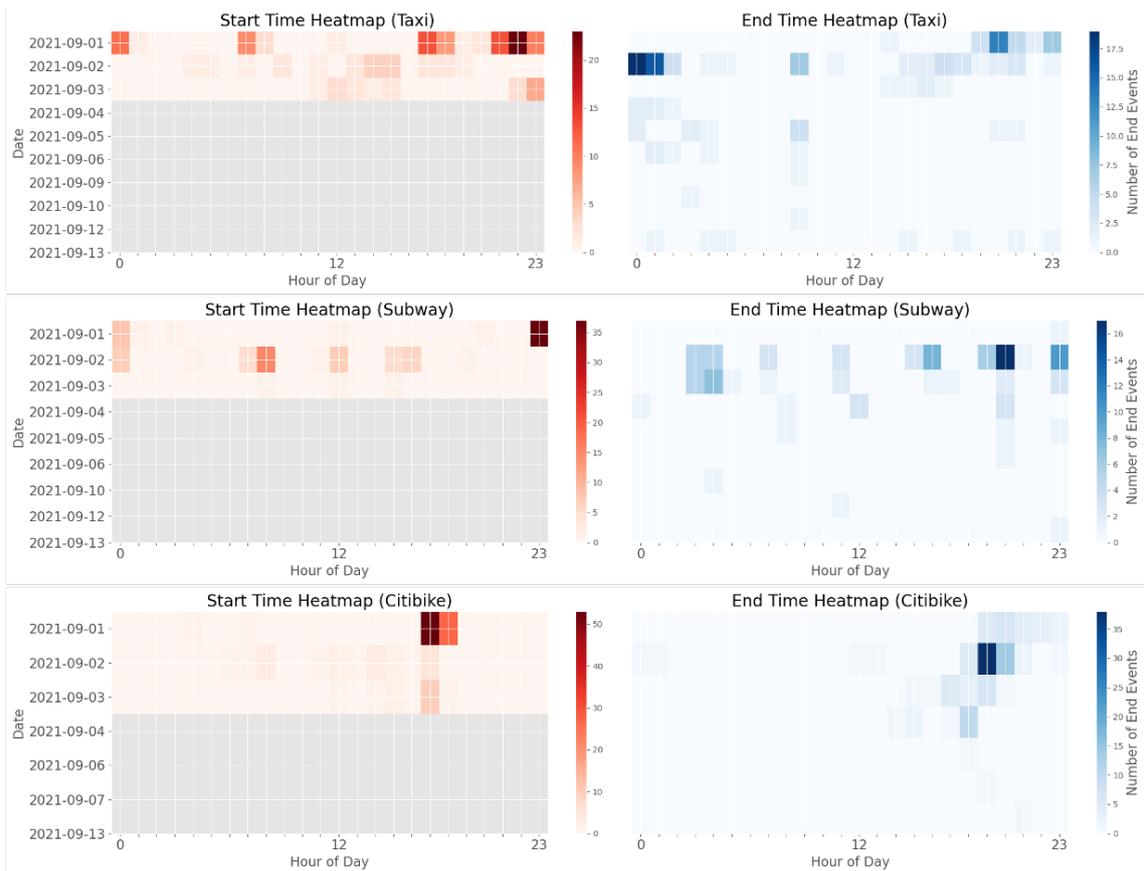


Figure 5.14: Heatmaps showing the temporal pattern of observed disruptions (start times) and recoveries (end times) by taxi zones.

**Resilience Metrics' Interrelationship** Correlation analysis reveals the main interdependencies between resilience metrics that vary considerably by transport mode and disruption severity. As shown in Figure 5.15, there is a strong and stable positive correlation between  $R$  and  $RT$  for most modes and conditions, suggesting that higher overall resilience tends to be associated with longer recovery times, possibly due to greater mobility volumes prior to disruption. An inconsistency is observed in the subway mode under the No Loss case, where  $R$  and  $RT$  have a negative but statistically insignificant relationship. This may indicate that when physical infrastructure is not damaged, shorter recovery times are more naturally associated with higher resilience. Under the No Loss case,  $R$  and  $RL$  have a mostly negative relationship across all modes, which suggests that smaller cumulative losses are conducive to higher resilience. This correlation is weakened for the Loss and Uncategorized case and even reverses in some instances, except for the Bike mode, which consistently shows a strong negative correlation between  $R$  and  $RL$ , indicating its adaptive nature.

The correlation between  $R$  and  $MD$  is largely non-significant, indicating that max-

imum disruption intensity is not by itself sufficient to predict overall resilience. In the same way, R is negatively correlated with the RA, as per the hypothesis that greater resilience will suffer less relative loss; the trend is maintained in a majority of cases, with the exception of the Overall–No Loss category. Ultimately, RT and RL turn out to have an acute negative correlation in No Loss regions for Taxi, Bike, and Overall metrics, which implies that large disruptions are being resolved quicker where no infrastructure was affected. In contrast, the Subway mode always shows a strong positive relationship between RT and RL, indicating that in fixed-infrastructure systems, higher losses imply longer recovery times. These findings emphasize the necessity for mode-specific and context-sensitive resilience planning in urban mobility.

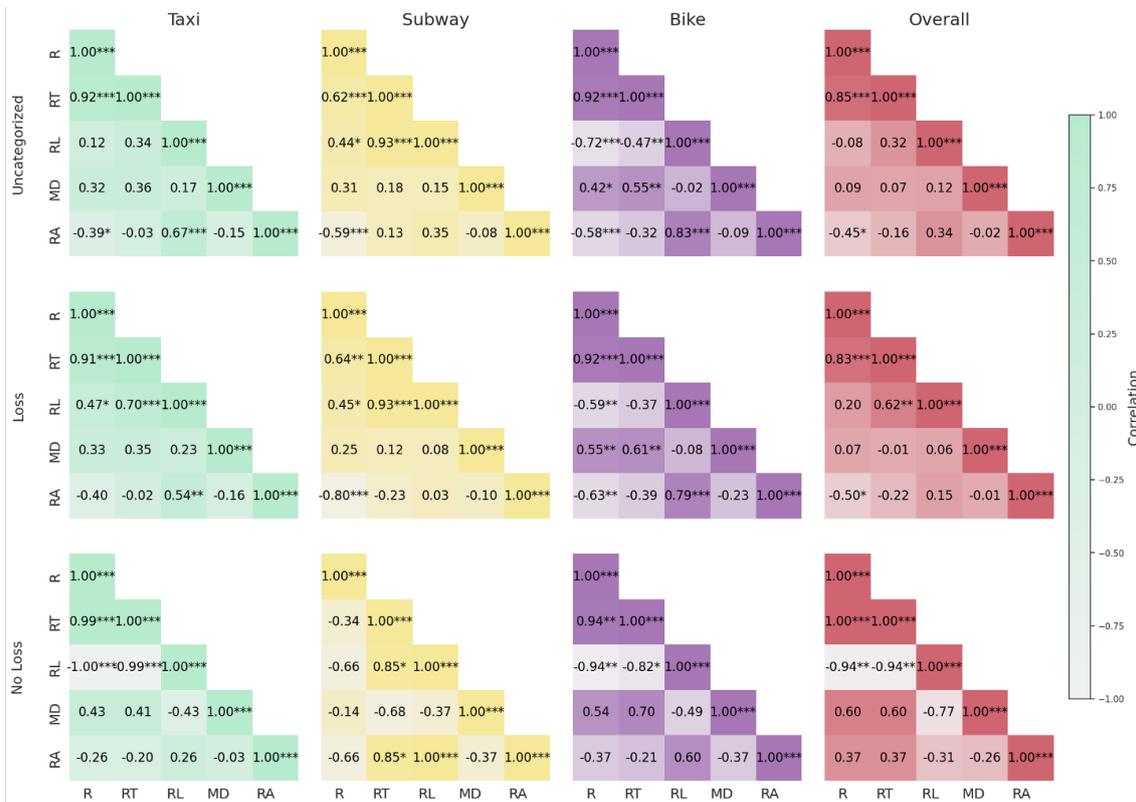


Figure 5.15: Correlation matrices of resilience metrics across modes and zones (No Loss, Loss, and Uncategorized). Significance levels are denoted as: \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

From the violin plot distributions (Figure 5.16) across resilience metrics, transport modes, and spatial types, several consistent patterns emerge that complement and reinforce previous findings from correlation analysis. The subway modes possess the most stable and narrow distributions for all resilience metrics, especially in No Loss areas, which is in line with previous findings of low variability and high system robustness. Bike metrics have the broadest distributions, how-

ever, especially in RL and RT, as expected of their high variability and sensitivity implied by their high positive inter-metric correlations. Taxi performance is intermediate, with moderate spread indicating heterogeneous neighborhood-level responses.

The difference between No Loss and Loss zones is immediately apparent: Loss zones all have broader and more scattered distributions for RL, RT, and MD, which aligns with correlation analyses that detected stronger correlations between disturbance severity and recovery time within these zones. No Loss zones, surprisingly, have high resilience and brief recovery times for every mode, and resilience ratios are frequently greater than 1, indicating compensatory surges for particular modes such as bike and taxi.

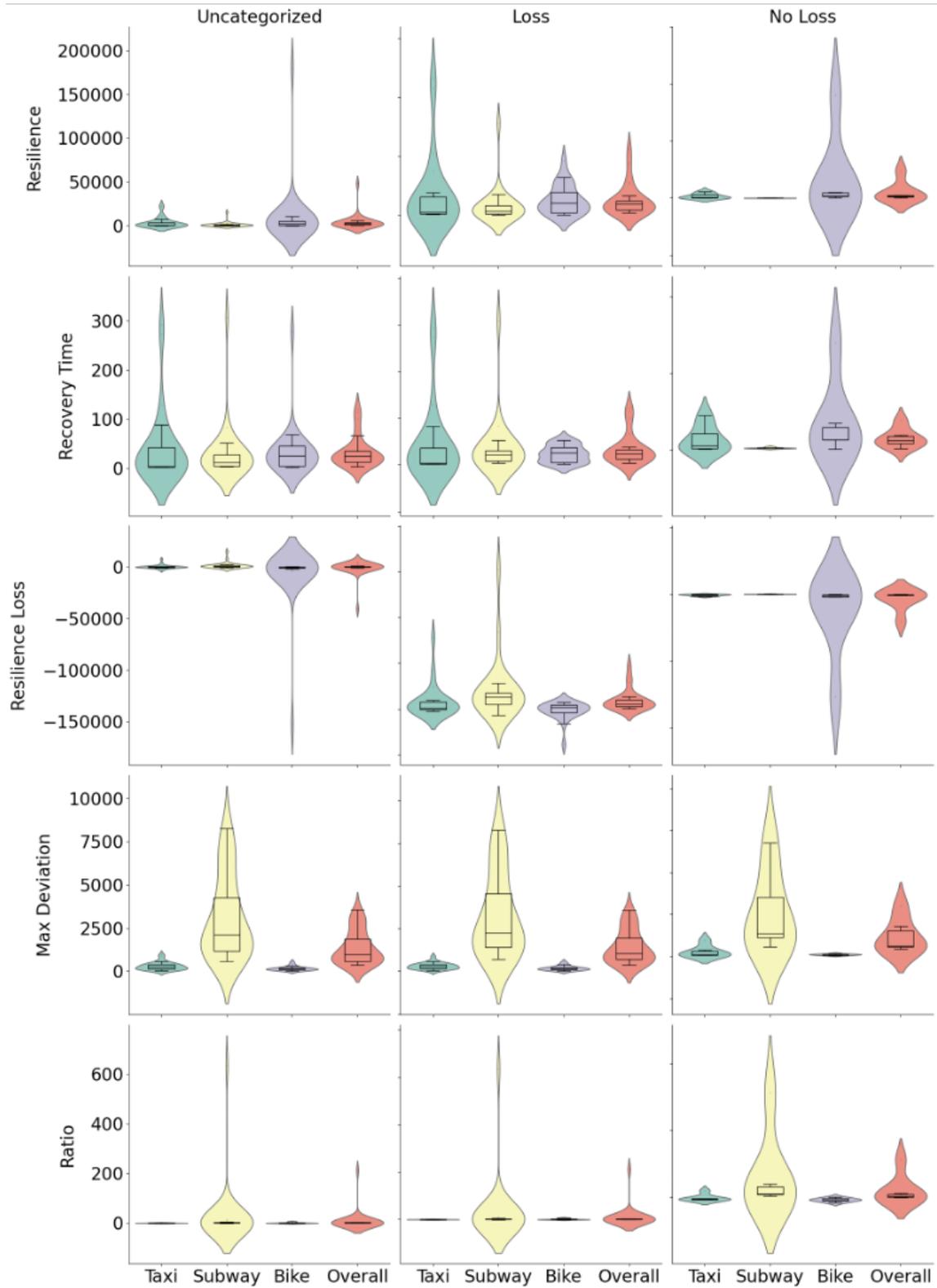


Figure 5.16: Violin plots of resilience metrics across modes and zones (No Loss, Loss, and Uncategorized).

