Center for Advanced Multimodal Mobility

Solutions and Education

Project ID: 2019 Project 10

Are Transportation Network Companies Synergistic with Other Shared Ride Mode Offerings? An Exploratory Analysis of Demand Data from NYC Utilizing High Resolution Spatiotemporal Models

Final Report

by

Nalini Ravishanker (ORCID ID: https://orcid.org/0000-0002-2028-4771)
Professor, Department of Statistics
University of Connecticut, Storrs CT 06269
Phone: 1-860-486-4760; Email: nalini.ravishanker@uconn.edu

Karthik Konduri (ORCID ID: https://orcid.org/0000-0003-2788-9455)
Research Scientist
University of Connecticut, Storrs CT 06269
Phone: 1-860-486-2992; Email: karthik.konduri@uconn.edu

for

Center for Advanced Multimodal Mobility Solutions and Education
(CAMMSE @ UNC Charlotte)
The University of North Carolina at Charlotte
9201 University City Blvd
Charlotte, NC 28223

September 2019

ACKNOWLEDGEMENTS

This project was funded by the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE @ UNC Charlotte), one of the Tier I University Transportation Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT), under the FAST Act. The authors are also very grateful for all of the time and effort spent by DOT and industry professionals to provide project information that was critical for the successful completion of this study.

DISCLAIMER

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation University Transportation Centers Program in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the U.S. Government. This report does not constitute a standard, specification, or regulation.

Table of Contents

List of Figures - Content

List of Figures - Appendices

List of Tables

EXCUTIVE SUMMARY

Spurred by technological advances, transportation networks and the mobility offerings for moving people and goods are undergoing transformative and significant changes. A significant operator in the area of moving people is transportation network companies (TNC), they are also commonly referred to by the name of the offering as dynamic ridesharing or ridesourcing. *A research need is to comprehensively understand the impacts of TNCs so that transportation systems can be planned and implemented, that effectively respond to changes it brings.* A unique feature of TNCs is the ease, efficiency, and effectiveness with which such services can be accessed and consumed, leveraged by technology and innovation. It is important to understand the demand for each of these services individually and to explore the interplay between these services so that policies and planning actions can be implemented to best promote these services and alleviate any negative impacts.

Our research consisted of a comprehensive exploration of all shared modes (subway, taxi, TNC, bikeshare) in a multivariate framework over multiple years, including incorporating the long-term patterns and incorporating the effects of short term shocks. This exploration was done using dynamic compositional models for time series, the data being aggregated across all of NY City, and would enable informed planning and operations decisions that positively impact all offerings within the shared mode landscape. Details are presented in the manuscript Toman et al. (2019). The next step in our research consisted of exploring the presence of spatial associations at taxi zone level in NYC, for which a comprehensive statistical analysis is scant in the ridesourcing literature. Together, the setup and outcomes of our research will be informative for building and estimating fine-scale spatio-temporal models for characterizing the existing system, as well as for short-term and long-term demand forecasting purposes.

Chapter 1. Introduction

1.1 Problem Statement

The United States is undergoing massive transformations not only in terms of the transportation infrastructure but also the mobility offerings. One of the most prominent sets of disruptive technologies in the transportation market is the Transportation Network Companies (TNCs), otherwise known as ridesourcing companies. These companies operate by using mobile devices such as smartphones to directly link commuters actively seeking transportation and drivers who act as owners/operators. The ridesourcing service has been rapidly adopted and has impressively penetrated the market since first being introduced by Uber in 2009. According to the data collected from the National Household Travel Survey (NHTS), the for-hire vehicle market has doubled from 2009 to 2017 due to the rapid expansion of TNCs, and about 10% of all Americans used ridesourcing services in any given month in 2017 (Conway, et al., 2018). The growing demand of TNCs is substantial. According to a recently released report, the total number of passengers transported by TNCs increased 37 percent from 1.90 billion in 2016 to 2.61 billion in 2017 (Schaller, 2018). Most of the passengers are serviced by Uber and Lyft. According to market share data from October 2018, the industry leader, Uber, accounts for 69 percent of the ridesourcing service market. 29 percent of the market is taken by Lyft, the second largest TNC in the US. The remaining TNCs, such as Via, Juno, and Gett, account for 2 percent of the US ridesourcing service market (Gessner, 2019). Although the adoption of TNCs is continuing to increase, the impact of these services on transportation network and travel behavior is still ambiguous. On one hand, the convenience and efficiency of TNCs hold promise for reducing vehicle ownership and promoting transit usage in urban areas. On the other hand, there is some concern that TNCs may take passengers away from public transit, increase vehicle miles traveled (VMT), and attenuate congestions.