Cumulatively, the combination of statistical correlations and distributional trends presents a clear picture of multi-modal resilience: subway networks show sys-

temic stability, taxi services demonstrate localized adaptability, and bike infrastructure remains highly sensitive to disruptions. Although the subway experienced the largest overall ridership decline during the disaster (as shown in Section 5.1.1), its standardized and infrastructure-based nature led to highly consistent responses across zones, resulting in stable and narrow resilience distributions. In contrast, taxi and bike modes exhibited greater spatial variation, reflecting their operational flexibility. The composite resilience metrics suggest that despite differences in mode-specific vulnerability, integrated multi-modal systems provide critical support for maintaining urban mobility under crisis conditions.

**Resilience Metrics Spatial Relationship** Figure 5.17 shows the spatial pattern of resilience metrics in New York City taxi zones. No Loss zones tend to have larger overall values of R, shorter RT, smaller RL, and smaller MD. These zones are mostly concentrated in central Manhattan and denser parts of Brooklyn and Queens. On the other hand, Loss neighborhoods, that is, those with high mobility losses due to Hurricane Ida, are concentrated in outer borough neighborhoods like southern Brooklyn, eastern Queens, and parts of the Bronx. They have lower recovery scores, longer recovery times, and higher deviations of services, showing more significant and extended disruption. Uncategorized neighborhoods, which are typically found in low-demand or data-poor neighborhoods, display a more dispersed distribution of resilience metrics.

Notably, Resilience and Resilience Loss maps exhibit complementary spatial patterns. Inner urban cores are the locations of clusters of high resilience, whereas high resilience loss is characteristic of outer boroughs. Recovery Time also exhibits strong spatial clustering, with outer boroughs featuring longer recovery times and shorter recoveries being more common in central areas. Maximum Deviation identifies localized hotspots of extreme mobility disruption, for instance, in Brooklyn and Queens. Lastly, the Ratio metric (Resilience to Resilience Loss) spatially matches the Resilience metric, with the greater values in the mid-city areas reflecting quicker recovery.

These geographic patterns powerfully reinforce the previous findings of violin plots and correlation analyses. The geographic concentration of high-scoring zones in No Loss areas and low-scoring zones in Loss areas graphically reinforces the quantitative differences found in previous sections. Outliers are also evident, such as pockets within resilient boroughs with very high RL or MD, demonstrating the presence of localized weaknesses despite broader trends. Overall, the spatial analysis confirms the view that urban mobility resilience during Hurri-

cane Ida was highly variable across the city, with strong performance in central business districts and peripheral locations demonstrating heightened vulnerability.

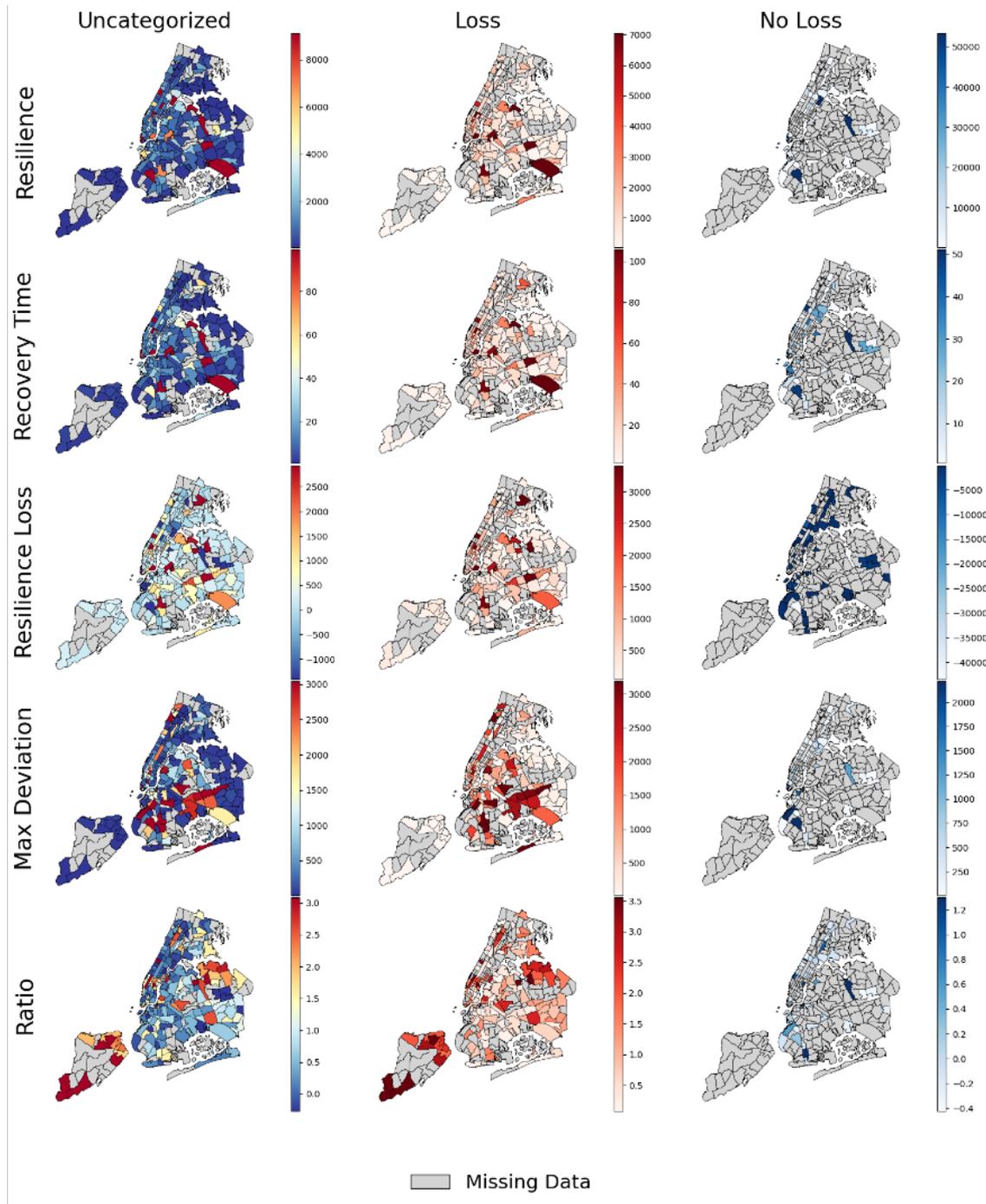


Figure 5.17: Spatial patterns of resilience metrics in NYC taxizones

### 5.2.3 Variable Transformation

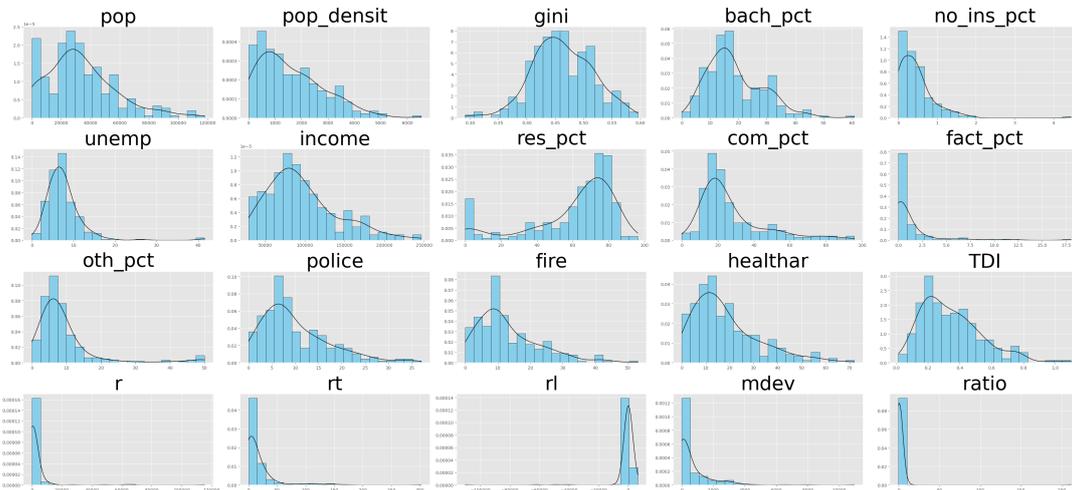
To ensure the validity and robustness of subsequent statistical and spatial analyses, it is necessary to address the distributional properties of the variables used in the study. As summarized in Table 4.4, many key variables, including demographic, socioeconomic, land use, and resilience indicators, exhibit significant skewness, heavy tails, or bimodal distributions. For example, variables such as population size, income, population density, and resilience-related metrics like resilience loss and recovery time are notably right-skewed.

Table 5.2: Variable transformation summary

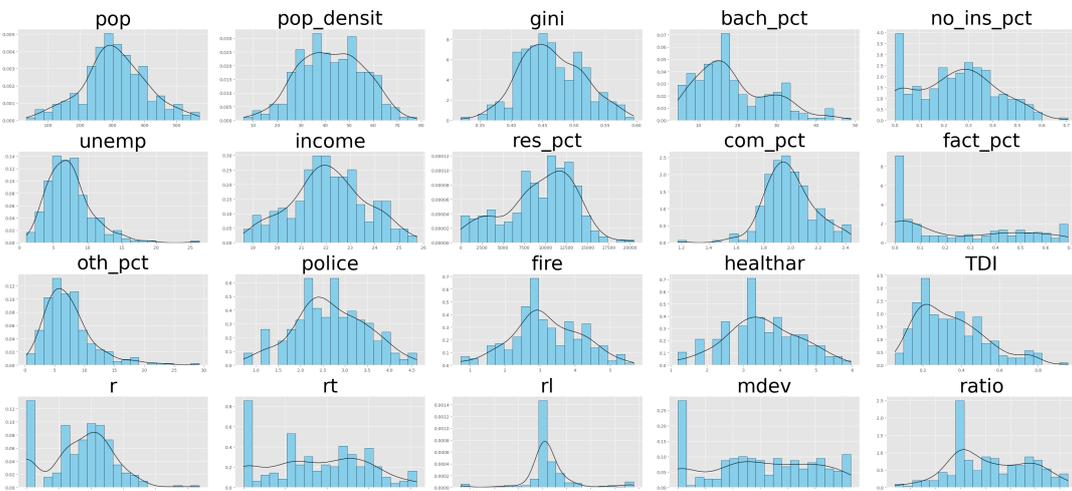
	pop	pop_densi	gini	bach_pct	no_ins_pct	unemp	income
W	✓						
B		✓	✓	✓	✓	✓	✓
Y							
	res_pct	com_pct	fact_pct	oth_pct	police	fire	healthar
W			✓				
B	✓	✓	✓	✓	✓	✓	✓
Y							
	TDI	r	rt	rl	mdev	ratio	
W			✓	✓	✓		
B	✓		✓		✓		
Y		✓		✓		✓	

*Note:* W stands for Winsorization; B stands for the Box-Cox transformation; and Y stands for the Yeo-Johnson transformation.

To mitigate these issues and improve the normality of variable distributions, three transformation techniques were applied in sequence: Winsorization, Box-Cox, and Yeo-Johnson. Winsorization was first used to reduce the impact of extreme values. Then, depending on the nature of the variable, either Box-Cox (for strictly positive variables) or Yeo-Johnson (for variables with non-positive values) transformations were applied. This stepwise approach enhances the distributional properties of the data and reduces heteroscedasticity, supporting more reliable statistical and spatial analyses. Figures 5.18.a and 5.18.b illustrate the histograms before and after the transformations, respectively.



(a) Histogram of original (untransformed) variable distributions.



(b) Histogram of transformed variable distributions after Winsorization, Box-Cox, and Yeo-Johnson transformations.

Figure 5.18: Comparison of variable distributions: (a) Original and (b) Transformed.

## 5.3 Variables Selection

### 5.3.1 Moran's I Result

Moran's I was used to test the level of global and local spatial autocorrelation for all the candidate variables. The bar plot in Figure 5.19 shows each variable's Moran's I value and p-value, indicating the significance and intensity of spatial clustering. A cut-off of Moran's I > 0.3 (at  $p < 0.001$ ) was used to keep only the variables that showed significant spatial autocorrelation. Meanwhile, Local Indicators of Spatial Association (LISA) cluster maps were generated (Figure 5.20), graphically displaying the real locations where high-high, low-low, or outlier cluster patterns were present. The coloring, red for strong clustering, yellow for moderate, blue for weak, and gray for nonsignificant, provides a quick visual representation of the spatial pattern of each variable.

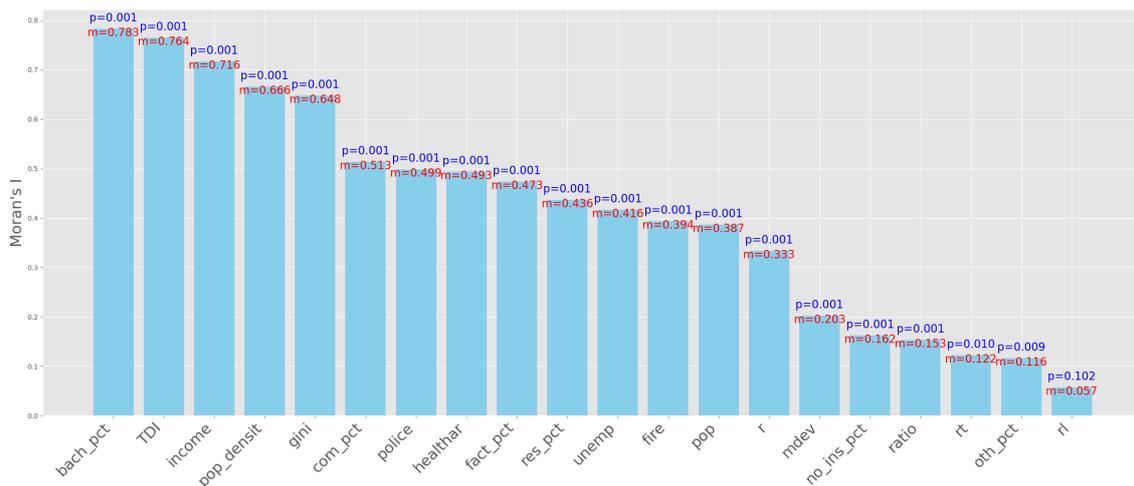


Figure 5.19: Bar plot of Moran's I values and p-values.

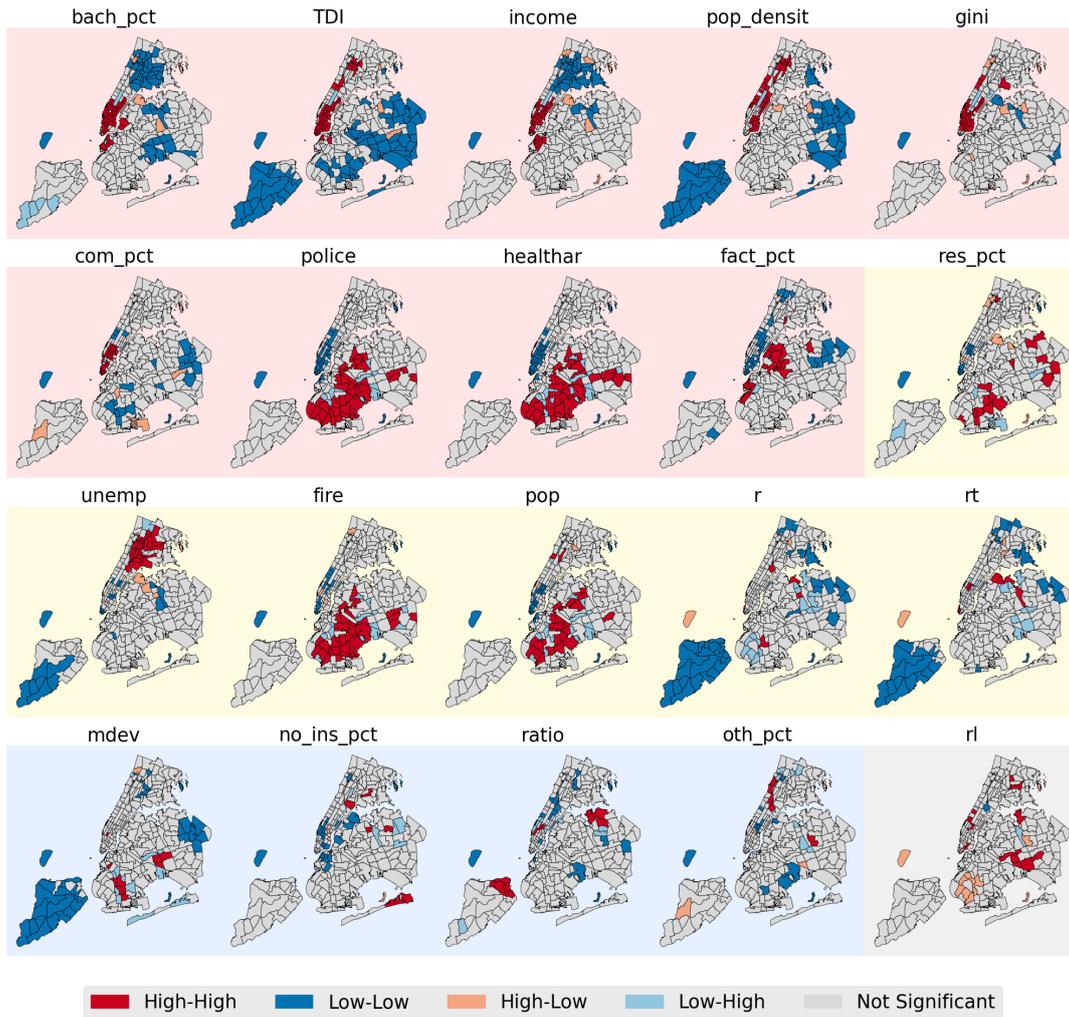


Figure 5.20: LISA maps of spatial clusters: red (strong), yellow (moderate), blue (weak), gray (nonsignificant).

### 5.3.2 Variance Inflation Factor Result

Then, the Variance Inflation Factor (VIF) testing was conducted to detect multicollinearity among the remaining variables. Figure 5.21 presents the VIF values of all candidate predictors. Several variables, such as “healthar” (VIF = 36.77), police (VIF = 28.48), and rt (VIF = 21.61), showed extreme multi-collinearity, well above the conventional threshold of 5, and were excluded to ensure model stability.

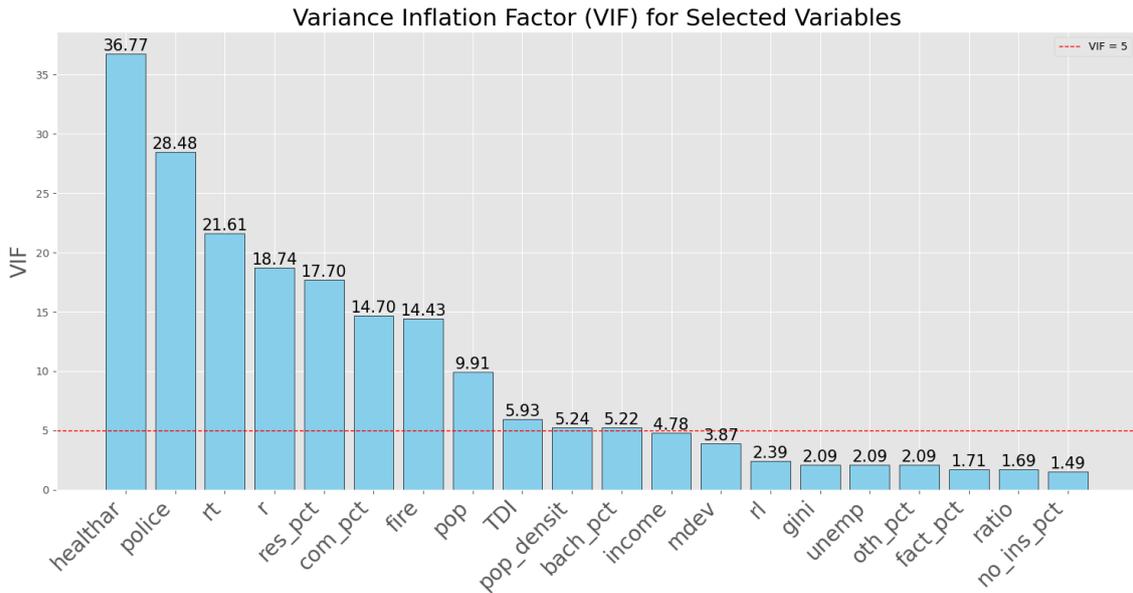


Figure 5.21: VIF values for candidate predictors.

Figure 5.22 displays the final selected variables for regression analysis. These variables were chosen based on high spatial autocorrelation (Moran's  $I > 0.3$ ,  $p < 0.001$ ) and low multi-collinearity ( $VIF < 5$ ) after filtering, ensuring they are both spatially meaningful and statistically stable.

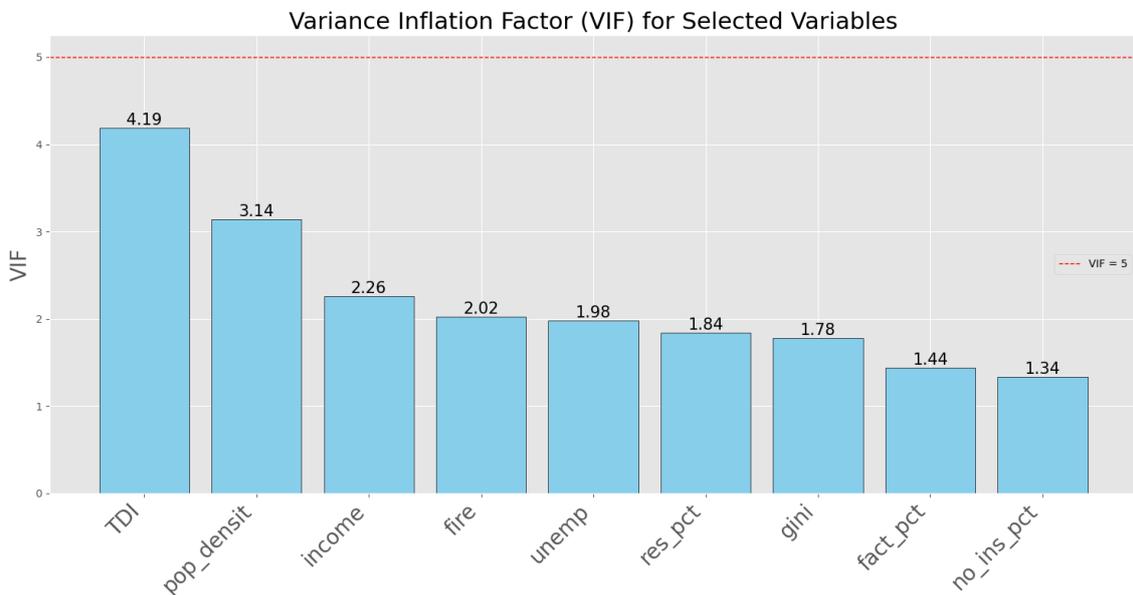


Figure 5.22: Final selected variables for regression analysis.

## 5.4 Statistical and Spatial Analysis

### 5.4.1 Correlation Between TDI and Resilience Metrics

Preliminary correlation analysis with non-parametric techniques—Kendall’s Tau, Spearman, distance correlation, and mutual information—demonstrated consistently weak or null correlations between the TDI and single resilience metrics (Figure 5.23). Regardless of the method, the strongest correlations identified were of moderate size (few exceeding 0.4). Stratifying by impact category (Loss vs. No Loss areas) yielded only moderate increases in the strength of association. Surprisingly, in the No Loss areas, the Spearman correlation between TDI and Maximum Deviation was 0.40, indicating a moderate relationship between transport diversity and the volatility of mobility patterns in unaffected areas. Overall, the correlation results suggest that the relationship between TDI and resilience is non-linear and inhomogeneous.

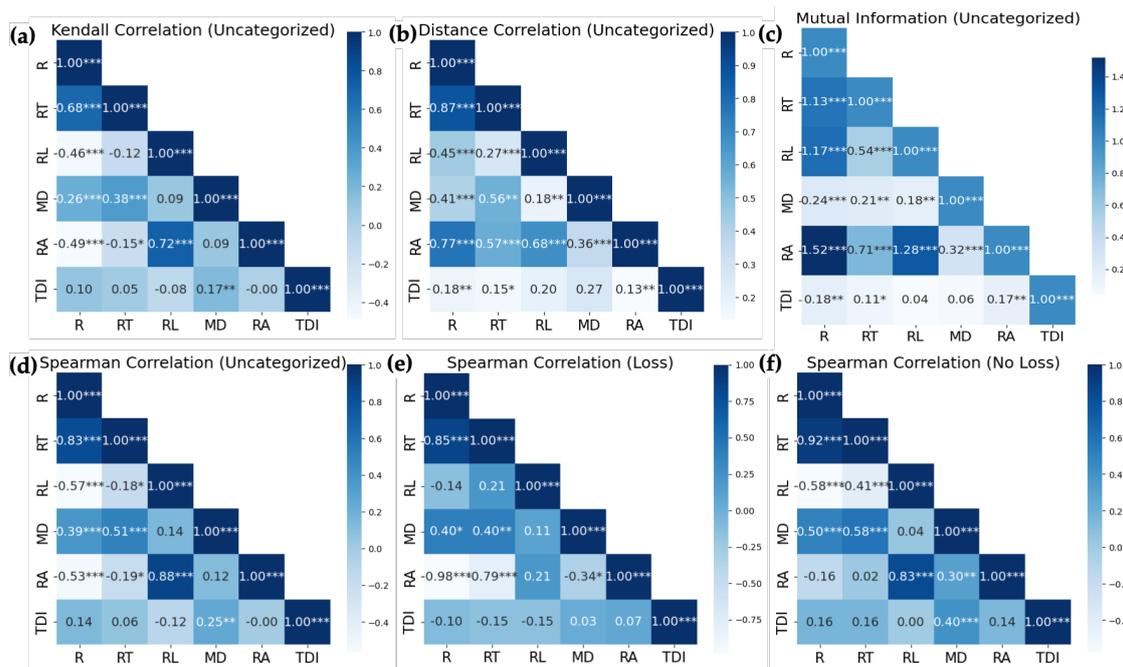


Figure 5.23: Preliminary correlation analysis between the TDI and resilience metrics.