The primary objective of the first part of this study is to understand the relationship of demand for TNCs and other shared modal offerings, namely taxis, subways and bikeshares, using a *multivariate modeling framework* to incorporate temporal patterns and effects of other exogenous factors. We formulate a vector autoregressive model with exogenous predictors (VARX) to explore the "substitutional" and "complementary" effects between TNCs, Taxis, and Citi Bike in New York City over the time span of time between April 2015 and June 2017. The response vector for the VARX model consists of transformed compositional time series, which is a multivariate data structure that allows one to model the *daily* demand for each mode as a proportion of the total. In addition, we fit a univariate DLM to total daily counts. We use both models to calculate fitted/predicted daily counts for each mode.

Exploring the role of TNCs in the shared mobility landscape is a useful research project. We carry out this research by not only assessing the spatiotemporal patterns for TNCs but also exploring the interplay with the demand for other shared ride modes in a given region. The growth of TNC is irrefutable and the direct impact of TNCs on mode choice behaviors of consumers is very evident. The current body of research on TNC is growing, and the literature is beginning to shine light on how they impact the demand for existing shared modes (e.g. bikeshare, transit). However, a comprehensive investigation of all shared ride modes in a particular region, including the temporal and spatial patterns of overall shared demand and its relationship to other shared modes, is lacking.

1.2 Objectives

The objectives are as outlined below:

1) Develop and apply alternative multivariate modeling methodologies for analyzing the spatiotemporal dependence patterns both within different shared ride mode offerings but also across shared ride modes, at high resolution.
2) Demonstrate the methods to analyze the spatiotemporal patterns within and across shared ride modes in greater New York City metropolitan area. In particular, the proposed research will focus

on the synergy (or lack thereof) between TNCs and other shared mode offerings including subway, bikeshare, and taxis.
3) Disseminate the approaches that have been developed to benefit planning.

1.3 Expected Contributions

To accomplish these objectives, several tasks have been undertaken.

1) We have done an exploratory analysis of the large ridesourcing data in order to understand the patterns of behavior of the counts of each model over time, by aggregating the data over all the zones in NYC.
2) We have carried out a comprehensive dynamic compositional statistical data analysis of the extensive ridesourcing data in order to understand patterns in each mode over time, as well as to understand patterns in the complementary and substitutability behaviors between the models.
3) We have done extensive exploratory analysis of the large ridesourcing data by taxi zone in NYC in order to understand the patterns of behavior of the counts of each mode in each taxi zone over time, and to understand any spatial patterns between the zones.

1.4 Report Overview

The remainder of this report is organized as follows: Chapter 2 presents a comprehensive review of the state-of-the-art and state-of-the-practice literature on the ridesourcing data and analysis. Chapter 3 provides a detailed model formulation for doing a dynamic compositional analysis of the multivariate time series data aggregated over the zones in NYC using R packages to analyze state space models. The results show interesting aspects relating the different ridesourcing modes. Chapter 4 describes the analysis of the spatio-temporal patterns in the ridesourcing data across different taxi zones in NYC. We analyzed both daily and weekly data using R packages. We first fit adequate time series (ARIMA/GARCH) models to the data in each zone and then studied the spatial correlation between the residuals after removing the temporal effects. We also carried out a spatial association analysis at taxi zone level by regressing the ARIMA/GARCH residuals on land use and demographic variables and computing Moran’s I statistics on the residuals from the resulting ARIMA/GARCH/MLR model. Chapter 5 summarizes the R code that was used to carry out the analysis. Finally, Chapter 6 concludes this report with a summary and a discussion of our project.

Chapter 2. Literature Review

2.1 Introduction

This chapter provides a review and synthesis of the state-of-the-art and state-of-the-practice literature on the ridesourcing problem. The growth of TNC is irrefutable and the direct impact of TNCs on mode choice behaviors of consumers is very evident. The current body of research on TNC is growing, and the literature is beginning to shine light on how they impact the demand for existing shared modes (e.g. bikeshare, transit).

2.2 Existing Research on Ridesourcing

The literature on ridesourcing has been growing rapidly in recent years given that more and more TNCs have made their data publicly available. Ridesourcing has been widely compared with traditional taxis since it exhibits similar characteristics and provides similar services. In fact, some studies point out that the TNCs provide better services compared to traditional taxis with respect to shorter waiting and travel times, and lower costs (Rayle, et al., 2016). Given the appealing advantages of ridesourcing services, it is not surprising that a large portion of the market share of taxis has been taken by the TNCs in many metropolitan cities. DeMay (2018) found that Seattleites used ridesourcing services 3.5 times more often than taxis. In New York City, Brodeur and Nield (2018) found that the number of taxi rides, number of passengers and fare income all decreased after Uber entered the market in May 2011, while Warerkar (2017) found that Uber had already overtaken the ridership of yellow cabs in 2017. In Washington D.C., the ridesourcing market has exploded since late 2015, which has coincided with a 31% drop in taxi ridership (Siddiqui, 2018).