### 5.4.2 Multivariate and Spatial Modeling

Given the inconclusive correlation results, we employed multivariate modeling to examine whether the relationship between the TDI and resilience metrics is influenced by additional socioeconomic and environmental factors, as well as

spatial structure. Specifically, we analyzed how resilience varies with TDI in conjunction with variables such as population characteristics, income levels, land use composition, and service accessibility.

To account for spatial dependence and heterogeneity, both global regression models (OLS, SAR, SEM) and local spatial models (GWR, MGWR) were applied, with 'Resilience' consistently used as the dependent variable. This approach facilitates a more comprehensive understanding of how the association between transportation diversity and resilience varies across different geographic and contextual settings.

Table 5.3: Regression results and statistical significance

	OLS	SAR	SEM	GWR	MGWR
<b>Regression Coefficient</b>					
<b>TDI</b>	22.01 **	18.84 *	22.22 **	1.38 **	0.65 *
<b>income</b>	1.16 ***	0.96 ***	1.08 ***	1.69 ***	1.39 *
<b>pop_densit</b>	0.00 ns	0.00 ns	0.00 ns	0.41 ns	1.14 ***
<b>gini</b>	17.84 **	14.41 *	17.33 **	0.90 **	1.22 *
<b>fact_pct</b>	2.71 ns	2.06 ns	2.29 ns	0.35 ns	0.25 *
<b>res_pct</b>	-0.00 *	-0.00 ns	-0.00 ns	-0.64 *	-0.54 *
<b>fire</b>	1.24 **	1.13 ***	1.19 ***	1.18 ***	0.72 *
<b>unemp</b>	0.13 ns	0.12 ns	0.12 ns	0.45 ns	0.27 *
<b>no_ins_pct</b>	1.01 ns	0.97 ns	0.93 ns	0.48 ns	0.16 *
<b>Model Performance</b>					
$R^2$	0.341	0.365	0.341	0.369	0.497
Adj. $R^2$	0.315	–	–	0.326	0.451
<b>Log-Likelihood</b>	-647.91	-644.625	-647.052	-642.689	-622.781
<b>AIC</b>	1316	1311.251	1314.104	1317.784	1300.888
<b>BIC</b>	1350	1349.353	1348.742	1373.906	1396.706

Note: Significance levels are denoted as: \* for  $p < 0.05$ ; \*\* for  $p < 0.01$ ; \*\*\* for  $p < 0.001$ ; ns for not significant.

### 5.4.2.1 Global Model: OLS

The OLS model provides a general standard to examine the linear relationship between the resilience indicator and explanatory factors. TDI is positive and significant ( $\beta = 22.01$ ,  $p < 0.01$ ), indicating that areas with greater diversity in transport are more likely to have greater mobility resilience. The remaining positively contributing variables are income (1.16,  $p < 0.001$ ), Gini index (17.84,  $p < 0.01$ ), and fire stations (1.24,  $p < 0.01$ ), indicating that economic wealth, moderate-income disparity, and availability of emergency services foster recovery. Residential land use percentage (*res\_pct*) is very weakly but negatively related ( $-0.00$ ,  $p < 0.05$ ).

But the  $R^2$  of 0.341 suggests modest explanatory power only, and residual spatial autocorrelation (Moran's  $I > 0.3$ ) is high, indicating failure of independence assumptions and suggesting the requirement for spatial modeling.

### 5.4.2.2 Spatial Dependence Models: SAR and SEM

In order to explain spatial dependence, SAR and SEM specifications were estimated. The SAR model fits better with a higher  $R^2$  of 0.365 and a statistically significant spatial lag coefficient ( $W_r = 0.283$ ,  $p = 0.006$ ), explaining that resilience in a particular zone is partially explained by that of its neighboring zones. Most of the variables remain significant, though a bit weaker than with OLS, e.g., TDI (18.84,  $p < 0.05$ ), Gini (14.41,  $p < 0.05$ ), income (0.96,  $p < 0.001$ ), and fire stations (1.13,  $p < 0.001$ ). The findings suggest that resilience has spatial spillover effects, which are effectively captured by SAR.

In contrast, the SEM model is not a considerable improvement over OLS ( $R^2 = 0.341$ ), and its spatial error term ( $\lambda = 0.179$ ,  $p = 0.146$ ) is insignificant, suggesting that the spatial patterns observed in resilience are more a reflection of the structure of the dependent variable rather than omitted spatial processes.

While TDI (22.22,  $p < 0.01$ ), Gini (17.33,  $p < 0.01$ ), and income (1.08,  $p < 0.001$ ) are still significant, the model is not an improvement over SAR, indicating that spatial dependency is more adequately explained by the outcome variable itself.

### 5.4.2.3 Local Spatial Models: GWR and MGWR

GWR estimates spatially varying coefficients and shows a moderate increase in explanatory power ( $R^2 = 0.369$ ). TDI (1.38,  $p < 0.01$ ), income (1.69,  $p < 0.001$ ),

and fire station density (1.18,  $p < 0.001$ ) are still significant predictors at most places, but with different local intensity. Importantly, the model also shows extremely strong spatial heterogeneity in the Gini coefficient, in line with potential neighborhood-level variation in how inequality affects recovery.

Yet, GWR uses one spatial bandwidth for all the variables, which restricts its capability to capture multiscale impacts. MGWR circumvents this by specifying variable-specific bandwidths. It performs the best ( $R^2 = 0.497$ , lowest AIC = 1300.89) with substantial improvement in model fit and explanatory capabilities. Here, TDI (0.65,  $p < 0.05$ ), income (1.39,  $p < 0.05$ ), and population density (1.14,  $p < 0.001$ ) are the global predictors—statistically significant, having reasonably consistent coefficients in space. Conversely, Gini (1.22,  $p < 0.05$ ), factory land use (0.25,  $p < 0.05$ ), and fire station (0.72,  $p < 0.05$ ) are strongly spatially heterogeneous, i.e., their influence is context-dependent and changes with the local infrastructure, industrialization, and urban density. The unemployment and uninsured population features are likewise locally significant but have variable implications in space, indicating that they are sensitive to neighborhood-level social structures and policy in modeling resilience.

#### 5.4.2.4 Spatial Effects and Multiscale Interpretations from MGWR

The Multiscale Geographically Weighted Regression (MGWR) analysis reveals distinct spatial patterns in the determinants of mobility resilience across New York City (Figures [5.24](#) and [5.25](#)).

**Citywide Predictors:** Both the TDI and income levels exhibit consistently positive and statistically significant effects on mobility resilience across nearly all taxi zones. Their spatially invariant coefficients and uniformly low p-values indicate that diversified transportation options and higher income levels are robust predictors of resilience throughout the city.

**Spatially Heterogeneous Variables:** Other variables demonstrate spatial heterogeneity in their relationships with resilience. For instance, the Gini coefficient shows strong positive associations with resilience in areas like Upper Manhattan, parts of Brooklyn, and Western Queens, while its influence diminishes in regions such as Staten Island or Eastern Queens. Similarly, population density has a more pronounced effect in central locations like Midtown Manhattan and Downtown Brooklyn, attenuating in peripheral areas. Factory land use percentage exhibits significant coefficients in South Brooklyn and the Bronx, highlighting industrial zones where land use can facilitate or hinder recovery. Variables such as resi-

dential land use percentage and fire station density have moderate yet intense influences focused on high-density urban cores, indicating that service accessibility and housing mix are essential in these areas.

**Localized Socio-economic & Environmental Factors:** The unemployment rate and the percentage of uninsured individuals emerge as highly localized variables. Their effects on resilience vary, being positive in certain low-income neighborhoods and negative or negligible in more affluent communities. These variations suggest that socioeconomic vulnerability operates in conjunction with resilience in a place-specific manner, shaped by local infrastructure, public services, and community capacity.

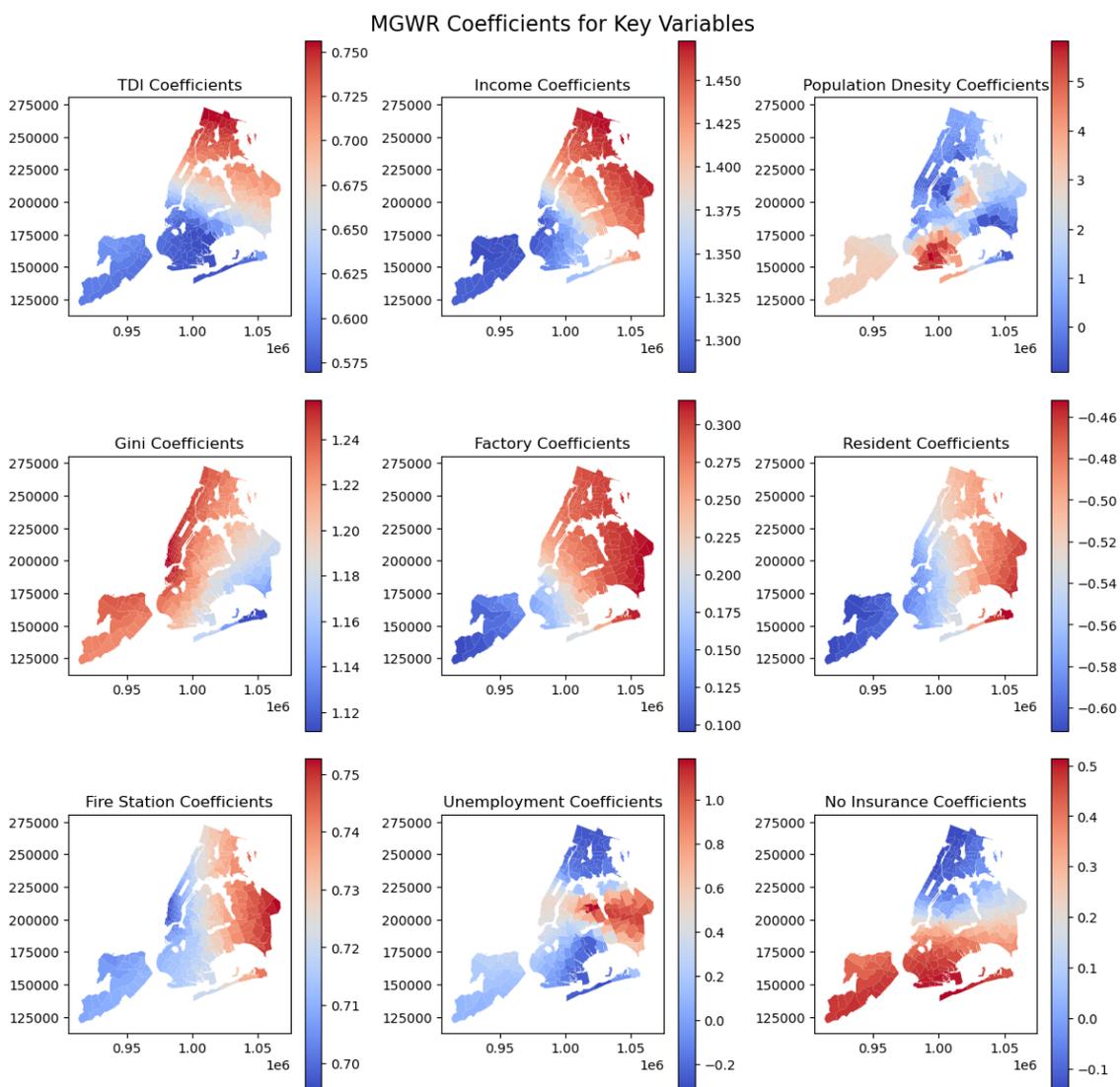


Figure 5.24: MGWR coefficient maps for key variables across NYC taxi zones.

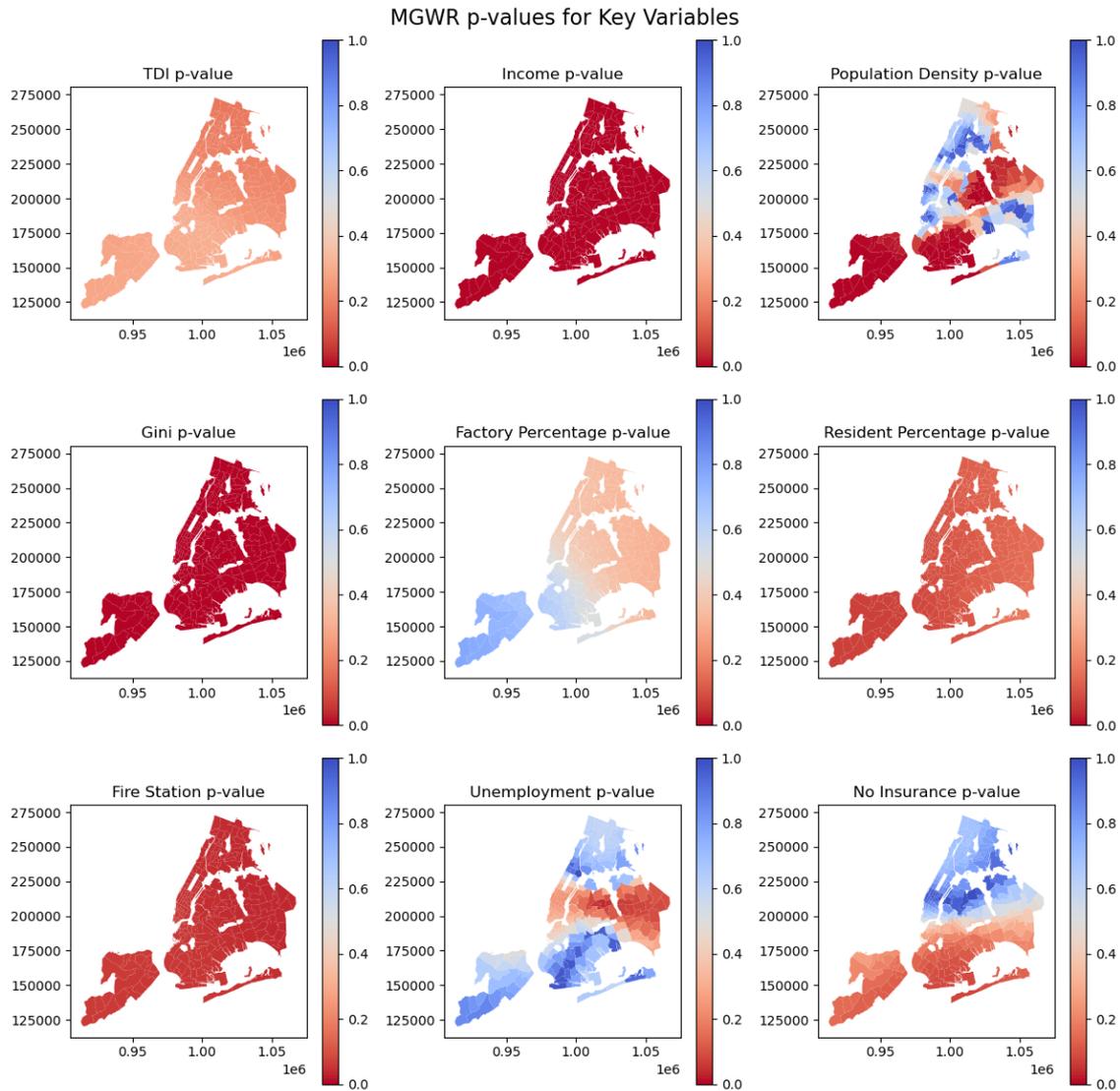


Figure 5.25: MGWR p-value maps for the same key variables, highlighting significant regions.

In conclusion, the multilevel modeling findings illustrate that global trends and local spatial heterogeneity influence how transportation diversity and mobility resilience are related in New York City. Although OLS and SAR models affirm the positive and consistent contributions of TDI, income, and access to emergency services to resilience, they also indicate the shortcomings of not accounting for spatial dependencies. MGWR, with the highest explanatory power, determines that certain variables, such as TDI and income, exhibit consistent, city-level effects. In contrast, others exhibit extremely variable effects across neighborhoods, including income inequality, land use, and social vulnerability indicators. These findings underscore that resilience is not uniform throughout the city but is locally mediated by socioeconomic context, infrastructure, and services.

## 6 Discussion

This chapter synthesizes the study's key findings regarding the three research questions while simultaneously outlining the methodological contributions made by the enhanced mobility resilience framework.

### 6.1 Impacts of Hurricane Ida on Transportation Modes

*RQ 1: How did Hurricane Ida impact different transportation modes in NYC, and which modes experienced the most significant disruptions?*

The research question concerned the extent and character of Hurricane Ida's impacts on the various transportation modes within New York City. While the event caused tremendous disruption throughout the city, the levels of impacts and recovery trajectories varied extensively by transportation systems.

The subway network was most severely affected. As is evident from the ridership time series and disruption detection output in Section 5.1.1 and Figure 5.3, the subway system saw an abrupt decline in usage on September 1st, corresponding to news reports of flooding and service suspension. This indicates the susceptibility of route-fixed, infrastructure-based networks to natural disasters. Service recovery developed gradually, with the long recovery time detected using the anomaly detection approach in Section 5.2.2, which verifies the low adaptive responsiveness of the subway system.

These empirical observations are mirrored in some official reports. On September 1st, 2021, for instance, the Manhattan 28th Street station was completely submerged as water poured in through ventilation grates and stairwells, rendering tracks and infrastructure useless. Likewise, the 145th Street 1 train station suffered flooding, putting it out of commission. These instances show how severe weather can immobilize even the most essential transit systems.

On the other hand, conventional for-hire taxis and large-volume for-hire taxis suffered modest losses and comparatively stable operations. As indicated by Figure 5.1, their use levels declined less steeply and bounced back more easily, due to reduced infrastructure dependence and real-time routing adaptability. These vehicles served as crucial alternatives when public transportation had been suspended.

Citi Bike patterns revealed a distinct response. Usage dropped significantly on the day of the storm but rebounded rapidly the following day, as shown in Figure 5.2. This rebound is further confirmed by the anomaly detection method in Section 5.2.2.4, which identified only brief periods of statistically significant disruption in most zones based on deviations from the Prophet-modeled mobility baseline.

A comparison of these modes shows that their resilience to disruption varies by system structure. Robustness in this case is affected by flexibility, redundancy, and infrastructure dependence. Transport modes with more spatial freedom and lower dependence on centralized infrastructure, i.e., bicycles and taxis, are more flexible in emergencies. These observations align with resilience theory, which emphasizes modularity, redundancy, and flexible operation of systems at different scales (Bruneau et al., 2003; Walker et al., 2004). In summary, the results emphasize the value of modal complementarity. There is no resilience in any transport mode in isolation; instead, resilience is fostered by functional diversity and the potential to change between systems during disruptions.

## 6.2 Spatial Variation and the Role of Transportation Diversity

*RQ 2: How did human mobility resilience and recovery time vary across different NYC neighborhoods or zones, and what role did transportation diversity play in mitigating these disruptions?*

The second research question examined spatial variation in mobility resilience throughout New York City and assessed how transportation diversity alleviated these variations. As evident in Section 5.2.2.4 and as indicated in Figure 5.17, the findings with respect to total resilience (R), recovery time (RT), loss of resilience (RL), and maximum deviation (MD) all revealed significant spatial inequality. High-resilience neighborhoods such as Midtown Manhattan, Downtown Brook-

lyn, and Long Island City experienced less disruption time and recovered faster. These regions scored highest on the Transportation Diversity Index (TDI), indicating a well-developed and strong multi-modal transport network.

By comparison, peripheral areas, comprising parts of the Bronx, Staten Island, and eastern Queens, experienced more extended service suspensions and took longer to recover. These areas tended to be more poorly ranked on the TDI. They were more reliant on single modes of transport, such as local buses or private automobiles, reflecting lower modal redundancy.

These spatial patterns are also corroborated in Section [5.4.2.4](#) by the MGWR results, which indicate that TDI has a consistent and statistically significant impact on resilience in almost all zones, with stable coefficients and p-values < 0.05. This attests that transport diversity is a structural buffer, particularly in areas where subway service was suspended. Zones served by well-connected bike or taxi substitutes recovered more quickly, illustrating the adaptive benefit of a diversified transport ecosystem.

### **6.3 Socio-Spatial Interactions and Urban Vulnerability**

*RQ 3: How do socio-economic, infrastructural, and environmental factors interact with transportation diversity to shape human mobility resilience in the face of Hurricane Ida?*

Extending the spatial analysis presented in Section [5.4.2](#), this analysis examines the relationship between transportation diversity and neighborhood-level infrastructural, socio-economic, and environmental conditions as joint determinants of human mobility resilience. Although the MGWR showed that the Transportation Diversity Index (TDI) consistently positively influences resilience outcomes in varying zones, the results indicate that this influence depends on a broader range of social and spatial factors.

One key insight is that income consistently boosts resilience, as shown across all models. Affluent neighborhoods had not only better access to varied transport options but also structural advantages like private car ownership, flexible work setups, and greater familiarity with e-mobility, which likely supported adaptation during disruption. These findings support past research showing that resilience is not purely infrastructural but closely tied to household resources and adaptive capacity.

In contrast, socio-economically disadvantaged areas, such as high unemployment or uninsured rates, did not tend to translate transport diversity into resilience. Although they may have physical infrastructure available, their ability to access or use alternatives appropriately appeared limited. This result aligns with findings in Section 5.2.2.4, where correlation and distributional analyses showed that high TDI values did not consistently correspond with high resilience scores, especially in socially vulnerable zones. The discrepancy highlights that the availability of infrastructure does not always mean functional access, in line with findings from transport justice research (Martens, 2016).

Land use and the presence of institutional services also affected resilience outcomes. Industrial land proportion was negatively associated with recovery, possibly due to lower population densities and more limited daily transit demand. Fire station density, in contrast, particularly in dense residential areas, was positively associated with resilience, suggesting that both logistical support and perceived safety form the foundations of post-disaster recovery. These patterns emphasize that urban form and service presence are active components of resilience processes, not inert backdrops.

Taken together, the findings suggest that diversity in transport provision is necessary but not sufficient. Its success is contingent upon its fit within the social, economic, and institutional context of particular neighborhoods. The notion of mobility resilience emerges not from physical networks alone but from the capacity of communities to access, navigate, and benefit from such networks during periods of duress. This view frames resilience as a relational and context-dependent phenomenon shaped by equity, infrastructure, and the demands of everyday life.

## 6.4 Performance of the Proposed Framework

This study proposes a new modeling framework for evaluating human mobility resilience, aiming to address critical limitations in existing approaches. While prior studies often rely on static baselines, focus on single transportation modes, and overlook intra-urban heterogeneity, the framework introduced here is dynamic, multi-modal, and spatially contextualized. It offers a more realistic and adaptable method for assessing how mobility systems respond to disruptive events such as Hurricane Ida.

First, the framework introduces dynamic baseline modeling using the Prophet time-series algorithm. This allows mode-specific baselines (for taxi, subway, and

bike systems) to capture both daily and weekly variation, providing a nuanced estimate of expected mobility under normal conditions. Disruptions are then identified based on deviations from these dynamic expectations, improving detection accuracy and temporal resolution.

Second, the framework adopts a multi-modal perspective by integrating subway, taxi, and Citi Bike data. This enables system-wide analysis that goes beyond isolated modal performance and captures substitution effects—how usage shifts across modes when one is disrupted. Such integration is essential for understanding the interdependencies within urban mobility systems.

Third, resilience is quantified across multiple dimensions, including overall resilience (R), resilience loss (RL), recovery time (RT), maximum deviation (MD), and resilience ratio (RA). This multi-metric approach moves beyond binary disruption/restoration assessments and provides a comprehensive understanding of both intensity and duration of impact.

Fourth, the model incorporates spatial heterogeneity through the use of Multiscale Geographically Weighted Regression (MGWR). By allowing explanatory variables to vary across space, the model uncovers how factors like transport diversity, income, and land use shape resilience outcomes differently across neighborhoods. These insights support more targeted and equitable planning interventions.

Overall, the enhanced model not only more accurately describes mobility progress in emergencies but also provides municipal planners with useful tools for identifying vulnerabilities, increasing infrastructure, and future-proofing the city.

## 6.5 Limitations

While the model provides insight, it has several limitations that need to be acknowledged. The research relied on publicly available mobility data, which were rich in detail but may not fully capture the response or informal activities associated with the disruption period. Analysis was similarly constrained to the short-term isolated impact of an intense weather event. It did not thus capture processes such as long-run behavioral adaptation or processes in recovering infrastructure. The socioeconomic data used in the spatial models were static and associated with the pre-disaster state without accounting for dynamic changes in vulnerability or resilience in the crisis period. Furthermore, within compos-

ite resilience measures, equal weightage may have been assigned to each mode, potentially at the expense of the varying importance or dependency on different modes during crisis periods.

# 7 Conclusion

This thesis explored how urban transportation systems respond to extreme weather disruptions, using New York City during Hurricane Ida as a case study. To support this analysis, it introduced an Improved Mobility Resilience Model that combines dynamic baseline forecasting (using the Prophet time-series algorithm), multi-modal integration of subway, taxi, and Citi Bike data, and spatial modeling. The model produced zone-level, time-sensitive indicators, offering new insights into how urban mobility adapted in the days leading up to, during, and following the storm.

## 7.1 General Conclusions

- **Mobility resilience varies by transportation mode and urban form.**

The analysis showed clear differences in how transit systems responded to Hurricane Ida. Subway systems, due to their fixed infrastructure and underground exposure, experienced the most severe and prolonged disruptions. In contrast, taxi services and Citi Bike exhibited faster recovery trajectories, demonstrating the greater adaptive capacity of decentralized and flexible mobility modes.

- **Transportation diversity contributes to resilience, but its effects are context-dependent.**

Areas with high transportation diversity, reflected by strong multi-modal integration, generally showed greater resilience. However, this effect was not consistent across all neighborhoods. In zones with high socio-economic and environmental vulnerability, even well-developed transport networks did not always lead to better outcomes. These areas lacked the resources or capacity to effectively use available modes, reinforcing that availability does not guarantee accessibility.

- **Spatial disparities in resilience reflect long-standing patterns of inequality.**

Central areas such as Midtown Manhattan and Downtown Brooklyn demonstrated high resilience, benefiting from dense infrastructure, multi-modal access, and institutional capacity. In contrast, peripheral neighborhoods, including Staten Island, the Bronx, and parts of eastern Queens, experienced longer recovery times and higher cumulative disruption. These patterns mirror structural spatial inequalities in mobility, access, and investment.