Many studies have found that TNCs have significant impacts on public transit ridership as well. However, unlike the evident competitive relationship between TNCs and taxis, the impacts of TNCs on public transit remains unclear. On the one hand, TNCs can service as an alternative mode that lures passengers away from public transit. Alternatively, Uber can be the solution to the first-last mile problem to help connect riders to public transportation options, which in turn could help increase transit ridership. Current studies have found both substitution and complementary impacts of TNCs on public transit demand. Contreras and Paz (2018) applied a multinomial linear regression to ridership data from Las Vegas after controlling for exogenous variables, TNCs showed a significant negative impact on taxicab ridership but a complementary impact on public transit ridership. Similarly, Hall et al. (2018) analyzed variations of Uber's impacts on public transit across US metropolitan areas and indicated that Uber complements the public transit and leads to an average of 5 percent increase in transit ridership after two years. On the contrary, Erhardt et al. (2019) conducted a longitudinal analysis to explore the reasons for the decline in public transit in major US cities from 2002 to 2018 which suggests that TNCs may be the main reason for the decrease of transit ridership. The ridership of heavy rail and buses is expected to decrease by 1.3 percent and 1.8 percent respectively for each year since TNCs entry into the market. Jin et al. (2019) analyzed Uber pickup data in New York City in 2014 and found that Uber both competes and complements with public transit; Uber competes with the ridership of public transit during most hours of the day in areas with good public transit coverage, while it complements the public transit services in areas with insufficient public transit service during the midnight hours.

The relationship of TNCs and bikeshare has not been extensively analyzed. Both Hoffman (2016) and Erhardt et al. (2019) noted the impacts of bikeshare on public transit demand, but did not mention the relationship between bikeshare and TNCs. To the best of our knowledge, the study by Gerte et al. (2019) is the only one exploring the influence of bikeshare on the demand of TNCs using data from New York City from 2015 to 2017. They found that bikeshare negatively influences the demand of TNCs, which could be because the both modes share the same user population. Depending on circumstances, such as weather, TNCs/bikeshare availability, and cost, the users may be switching back and forth. Gerte et al. (2019) explore the relationship between TNCs and other shared modes (subway, taxi, bikeshare), their univariate analysis for each modal demand fails to incorporate the correlations between modes. In this article, we build a multivariate VARX model to time series of transformed compositions of daily modal demand in order to incorporate the relationship between demand patterns of TNCs and other shared modes (subway, taxi, and bikeshare) in New York City.

2.3 Summary

A comprehensive review and synthesis of the current and existing research and development of the ridesourcing problem has been discussed in the preceding section.

Chapter 3. Dynamic Predictive Models for Ridesourcing Services Using Daily Compositional Data

3.1 Introduction

The primary objective of this study was to understand the relationship of demand for TNCs and other shared modal offerings, namely taxis, subways and bikeshares, using a *multivariate modeling framework* to incorporate temporal patterns and effects of other exogenous factors. We formulated a vector autoregressive model with exogenous predictors (VARX) to explore the "substitutional" and "complementary" effects between TNCs, Taxis, and Citi Bike in New York City over the time span of time between April 2015 and June 2017. The response vector for the VARX model consisted of transformed compositional time series, which is a multivariate data structure that allows one to model the *daily* demand for each mode as a proportion of the total. In addition, we fit a univariate DLM to total daily counts. We used both models to calculate fitted/predicted daily counts for each mode.

The format of this section is as follows. Section 3.2 provides a description of the data as well as an exploratory analysis. Sections 3.3 presents the VARX modeling for the transformed compositions of daily modal demand. Section 3.4 shows the univariate DLM model for total daily counts. Section 3.5 describes forecasting of daily modal counts based on both models. Finally, Section 3.6 gives an overall summary and discussion of the results and conclusions that can be drawn from this analysis. Details have been presented in the manuscript Toman et al. (2019), which is under review in a peer reviewed journal.

3.2 Data Description and Exploratory Analysis

Our data follows the format of Gerte et al. (2019). Missing data and data quality issues related to Green Cabs, subway, or Citi Bike were rectified via imputation using local averaging or deleting observations. Verification of this data was done by cross referencing yearly totals used in this paper with MTA's reported subway ridership and other published records. Our analysis is restricted to the time period spanning from 04/01/2015-06/30/2017 because Via and Lyft had not yet started their services prior to April 1st, 2015. The seven transportation providers have been aggregated into four mode categories:

1. TNC (consolidating Uber, Lyft, and Via) (TLC, 2017)
2. Taxi (consolidating Yellow cabs and Green cabs) (MTA, 2017)
3. Citi Bike (Citi Bike, 2017)
4. Subway (MTA, 2017)

Table 3.1 provides a numerical description of daily observed counts for each mode as well as the total. Figures 3.1 and 3.2 show time series plots for all four modes and all modes except subway, respectively.