- **The proposed model enables more realistic and localized resilience assessment.**

By integrating dynamic baseline estimation using Prophet, multi-modal mobility data, and spatial regression models (SAR, SEM, MGWR), this study offers a robust framework for identifying disruptions and measuring recovery over time and across space. The approach moves beyond static, single-mode methods by accounting for temporal variation, modal complementarity, and spatial heterogeneity in resilience performance.

- **Equity should be central to resilience planning.**

The findings reinforce that transportation diversity alone is not sufficient. Its benefits depend on local socio-economic and infrastructural conditions. Building resilience thus requires not only investing in multi-modal systems but also ensuring that all communities, especially those in vulnerable or underserved areas, have the capacity to access and benefit from these systems during times of crisis.

*In sum, urban mobility resilience is not only a function of infrastructure and diversity, but of spatial equity, accessibility, and the capacity of neighborhoods to adapt under stress.*

## 7.2 Future Work

Future research can include real-time behavioral data regarding mobile traces, application usage, or social media, to better represent adaptive journey decisions in response to disruption. Increasing the temporal analysis window for effects persisting into the long run or occurring repeatedly would enable the detection of long-term patterns of adaptation and resilience. Examining testing in diverse urban environments would make the model more useful, in addition to showing how the response of resilience intervention needs to be designed specifically in terms of local geographic and social contexts. Incorporating flexible

modal weights and dynamic socio-economic and environmental parameters can enhance the explanatory ability of the model. At the same time, agent-based modeling can show how different groups of people react under mobility stress with different mobility constraints. These would enhance the present study's contributions and enable the creation of more responsive, equitable, and resilient mobility systems in response to mounting uncertainty in the climate.

# References

- Aghababaei, M. T., Costello, S. B., and Ranjitkar, P. 2021. [Measures to evaluate post-disaster trip resilience on road networks](#). *Journal of Transport Geography*, 95:103154.
- Ayesha, B., Jeewanthi, B., Chitraranjan, C., Perera, A. S., and Kumarage, A. S. 2021. [User Localization Based on Call Detail Records](#). *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11871 LNCS:411–423.
- Barreras, F. and Watts, D. J. 2024. [The exciting potential and daunting challenge of using GPS human-mobility data for epidemic modeling](#). *Nature Computational Science* 2024 4:6, 4(6):398–411.
- Bengtsson, L., Lu, X., Thorson, A., Garfield, R., and Schreeb, J. 2011. [Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti](#). *PLOS Medicine*, 8(8):e1001083.
- Bill de Blasio, . 2015. [OneNYC: A Plan for a Sustainable Global City A Plan for a Sustainable Global City on JSTOR](#). Technical report.
- Brown, J. T. 2014. [The Hurricane Sandy Rebuilding Strategy: In Brief](#).
- Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A., and Von Winterfeldt, D. 2003. [A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities](#). *Earthquake Spectra*, 19(4):733–752.
- Cats, O. 2024. [Identifying human mobility patterns using smart card data](#). *Transport Reviews*, 44(1):213–243.
- Chang, S. E. and Nojima, N. 2001. [Measuring post-disaster transportation system performance: the 1995 Kobe earthquake in comparative perspective](#). *Transportation Research Part A: Policy and Practice*, 35(6):475–494.

- Chelleri, L., Waters, J. J., Olazabal, M., and Minucci, G. 2015. [Resilience trade-offs: addressing multiple scales and temporal aspects of urban resilience](#). *Environment and Urbanization*, 27(1):181–198.
- Cimellaro, G. P., Reinhorn, A. M., and Bruneau, M. 2010. [Framework for analytical quantification of disaster resilience](#). *Engineering Structures*, 32(11):3639–3649.
- Cox, A., Prager, F., and Rose, A. 2011. [Transportation security and the role of resilience: A foundation for operational metrics](#). *Transport Policy*, 18(2):307–317.
- Cutter, S. L., Burton, C. G., and Emrich, C. T. 2010. [Disaster Resilience Indicators for Benchmarking Baseline Conditions](#). *Journal of Homeland Security and Emergency Management*, 7(1).
- Deng, H., Aldrich, D. P., Danziger, M. M., Gao, J., Phillips, N. E., Cornelius, S. P., and Wang, Q. R. 2021. [High-resolution human mobility data reveal race and wealth disparities in disaster evacuation patterns](#). *Humanities and Social Sciences Communications* 2021 8:1, 8(1):1–8.
- Donovan, B. and Work, D. B. 2015. [Using coarse GPS data to quantify city-scale transportation system resilience to extreme events](#). *Transportation Research Part C: Emerging Technologies*, 79:333–346.
- Fotheringham, A. S., Brunson, C., and Charlton, M. 2006. [Geographically weighted regression : the analysis of spatially varying relationships](#).
- Fotheringham, A. S., Yang, W., and Kang, W. 2017. [Multiscale Geographically Weighted Regression \(MGWR\)](#). *Annals of the American Association of Geographers*, 107(6):1247–1265.
- Gao, S., Yang, J.-A., Yan, B., Hu, Y., Janowicz, K., and McKenzie, G. 2014. [Detecting Origin-Destination Mobility Flows From Geotagged Tweets in Greater Los Angeles Area](#).
- Godschalk, D. R. and Baxter, S. [Urban Hazard Mitigation: Creating Resilient Cities](#).
- Griffith, D. A. 2009. [Spatially Autoregressive Models](#). *International Encyclopedia of Human Geography: Volume 1-12*, 1-12:10–396.
- Haraguchi, M., Nishino, A., Kodaka, A., Allaire, M., Lall, U., Kuei-Hsien, L., Onda, K., Tsubouchi, K., and Kohtake, N. 2022. [Human mobility data and](#)

- analysis for urban resilience: A systematic review. *Environment and Planning B: Urban Analytics and City Science*, 49(5):1507–1535.
- He, Z., Hu, Y., Duan, L. L., and Michailidis, G. 2024. Returners and explorers dichotomy in the face of natural hazards. *Scientific Reports*, 14(1):13184.
- Henry, D. and Emmanuel Ramirez-Marquez, J. 2012. Generic metrics and quantitative approaches for system resilience as a function of time. *Reliability Engineering & System Safety*, 99:114–122.
- Hsieh, C. H. and Feng, C. M. 2020. The highway resilience and vulnerability in Taiwan. *Transport Policy*, 87:1–9.
- Hu, T., Wang, S., She, B., Zhang, M., Huang, X., Cui, Y., Khuri, J., Hu, Y., Fu, X., Wang, X., Wang, P., Zhu, X., Bao, S., Guan, W., and Li, Z. 2021. Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges. *International Journal of Digital Earth*, 14(9):1126–1147.
- Hunter, R. F., Akaraci, S., Wang, R., Reis, R., Hallal, P. C., Pentland, S., Millett, C., Garcia, L., Thompson, J., Nice, K., Zapata-Diomed, B., and Moro, E. 2024. City mobility patterns during the COVID-19 pandemic: analysis of a global natural experiment. *The Lancet Public Health*, 9(11):e896–e906.
- Ilbeigi, M. 2019. Statistical process control for analyzing resilience of transportation networks. *International Journal of Disaster Risk Reduction*, 33:155–161.
- James, G., Witten, D., Hastie, T., Tibshirani, R., and Taylor, J. 2023. *An Introduction to Statistical Learning*.
- Kang, C., Ma, X., Tong, D., and Liu, Y. 2012. Intra-urban human mobility patterns: An urban morphology perspective. *Physica A: Statistical Mechanics and its Applications*, 391(4):1702–1717.
- Kaufman, S. M., Qing, C., Levenson, N., Hanson, M., Management, R. C. f. T. P., and Service, N. Y. U. R. F. W. G. S. o. P. 2012. Transportation during and after Hurricane Sandy.
- Kontou, E., Murray-Tuite, P., and Wernstedt, K. 2016. Commuter Adaptation in Response to Hurricane Sandy's Damage. *Natural Hazards Review*, 18(2):04016010.
- Kontou, E., Murray-Tuite, P., and Wernstedt, K. 2017. Duration of commute travel changes in the aftermath of Hurricane Sandy using accelerated failure time modeling. *Transportation Research Part A: Policy and Practice*, 100:170–181.

- Li, H., Calder, C. A., and Cressie, N. 2007. [Beyond Moran's I: Testing for Spatial Dependence Based on the Spatial Autoregressive Model](#). *Geographical Analysis*, 39(4):357–375.
- Loreti, S., Keiler, M., and Zischg, A. P. 2025. [A severe local flood and social events show a similar impact on human mobility](#). *npj Complexity* 2025 2:1, 2(1):1–18.
- Lu, P. and Stead, D. 2013. [Understanding the notion of resilience in spatial planning: A case study of Rotterdam, The Netherlands](#). *Cities*, 35:200–212.
- Martens, K. 2016. [Transport Justice : Designing fair transportation systems](#). *Transport Justice*.
- Mattsson, L. G. and Jenelius, E. 2015. [Vulnerability and resilience of transport systems – A discussion of recent research](#). *Transportation Research Part A: Policy and Practice*, 81:16–34.
- Meerow, S., Newell, J. P., and Stults, M. 2016. [Defining urban resilience: A review](#). *Landscape and Urban Planning*, 147:38–49.
- Meister, A., Mondal, A., Asmussen, K. E., Bhat, C., and Axhausen, K. W. 2022. [Modeling Urban Mode Choice Behavior During the COVID-19 Pandemic in Switzerland Using Mixed Multiple Discrete-Continuous Extreme Value Models](#). *Transportation Research Record: Journal of the Transportation Research Board*, page 036119812210895.
- Mirjalili, R., Barati, H., and Yazici, A. 2023. [Resilience Analysis of New York City Transportation Network After Snow Storms](#). *Transportation Research Record*, 2677(1):694–707.
- Montazer, S. and Young, M. 2024. [The Long Haul Home: The Relationship between Commuting Distance, Work Hours, Work-to-Family Conflict, and Psychological Distress](#). *Socius*, 10.
- Moran, P. A. P. 1950. [Notes on Continuous Stochastic Phenomena](#). *Biometrika*, 37(1/2):17.
- Murray-Tuite, P. M. 2006. [A comparison of transportation network resilience under simulated system optimum and user equilibrium conditions](#). *Proceedings - Winter Simulation Conference*, pages 1398–1405.
- Otsuka, R. P., Work, D. B., and Song, J. 2016. [Estimating post-disaster traffic conditions using real-time data streams](#). *Structure and Infrastructure Engineering*, 12(8):904–917.

- Rahimi-Golkhandan, A., Garvin, M. J., and Brown, B. L. 2019. [Characterizing and measuring transportation infrastructure diversity through linkages with ecological stability theory](#). *Transportation Research Part A: Policy and Practice*, 128:114–130.
- Rahimi-Golkhandan, A., Garvin, M. J., and Wang, Q. 2021. [Assessing the Impact of Transportation Diversity on Postdisaster Intraurban Mobility](#). *Journal of Management in Engineering*, 37(1):04020106.
- Ren, Z., Zhu, J., Education, U. o. N. C. a. C. C. f. A. M. M. S., , , Engineering, U. o. C. D. o. C., and Environmental, . 2022. [Disaster Resilience Through Diverse Evacuation and Emergency Transportation Systems \(Phase II\) Final Report](#).
- Rey, S. J. and Franklin, R. S. 2022. [Handbook of spatial analysis in the social sciences](#). *Handbook of Spatial Analysis in the Social Sciences*, pages 1–588.
- Rodrigue, J. P. 2024. [The geography of transport systems](#). *The Geography of Transport Systems*, pages 1–402.
- Rose, A. 2004. [Defining and measuring economic resilience to disasters](#). *Disaster Prevention and Management: An International Journal*, 13(4):307–314.
- Roy, A. and Kar, B. 2022. [Effect of Social Vulnerability on Taxi Trip Times during Hurricane Sandy](#). *Findings*, 2022.
- Roy, K. C., Cebrian, M., and Hasan, S. 2019. [Quantifying human mobility resilience to extreme events using geo-located social media data](#). *EPJ Data Science* 2019 8:1, 8(1):1–15.
- Santana, C., Botta, F., Barbosa, H., Privitera, F., Menezes, R., and Di Clemente, R. 2023. [COVID-19 is linked to changes in the time–space dimension of human mobility](#). *Nature Human Behaviour* 2023 7:10, 7(10):1729–1739.
- Shen, Z., Ji, C., and Lu, S. 2024. [Transportation network resilience response to the spatial feature of hazards](#). *Transportation Research Part D: Transport and Environment*, 128:104121.
- Sun, J., Chow, A. C., and Madanat, S. M. 2020a. [Multimodal transportation system protection against sea level rise](#). *Transportation Research Part D: Transport and Environment*, 88:102568.
- Sun, W., Bocchini, P., and Davison, B. D. 2020b. [Resilience metrics and measurement methods for transportation infrastructure: the state of the art](#). *Sustainable and Resilient Infrastructure*, 5(3):168–199.

- Tang, J., Liu, F., Wang, Y., and Wang, H. 2015. [Uncovering urban human mobility from large scale taxi GPS data.](#) *Physica A: Statistical Mechanics and its Applications*, 438:140–153.
- Gun, J. P., Pel, A. J., and Arem, B. 2016. [A general activity-based methodology for simulating multimodal transportation networks during emergencies.](#) *European Journal of Transport and Infrastructure Research*, 16(3):490–511.
- Veloso, M., Phithakkitnukoon, S., and Bento, C. 2011. [Urban mobility study using taxi traces.](#) *TDMA'11 - Proceedings of the 2011 International Workshop on Trajectory Data Mining and Analysis*, pages 23–30.
- Walker, B., Holling, C. S., Carpenter, S. R., and Kinzig, A. 2004. [Resilience, adaptability and transformability in social-ecological systems.](#) *Ecology and Society*, 9(2):5.
- Wan, Z., Lang, Q., Zhang, Y., Zhang, J., Chen, Y., Liu, G., and Liu, H. 2025. [Improving the resilience of urban transportation to natural disasters: the case of Changchun, China.](#) *Scientific Reports 2025 15:1*, 15(1):1–22.
- Wang, Q., Phillips, N. E., Small, M. L., and Sampson, R. J. 2018. [Urban mobility and neighborhood isolation in America's 50 largest cities.](#) *Proceedings of the National Academy of Sciences of the United States of America*, 115(30):7735–7740.
- Wang, Y., Rahimi-Golkhandan, A., Chen, C., Taylor, J. E., and Garvin, M. J. 2019. [Measuring the Impact of Transportation Diversity on Disaster Resilience in Urban Communities: Case Study of Hurricane Harvey in Houston, TX.](#) *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2019*, pages 555–562.
- Xia, F., Wang, J., Kong, X., Wang, Z., Li, J., and Liu, C. 2018. [Exploring Human Mobility Patterns in Urban Scenarios: A Trajectory Data Perspective.](#) *IEEE Communications Magazine*, 56(3):142–149.
- Xu, Z. and Chopra, S. S. 2023. [Interconnectedness enhances network resilience of multimodal public transportation systems for Safe-to-Fail urban mobility.](#) *Nature Communications*, 14(1):1–11.
- Yang, Y., Heppenstall, A., Turner, A., Comber, A., Heppenstall@leeds, A. J. A., Uk, A. H., Turner@leeds, A. G. D. A., Uk, . A. T., Comber@leeds, A. A., and Uk, A. C. 2019. [Who, Where, Why and When? Using Smart Card and Social](#)

- Media Data to Understand Urban Mobility. *ISPRS International Journal of Geo-Information* 2019, Vol. 8, Page 271, 8(6):271.
- Yang, Z., Liu, X., Wang, J., Yan, X., Shen, R., and Huo, Z. 2025. Assessment on resilience of urban agglomeration transportation system considering passenger choice and load-capacity factor. *Reliability Engineering & System Safety*, 253:110527.
- Yao, Q., Shan, X., Li, M., and Wang, J. 2024a. The Impact of Floods on the Mobility of Automobile Commuters in Shanghai Under Climate Change: Yao et al. The Impact of Floods on the Mobility of Automobile Commuters in Shanghai. *International Journal of Disaster Risk Science*, 15(6):986–1000.
- Yao, Y., Liang, L., Zhang, Y., Wang, Y., Hu, Z., Fan, Y., Guan, Q., Jiang, R., and Shibasaki, R. 2024b. Resilience Patterns of Multiscale Human Mobility Under Extreme Rainfall Events Using Massive Individual Trajectory Data. *Annals of the American Association of Geographers*.
- Yuan, A., Spira-Cohen, A., Olson, C., and Lane, K. 2024. Immediate Injury Deaths Related to the Remnants From Hurricane Ida in New York City, September 1-2, 2021. *Disaster Medicine and Public Health Preparedness*, 18(e55):e55.
- Zhao, K., Tarkoma, S., Liu, S., and Vo, H. 2016. Urban human mobility data mining: An overview. *Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016*, pages 1911–1920.
- Zheng, Y., Wang, S., Liu, L., Aloisi, J., and Zhao, J. 2024. Impacts of remote work on vehicle miles traveled and transit ridership in the USA. *Nature Cities* 2024 1:5, 1(5):346–358.
- Zhu, K., Cheng, Z., and Wang, J. 2024. Measuring Chinese mobility behaviour during COVID-19 using geotagged social media data. *Humanities and Social Sciences Communications* 2024 11:1, 11(1):1–12.

# 8 Appendix

Table 8.1: Prophet cross-validation summary by taxizone (Part 1)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
1	0.952	33.678	27.343	-	-	-	-	-	-
2	0.96	0.497	0.318	-	-	-	-	-	-
3	0.972	19.544	14.981	0.965	357.675	274.223	-	-	-
4	0.931	53.186	37.659	-	-	-	0.917	32.044	24.939
5	0.924	4.176	3.32	-	-	-	-	-	-
6	0.945	7.433	5.85	-	-	-	-	-	-
7	0.938	91.682	68.397	0.929	491.672	425.377	0.924	37.782	29.878
8	0.991	3.079	2.252	-	-	-	0.968	7.568	5.762
9	0.945	10.136	8.231	-	-	-	-	-	-
10	0.952	28.087	21.639	-	-	-	-	-	-
11	0.959	11.394	8.967	-	-	-	-	-	-
12	0.919	7.132	5.794	-	-	-	0.974	7.579	5.776
13	0.938	102.271	78.539	-	-	-	0.979	74.173	59.676
14	0.952	52.604	39.084	0.988	411.341	337.693	0.973	-	-
15	0.945	8.601	6.837	-	-	-	-	-	-
16	0.924	21.662	16.57	-	-	-	-	-	-
17	0.938	96.984	73.187	0.959	951.404	833.147	0.931	35.645	29.191
18	0.986	47.481	37.461	0.976	1201.285	1057.84	0.957	-	-
19	0.986	8.634	6.975	-	-	-	-	-	-
20	0.979	27.167	21.591	-	-	-	0.947	-	-
21	0.966	23.961	18.641	0.952	162.892	139.663	-	-	-
22	0.959	28.386	21.962	0.964	761.435	659.446	-	-	-
23	0.924	14.263	11.51	-	-	-	-	-	-
24	0.966	20.36	15.776	0.988	144.54	115.687	0.944	12.541	9.571
25	0.952	51.996	39.633	0.96	1684.294	1359.485	0.93	35.725	28.655
26	0.966	85.98	67.732	0.976	1784.834	1497.125	0.964	-	-
27	0.899	13.369	8.778	-	-	-	-	-	-
28	0.972	18.757	14.476	0.988	316.391	267.816	-	-	-
29	0.952	23.016	18.056	1.0	1707.766	1441.095	-	-	-
30	1	3.37	1.889	1.0	251.196	186.814	-	-	-
31	0.993	13.815	10.522	-	-	-	-	-	-
32	0.972	23.701	19.004	0.976	140.607	118.222	-	-	-
33	0.931	56.735	44.178	0.949	823.872	729.123	0.965	35.885	28.713
34	0.959	13.567	10.409	-	-	-	0.928	13.891	10.839
35	0.966	58.598	44.803	0.97	1248.367	1074.395	-	-	-

Table 8.2: Prophet cross-validation summary by taxizone (Part 2)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
36	0.931	101.317	75.069	0.99	2219.026	1930.831	0.924	28.445	22.6
37	0.952	168.384	118.892	1.0	2641.525	2330.887	0.924	45.39	36.82
38	0.945	10.659	8.232	-	-	-	-	-	-
39	0.924	69.768	51.159	0.941	1608.414	1160.764	-	-	-
40	0.924	32.221	24.41	0.988	94.416	76.713	0.942	19.487	15.232
41	0.966	67.455	52.512	0.953	711.245	576.345	0.972	37.903	30.143
42	0.966	94.941	71.655	1.0	531.006	464.331	0.938	27.74	22.303
43	0.993	61.193	47.329	0.97	1583.651	1410.436	0.993	102.145	76.463
44	0.957	6.093	4.719	-	-	-	-	-	-
45	0.917	29.094	22.317	0.988	536.909	407.514	0.931	25.685	20.485
46	0.986	6.412	4.903	-	-	-	-	-	-
47	0.993	35.573	28.543	-	-	-	0.964	5.437	4.234
48	0.931	133.158	102.008	0.941	343.515	278.611	0.924	72.352	57.532
49	0.966	64.584	48.722	1.0	2736.433	2557.469	0.917	33.779	27.533
50	0.931	78.953	62.026	-	-	-	0.945	35.327	28.634
51	0.979	29.728	23.407	0.976	151.372	120.033	-	-	-
52	0.959	16.099	12.496	-	-	-	0.969	9.576	7.584
53	0.979	13.958	10.911	-	-	-	-	-	-
54	0.958	7.714	6.035	-	-	-	0.964	15.943	12.304
55	0.931	44.172	32.284	1.0	1317.205	1087.606	-	-	-
56	0.986	25.553	19.343	-	-	-	-	-	-
57	0.979	3.697	2.944	-	-	-	-	-	-
58	0.924	5.012	3.836	-	-	-	-	-	-
59	0.958	6.366	4.349	-	-	-	0.965	2.673	2.031
60	0.979	24.953	19.515	0.988	139.548	121.485	0.957	4.762	3.643
61	0.945	142.776	102.586	0.98	2026.408	1839.06	0.924	26.405	21.695
62	0.966	38.53	29.981	0.965	1018.185	831.385	0.985	6.86	5.256
63	0.979	32.901	24.821	0.919	1594.814	1231.319	-	-	-
64	0.931	6.933	5.547	-	-	-	-	-	-
65	0.952	46.023	34.979	1.0	1506.872	1271.574	0.979	23.078	18.066
66	0.903	44.671	34.909	0.929	1137.966	913.025	0.944	38.455	30.379
67	0.945	16.937	13.142	-	-	-	-	-	-
68	0.91	134.159	106.35	0.964	207.356	174.898	0.952	142.818	114.53
69	0.986	50.568	39.487	1.0	336.514	246.618	0.966	9.559	7.516
70	0.959	17.503	13.398	-	-	-	-	-	-

Table 8.3: Prophet cross-validation summary by taxizone (Part 3)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
71	0.966	46.936	34.053	1.0	82.737	67.55	-	-	-
72	0.959	51.493	38.534	-	-	-	-	-	-
73	0.966	9.934	7.954	-	-	-	-	-	-
74	1	69.491	51.44	0.941	263.271	220.062	0.952	29.157	22.893
75	0.979	82.18	62.097	0.953	349.683	295.602	0.952	31.016	24.417
76	0.959	88.826	69.165	0.949	1876.988	1629.719	-	-	-
77	0.945	31.995	25.295	0.976	248.486	199.317	-	-	-
78	0.979	41.769	31.67	0.988	177.175	149.66	-	-	-
79	0.924	298.339	212.21	1.0	298.721	261.304	0.959	146.132	120.089
80	0.903	148.5	110.03	1.0	985.039	868.044	0.924	40.188	32.772
81	0.966	24.619	18.881	0.941	74.672	58.738	-	-	-
82	0.959	50.336	38.518	1.0	809.267	653.485	-	-	-
83	0.986	16.243	12.881	-	-	-	-	-	-
84	0.959	6.747	5.164	-	-	-	-	-	-
85	0.966	29.155	21.899	0.941	38.264	30.256	-	-	-
86	0.952	17.545	13.916	0.976	693.636	592.122	-	-	-
87	0.897	75.79	59.766	0.97	660.874	591.346	0.931	53.225	42.688
88	0.952	32.228	25.884	0.917	1480.848	1242.648	0.985	13.039	10.093
89	0.938	64.831	48.66	0.949	4697.932	4227.533	0.953	-	-
90	0.938	107.345	82.151	1.0	472.425	411.793	0.924	47.635	36.771
91	0.966	43.614	32.441	-	-	-	-	-	-
92	0.959	32.284	25.09	1.0	143.169	110.471	-	-	-
93	0.828	30.285	21.642	0.99	255.426	216.983	-	-	-
94	1	24.72	19.546	0.929	208.124	173.792	0.976	-	-
95	0.959	53.793	40.386	0.988	322.715	268.552	-	-	-
96	0.958	6.214	4.249	-	-	-	-	-	-
97	0.972	70.464	50.732	0.99	1148.641	961.21	0.966	47.776	38.66
98	0.966	9.38	7.479	-	-	-	-	-	-
99	0.879	1.008	0.763	-	-	-	-	-	-
100	0.966	77.025	62.603	0.96	2431.08	2147.578	0.89	41.959	32.527
101	0.951	7.421	5.994	-	-	-	-	-	-
102	0.986	19.845	15.535	-	-	-	-	-	-
105	1	0.266	0.262	-	-	-	-	-	-
106	0.91	24.726	19.011	0.952	1482.706	1221.977	0.935	12.694	10.057
107	0.924	127.626	99.017	0.965	227.813	197.121	0.917	47.869	37.576