Table 3.1 Five Number Summary of Observed Modal Demand Counts

	TNC	Taxi	Citi Bike	Subway	Total
Min.	59075.00	106071.00	1997.00	1334767.00	156701.00
Q1	147538.00	363319.00	25920.00	3503973.00	4269713.00
Median	219908.00	401732.00	36224.00	5525581.00	6159584.00
Q3	295363.00	439449.00	47023.00	5870137.00	6566306.00
Max	539267.00	580465.00	69772.00	6233796.00	6975008.00

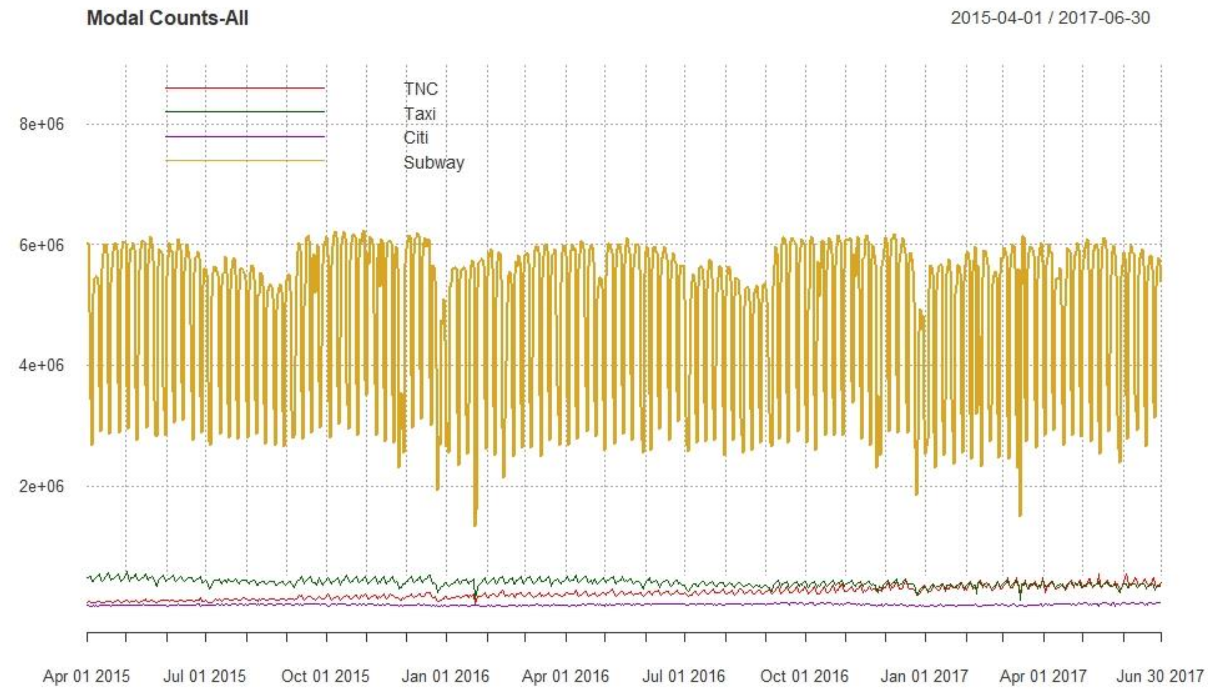


Figure 3.1. Demand of Counts for Four Modes

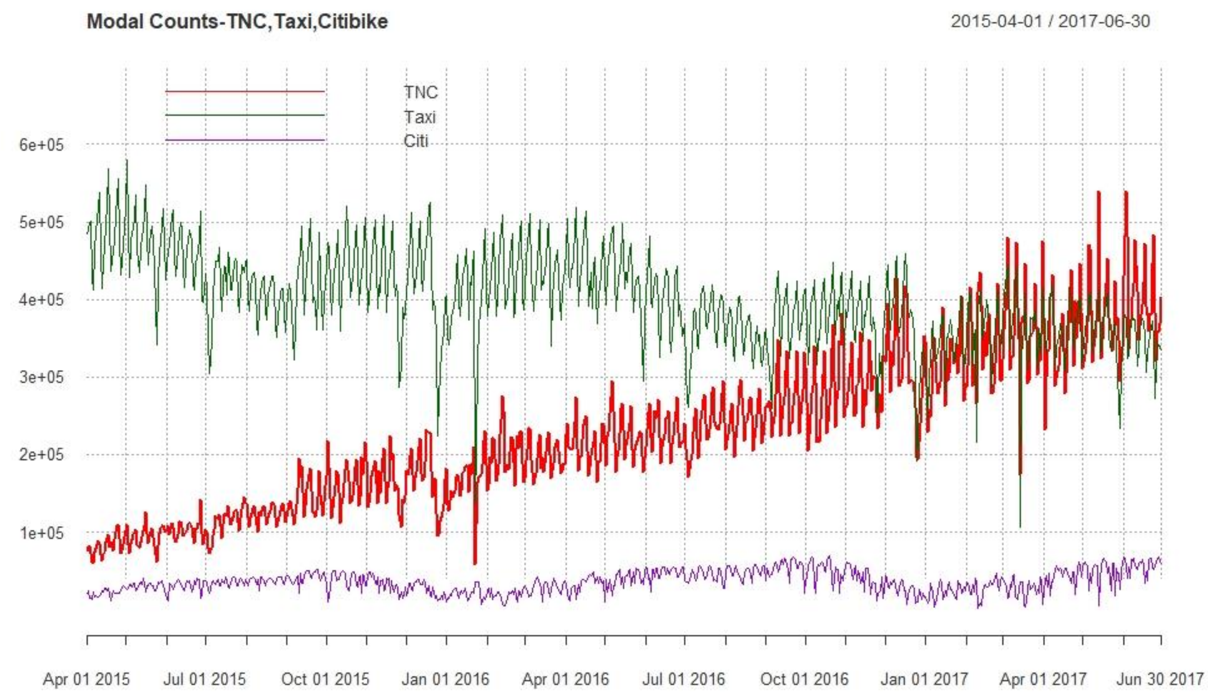


Figure 3.2. Modal Demand Counts-TNC, Taxi, and Citi Bike

3.3 VARX modeling for the transformed compositions

- We first defined proportions of daily demand for each transportation mode (Citi Bike, subway, taxi, and TNC).
- We used the Additive Log Ratio (ALR) transformation (Aitchison, 1986) to convert the G=4 proportions into compositions $\boldsymbol{Y_t} = (\boldsymbol{Y_{t,1}}, \boldsymbol{Y_{t,2}}, \boldsymbol{Y_{t,3}})^{\boldsymbol{T}}$, defined on a 3-dimensionalsimplex. We use subway as the baseline component, $\boldsymbol{X_{t,G}}$.
- We investigated cross-correlations between thee components of the ALR transformed counts.
- We fit a vector autoregressive model with exogenous predictors to the vector of ALR components. The exogenous predictors included a set of indicators functions corresponding to the day of the week, federal holidays, peak NYC travel season (Sep-Dec), and Peak Citi Bike Usage (May-Oct). In addition, covariates for the number of city-issued event permits on a particular day and daily precipitation measured in inches were included in the model. This model was fit using conditional least squares
 implemented via the *vars* (Pfaff, 2008) package in R. Note that the model includes coefficients for all seven days and no intercept.
- We fit VARX(p) models for $\boldsymbol{1 \leq p \leq 10}$ using the same set of exogenous predictors for every p. To select the best model, i.e., the best value of p, we used the Bayesian Information Criterion (BIC) and Mean Absolute Prediction Error (MAPE).

We show results from a VARX(1) model fit in Table 3.2.

Table 3.2 Estimated Coefficients from the VARX(1) Model Fit

	TNC(*t*)	Taxi(*t*)	Citi(*t*)
TNC (*t*-1)	0.699***	-0.043*	-0.020
Taxi (*t*-1)	-0.544***	0.299***	-0.189
Citi Bike (*t*-1)	0.019*	-0.017**	0.432***
Trend	0.0004***	-0.0002***	0.0003
Wednesday	-2.509***	-2.029***	-3.675***
Thursday	-2.430***	-1.985***	-3.707***
Friday	-2.381***	-1.943***	-3.796***
Saturday	-1.875***	-1.442***	-3.526***
Sunday	-1.900***	-1.474***	-3.297***
Monday	-2.719***	-2.232***	-3.712***
Tuesday	-2.529***	-2.046***	-3.710***
Precipitation	0.115***	0.021**	-0.348***
Holiday	0.359***	0.342***	0.155***
Peak Travel	-0.014	-0.021***	0.099***
Events	-0.00004	-0.00004	0.001***
Citi Bike Peak	-0.024**	-0.010	0.208***
Observations	736	736	736
Adjusted R^2	0.999	0.999	0.997
Residual Std. Error	0.098	0.070	0.288
F-Stat (df=16;720)	50,027.28***	57,473.750 ***	13,835.260***

- We next fit a dynamic linear model (DLM) for the daily total count.
- We then predicted rider counts by modal type by combining the results from the compositional analysis with the model from total counts. The forecasts are shown in Figure 3.3.