Table 8.4: Prophet cross-validation summary by taxizone (Part 4)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
108	0.924	14.982	11.74	0.929	990.583	824.228	-	-	-
109	0.952	8.782	6.76	-	-	-	-	-	-
110	1	0.314	0.215	-	-	-	-	-	-
111	0.941	2.17	1.602	-	-	-	0.957	3.136	2.573
112	0.931	94.353	72.219	0.98	392.1	285.705	0.945	56.356	45.223
113	0.897	83.609	65.735	0.889	1064.889	934.231	0.917	69.349	55.315
114	0.89	115.369	86.45	0.828	251.53	209.023	0.91	42.09	33.339
115	0.972	8.446	6.655	-	-	-	-	-	-
116	0.966	47.977	35.879	0.899	1688.398	1398.81	0.945	15.241	11.955
117	0.959	19.549	14.615	0.98	1027.844	900.749	-	-	-
118	0.91	10.898	8.579	-	-	-	-	-	-
119	0.986	39.798	30.536	-	-	-	0.958	5.207	4.076
120	0.979	3.777	2.72	-	-	-	0.967	4.235	3.124
121	0.979	20.006	15.578	-	-	-	-	-	-
122	0.966	11.571	8.96	-	-	-	-	-	-
123	0.972	36.754	28.417	1.0	221.417	195.204	-	-	-
124	0.972	12.095	9.357	-	-	-	-	-	-
125	0.869	57.676	44.598	0.899	286.217	245.797	0.972	51.84	41.865
126	0.986	30.503	23.392	0.929	354.268	287.067	0.964	4.699	3.521
127	0.945	46.121	34.398	0.988	1254.951	1003.886	0.978	-	-
128	0.965	4.885	3.803	-	-	-	0.983	-	-
129	0.945	64.346	48.332	0.939	1241.186	1056.889	-	-	-
130	0.945	49.405	38.568	0.893	1035.056	875.245	-	-	-
131	0.972	14.947	11.799	-	-	-	-	-	-
132	0.897	169.504	137.666	1.0	733.705	545.312	-	-	-
133	0.938	19.265	14.667	1.0	659.381	395.905	0.984	-	-
134	0.931	20.672	16.036	0.988	369.403	320.517	-	-	-
135	0.945	15.417	12.172	-	-	-	-	-	-
136	0.966	27.877	21.686	-	-	-	0.946	-	-
137	0.924	108.25	79.504	-	-	-	0.952	51.531	41.125
138	0.848	179.91	145.951	-	-	-	-	-	-
139	0.931	16.431	12.944	-	-	-	-	-	-
140	0.979	113.483	88.684	-	-	-	0.965	13.166	10.197
141	0.945	116.482	90.439	-	-	-	0.931	54.421	41.85
142	0.945	103.908	82.486	0.918	374.174	303.6	0.972	30.963	23.291

Table 8.5: Prophet cross-validation summary by taxizone (Part 5)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
143	0.938	69.964	54.482	1.0	62.004	50.593	0.959	54.672	43.55
144	0.903	92.533	69.936	0.96	1668.921	1518.032	0.952	52.816	41.904
145	0.931	77.627	57.484	0.949	3368.168	2844.235	0.945	34.892	27.158
146	0.931	25.793	19.133	1.0	369.611	292.201	0.938	15.506	11.921
147	0.979	24.073	19.097	0.909	406.607	336.521	0.971	4.349	3.347
148	0.903	239.342	167.843	0.988	1149.546	1000.769	0.959	84.061	69.05
149	0.959	22.857	18.121	0.976	924.88	810.236	–	–	–
150	0.924	15.742	12.02	–	–	–	–	–	–
151	0.959	40.309	31.187	1.0	131.099	111.236	0.952	33.738	26.624
152	0.917	25.971	19.908	0.953	621.209	505.157	0.952	13.5	10.603
153	0.959	11.133	8.797	0.976	122.506	99.411	–	–	–
154	0.993	4.476	3.211	–	–	–	–	–	–
155	0.952	24.984	19.292	–	–	–	–	–	–
156	0.897	11.716	9.113	–	–	–	–	–	–
157	0.959	30.191	22.676	–	–	–	–	–	–
158	0.897	123.82	95.112	–	–	–	0.958	77.432	63.095
159	0.986	48.384	38.169	0.965	273.536	222.943	0.917	8.361	6.69
160	0.966	19.341	14.85	–	–	–	–	–	–
161	0.959	224.236	179.521	0.98	503.743	412.28	0.907	35.428	27.105
162	0.903	186.097	144.734	0.918	234.794	193.199	0.915	36.209	26.232
163	0.938	136.294	112.341	1.0	306.28	243.39	0.945	56.226	43.677
164	0.883	139.663	112.064	0.939	808.155	658.481	0.945	32.544	25.464
165	0.986	37.881	29.837	0.939	1550.615	1394.564	–	–	–
166	0.966	42.207	32.352	0.98	550.554	469.984	0.952	17.289	13.386
167	0.966	39.581	31.707	0.964	195.757	166.588	0.959	5.925	4.5
168	0.986	58.88	45.114	0.988	1085.517	925.41	0.938	12.746	10.052
169	0.993	54.557	43.197	0.949	622.449	534.627	0.972	4.984	3.867
170	0.924	165.861	133.954	–	–	–	0.91	68.846	51.214
171	0.979	18.063	14.271	–	–	–	–	–	–
172	0.931	6.857	5.262	–	–	–	–	–	–
173	1	23.058	17.988	0.871	394.464	312.452	–	–	–
174	0.972	45.838	35.334	0.919	480.251	372.628	0.937	–	–
175	0.945	7.373	5.705	–	–	–	–	–	–
176	0.965	5.037	3.935	–	–	–	–	–	–
177	0.972	42.538	32.265	0.98	1513.253	1302.36	0.964	4.498	3.518

Table 8.6: Prophet cross-validation summary by taxizone (Part 6)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
178	1	15.331	11.949	0.976	427.315	370.997	-	-	-
179	0.966	31.644	24.36	-	-	-	0.952	15.234	12.121
180	0.966	13.068	10.169	0.976	1038.891	856.291	-	-	-
181	0.897	99.56	74.465	0.859	2963.395	2539.088	0.931	62.909	51.618
182	0.979	20.31	15.631	-	-	-	-	-	-
183	0.938	13.143	10.277	0.929	343.583	278.117	-	-	-
184	0.903	17.576	11.954	-	-	-	-	-	-
185	0.986	31.967	24.829	0.98	319.477	257.55	-	-	-
186	0.897	123.123	98.084	0.857	2238.29	1907.883	0.938	48.766	37.907
187	0.972	7.498	5.848	-	-	-	-	-	-
188	0.959	76.565	57.566	0.929	517.826	422.24	0.941	8.458	6.475
189	0.931	46.224	33.112	0.952	566.542	499.529	0.95	26.421	21.277
190	0.972	22.39	15.531	0.929	63.855	55.863	0.965	28.317	21.804
191	0.938	21.999	17.165	-	-	-	-	-	-
192	0.952	12.951	9.807	-	-	-	-	-	-
193	0.972	14.427	11.082	1.0	279.064	227.752	0.958	8.915	6.67
194	1	7.696	5.599	-	-	-	0.964	11.337	8.357
195	0.938	24.396	18.822	-	-	-	0.985	14.899	11.475
196	0.972	18.951	14.646	0.988	81.111	68.447	-	-	-
197	0.945	44.875	33.315	0.798	1615.057	1376.279	-	-	-
198	0.931	64.998	49.108	0.99	1925.093	1743.877	0.917	10.935	8.665
199	1	0	0	-	-	-	-	-	-
200	0.993	18.457	13.767	-	-	-	-	-	-
201	0.883	19.54	14.04	0.99	1158.793	1048.44	-	-	-
202	0.952	7.672	5.851	1.0	543.23	454.757	0.984	13.684	10.183
203	0.938	14.197	11.214	-	-	-	-	-	-
204	0.943	4.296	3.362	-	-	-	-	-	-
205	0.972	34.213	26.604	-	-	-	-	-	-
206	0.938	19.176	15.055	-	-	-	-	-	-
207	0.951	6.974	5.485	-	-	-	-	-	-
208	0.938	32.432	24.798	-	-	-	-	-	-
209	0.945	26.963	21.326	0.94	520.147	439.993	0.912	10.652	8.43
210	0.959	30.264	23.521	0.988	781.819	699.136	-	-	-
211	0.917	79.905	61.727	1.0	1835.158	1611.205	0.944	24.993	19.847
212	0.966	25.069	19.324	0.918	367.1	296.333	-	-	-

Table 8.7: Prophet cross-validation summary by taxizone (Part 7)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
213	0.993	41.664	32.466	-	-	-	-	-	-
214	0.966	10.357	8.293	-	-	-	-	-	-
215	0.959	26.541	20.5	-	-	-	-	-	-
216	0.959	51.382	37.234	0.899	1384.3	1087.464	-	-	-
217	0.952	35.011	27.57	0.988	794.373	610.603	0.918	5.022	3.652
218	0.979	18.127	14.076	-	-	-	-	-	-
219	0.945	15.887	12.569	-	-	-	-	-	-
220	0.952	27.64	21.368	0.899	2347.808	1997.74	-	-	-
221	0.924	13.492	10.747	-	-	-	-	-	-
222	0.966	12.131	9.436	-	-	-	-	-	-
223	0.959	47.238	36.752	0.824	948.642	785.419	0.916	22.038	17.404
224	0.876	22.591	16.825	-	-	-	0.903	33.962	26.396
225	0.966	98.562	71.602	1.0	809.241	649.405	0.897	16.051	12.998
226	0.945	49.441	37.456	0.97	650.558	561.433	0.963	6.054	4.576
227	0.979	24.051	18.654	0.952	337.276	273.149	0.94	4.791	3.93
228	0.979	54.101	41.336	0.905	3866.573	3199.378	0.95	13.505	10.539
229	0.938	101.143	81.138	1.0	125.623	104.244	0.91	32.006	24.226
230	0.793	152.214	121.733	0.878	926.681	772.301	0.924	34.266	26.286
231	0.924	142.829	113.92	0.96	4518.698	4000.964	0.917	84.959	67.769
232	0.841	71.646	57.245	0.964	128.012	102.324	0.952	57.81	46.209
233	0.938	85.464	68.06	-	-	-	0.951	42.847	33.202
234	0.91	164.507	131.468	0.878	2970.252	2519.64	0.89	73.502	56.406
235	0.986	45.256	35.615	0.906	270.201	203.953	0.944	-	-
236	0.993	185.376	143.391	0.953	186.316	154.243	0.951	27.773	21.0
237	0.993	199.26	157.553	0.97	771.607	589.391	0.919	18.586	13.355
238	0.959	91.663	71.987	0.952	459.036	368.408	0.952	32.35	25.218
239	0.883	104.528	81.33	0.869	260.737	221.232	0.979	53.45	41.594
240	0.993	9.027	6.583	-	-	-	-	-	-
241	0.993	36.288	28.556	0.918	247.49	200.691	0.97	-	-
242	0.966	50.126	39.281	0.906	83.714	65.616	-	-	-
243	0.972	58.399	44.598	0.879	1747.766	1504.694	0.965	10.648	8.319
244	0.966	91.108	68.756	0.96	1588.892	1361.673	0.986	13.93	10.713
245	0.959	9.169	7.185	-	-	-	-	-	-
246	0.924	141.985	114.171	-	-	-	0.972	80.357	64.154
247	0.986	56.21	41.881	0.98	859.936	729.617	0.951	12.389	9.698

Table 8.8: Prophet cross-validation summary by taxizone (Part 8)

taxizone	taxi_accuracy	taxi_rmse	taxi_mae	sub_accuracy	sub_rmse	sub_mae	bike_accuracy	bike_rmse	bike_mae
248	0.993	24.835	19.313	1.0	133.141	110.351	–	–	–
249	0.89	185.211	138.05	1.0	1409.2	1203.464	0.945	73.43	60.451
250	0.986	25.315	20.037	0.988	255.201	215.901	–	–	–
251	0.924	8.59	6.81	–	–	–	–	–	–
252	0.959	11.974	9.292	–	–	–	–	–	–
253	0.872	5.531	3.313	–	–	–	–	–	–
254	0.986	39.123	29.977	1.0	242.389	207.709	–	–	–
255	0.903	130.26	100.076	1.0	341.322	265.555	0.924	85.464	69.924
256	0.931	113.429	84.132	0.988	335.96	288.905	0.958	53.573	43.6
257	0.91	17.274	13.106	0.905	249.722	203.348	0.956	7.916	6.077
258	0.959	29.011	22.35	0.917	2205.287	1875.237	–	–	–
259	0.979	28.566	21.852	0.976	124.278	95.976	–	–	–
260	0.959	36.666	28.522	0.96	601.737	498.859	0.939	3.887	2.922
261	0.959	41.318	33.574	0.919	3960.417	3197.255	0.923	25.646	20.394
262	0.966	62.812	49.049	–	–	–	0.957	16.774	13.132
263	0.945	90.077	69.182	–	–	–	0.945	33.654	25.907
264	0.889	17.753	14.038	–	–	–	–	–	–
265	0.91	185.838	141.792	–	–	–	–	–	–
103	–	–	–	–	–	–	1.0	–	–

## 9 Personal Declaration

I declare that the submitted thesis is the result of my independent work. All external sources are explicitly acknowledged in the thesis.

---

Mengqi Li

---

Mengqi LI

Zurich, 15.04.2025