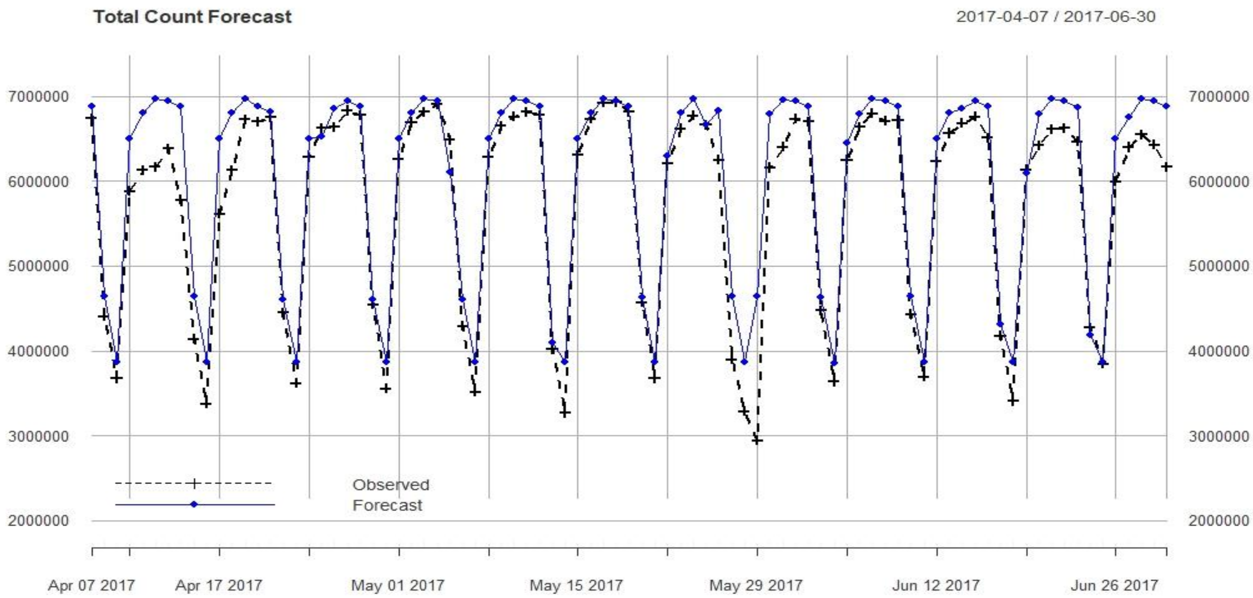


Figure 3.3. 85 Day Ahead Forecast for Total Daily Demand Counts

3.4 Summary

The primary goal of this analysis was to explore the dynamic relationships in demand patterns between four modal offerings, TNC, Taxi, Citi Bike, and Subway and discuss how the usage of different modes may be changing over time in New York City. A compositional time series approach was used to study the dynamic relationships between the ALR transformed compositions of TNC, Taxi, and Citi Bike, relative to the baseline mode, Subway. This model helps us to estimate temporal patterns present in the daily demand for TNC, Taxi, and Citi Bike and how they have changed over the time period from 04/01/2015 up until 04/06/2017.

The results show that over this time period, TNC's proportion relative to the subway system steadily increased, while taxi services saw a steady decrease. These results suggest that there is a substitutionary relationship between the two modes as TNC poaches many of the same customers who would utilize taxi services. The compositional analysis also reveals that all three modes have a strong weekly seasonality, with TNC and taxi services seeing a large increase in usage between Thursday and Sunday. Furthermore, exogenous predictors such as major holidays and peak travel season were all found to be statistically significant in predicting demand proportions. This model enables us to forecast several steps ahead. By contrast, the univariate DLM showed that total daily ridership counts remain constant over the same time period and that the exogenous predictors (peak travel season, precipitation, and major holidays) are statistically significant predictors. This supports the conjecture of a strong substitutionary relationship between TNC and taxis. More explicitly, since the total count of ridership does not appear to be increasing, it seems plausible that the gains of TNC have come mostly at the expense of taxi services. The final step of the analysis involves combining the predicted modal proportions from Section 4 and the predicted total counts from Section 5 to obtain forecasts for the modal counts over the holdout period. The count forecasts in the holdout period for subway and taxi are more accurate than forecasts for counts of TNC and Citi Bike.

While research is ongoing regarding the effects of TNCs in the ridesourcing marketplace, many questions about the substitutionary and complementary dynamics between TNCs and other shared ride modes remain unresolved. Recent research on this topic by Erhardt et al., 2019 used a longitudinal random-effects model to study the effects of TNCs and bikesharing on public transport. They examined monthly aggregated rideshares in 22 metropolitan areas in the US from January 2002 to April 2018 and concluded that the introduction of TNCs had a negative association with motor bus and heavy rail ridership. Their findings also indicated that bikeshare had a substitutionary impact on motor buses and a complementary impact on heavy and light rail ridership. Ostensibly, their research goals are like ours in terms of assessing substitutionary and complementarity relationships between ridesourcing modes over time. However, there are some differences between the two analyses.

(i) First, our study focuses on the *dynamic* relationships between all forms of ridesourcing in a *joint framework* with an emphasis on modeling cross-correlations between the modes.

(ii) Secondly, our analysis is performed on daily data which allows us to gain insights into dynamic relationships at a finer temporal resolution. As a result, we can use the dynamic compositional analysis to not only draw inferences about relationships between modal offerings but also use this framework to generate useful short and medium term forecasts for use in public policy settings in contrast to the analysis in Schaller (2018) and Erhardt (2019).

In summary, the results of our compositional analysis indicate that the overall usage of shared ride modes does not show any appreciable increase over the study time period. However, at the modal level, there does seem to be a significant substitutionary dynamic between TNC and taxi as they vie for the same user base, and we can quantify this effect over time.

Chapter 4. Spatio-Temporal Analysis of Ridesourcing Data by Taxi-zones

4.1 Introduction

The primary purpose of this section is to discuss temporal modeling of ridesourcing data using suitable time series models and subsequent spatial analysis of resulting residuals from these models to understand associations between taxi zones in NYC. In this study, we fit time series models to each of the taxi and TNC demand data using autoregressive integrated moving average/Generalized autoregressive conditionally heteroscedastic (ARIMA/GARCH) models to trip counts of TNC and Taxi in New York City from January 1, 2015 to June 30, 2017. The analysis was conducted at a daily level and weekly level of aggregation. We then analyzed the residuals from these time series models to investigate spatial associations after accounting for land use and demographic information at the taxi zone level. Specifically, we carried out the following tasks:

- Data Processing and EDA.
- ARIMA/GARCH model fitting.
- Spatial Association Analysis on ARIMA/GARCH residuals at the taxi zone level.
- Multiple Linear Regression (MLR) analysis of time series residuals on demographic and land use variables.
- Spatial Association Analysis on ARIMA/GARCH/MLR residuals at the taxi zone level.

The weekly level analysis is shown in this report. Both weekly and daily levels of analysis will be included/summarized in our manuscript to be submitted for publication.

4.2 Data Processing and EDA

There are total of 263 taxi zones in New York City, however, some of the taxi zones have very low trip counts (<10) of either TNC or Taxi. For example, the zone of Governor's Island/Ellis Island/Liberty Island and the zone of Central Park. Similar as daily level data processing, we first removed the time period where the minimum daily trip counts of either mode are equal to zero. Then we aggregate the daily trip counts to weekly trip counts. After aggregation, there are total of 129 complete weeks in our dataset. We further removed the zones with mean weekly trip count less than 10. Our final dataset includes TNC trip counts of 229 zones and Taxi trip counts of 212 zones for 129 weeks starting from 2015-01-11 to 2017-06-25. Besides of modal data, the values of three exogenous variables are also aggregated to weekly level namely, weekly average precipitation in inch, weekly count of city permitted events, and a dummy variable to indicate if any holiday is included in a week. In most of the zones, TNC trip counts show a continuous increase over time, while the taxi trip counts exhibit a decrease trend. Figure 1 shows the weekly trip counts of TNC and Taxi at Allerton/Pelham Gardens, Bronx.

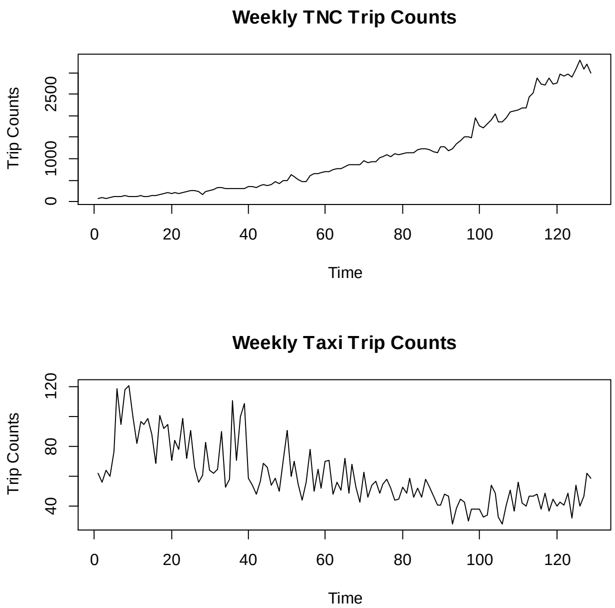


Figure 4.1. Weekly Trip Counts of TNC and Taxi at Allerton/Pelham Gardens, Bronx

4.3 ARIMA/GARCH Model Fitting Result

The purpose of fitting ARIMA and GARCH model is to remove the temporal pattern from trip counts of TNC and Taxi. Similar as daily-level data modeling, we fit ARIMA models on weekly TNC and Taxi trip counts for each taxi zone by applying auto.arima function in R. A log-transformation is applied for both TNC and Taxi trip counts in order to stabilize the variance. Ljung-Box and McLeod Li tests are performed to evaluate if the residuals are temporally clean. For the zones that are failing McLeod Li test, a GARCH (1,1) model is fitted on the residuals of ARIMA model to further remove the temporal pattern. Table 4.1 shows the final results of Ljung-Box and McLeod Li tests. There are a total of 213 taxi zones with temporally clean residuals of TNC data and 189 taxi zones with temporally clean residuals of Taxi data. Figure 4.2 shows the map of these taxi zones for both modes.

Table 4.1. Ljung-Box and McLeod Li Test Results

Mode	If include Exogenous Variables	Number of Taxi Zones Fail Ljung-Box Test ($p<0.05$ and Lag =12)	Number of Taxi Zones Fail McLeod Li Test ($p<0.05$ and Lag =12)	Total Number of Temporally Clean Zones	Total Number of Taxi Zones
TNC	No	16	2	213	229
Taxi	No	14	2	189	212

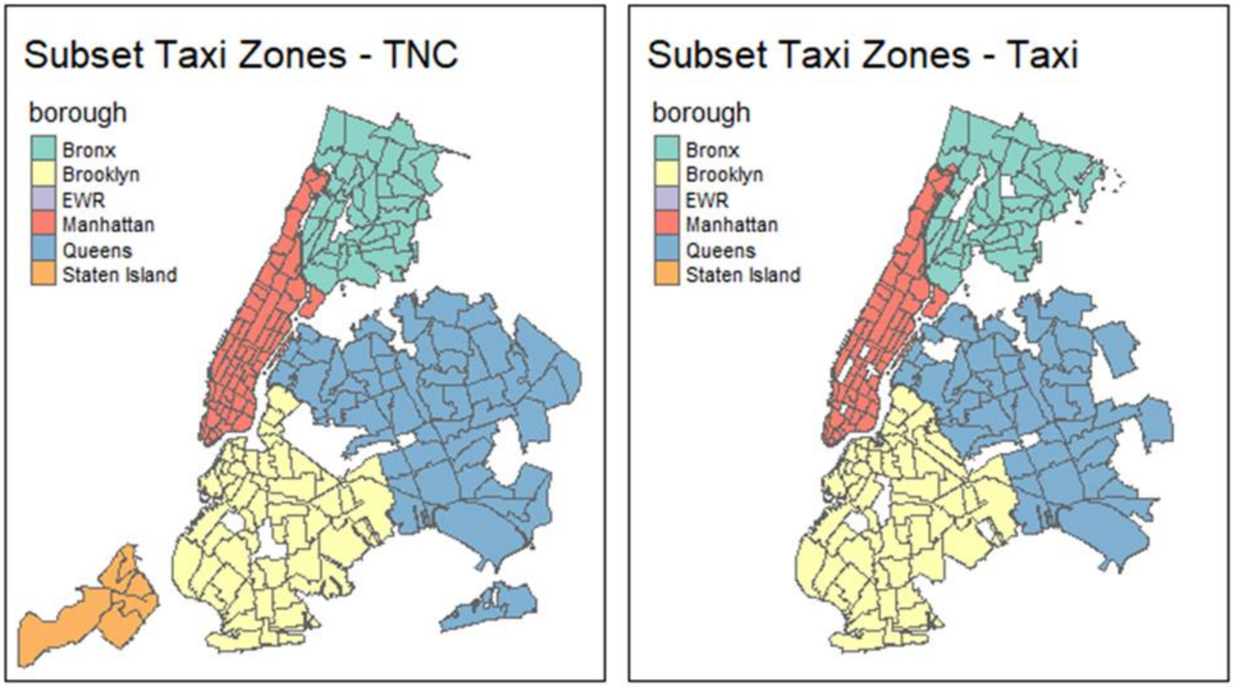


Figure 4.2. Taxi Zones with Temporally Clean Residuals for Taxi and TNC

4.4 Spatial Analysis of ARIMA/GARCH Residuals

The residual data of both TNC and taxi counts after fitting and removing temporal dependence are used for spatial analysis. Moran's I is calculated to quantity the correlation of a zone with its neighborhood. Both global and local Moran's I values are calculated to investigate the overall and local spatial correlation. R package 'spdep' is used for Moran's I calculation (Bivand, 2019).

4.4.1. Global Moran's I

Moran's I is a cross-product statistic between a variable and its spatial lag, with the variable expressed in deviations from its mean. Global Moran's I is calculated as the equation below (Gimond, 2019):

$$I = \frac{\sum_i \sum_j w_{ij} z_i * z_j / S_o}{\sum_i {z_i}^2 / n}$$

- w_{ij} : is the spatial weights of location i and j
- z_i: is the observation at location i
- $S_o = \sum_i \sum_j w_{ij}$ is the sum of all weights
- n: is the number of observations

Besides of Moran's I, a pseudo P value is also calculated to inform if the spatial correlation is statistically significant.

For our analysis, zones that share the same boundary are considered neighbors. All neighbors are weighted equally. Since Moran's I is often applied on cross-sectional data instead of time series, we aggregate the residuals at the following temporal resolutions (we do a similar aggregation for the as daily level data analysis as well):

- Full time series length
- 1/4 of the time series length
- Season – Spring, Summer, Fall, Winter
- Month

The residuals of both TNC and taxi are segmented into four equal sized sequential segments of residuals of 32 weeks (the first week was removed to get an integer number of weeks for each segment). Also, we aggregate all the series over the full length of time as a reference value. The results in Table 4.2 indicate that for all four segments of both modes, Moran's I rejects the null in favor of the alternative hypothesis, indicating a significant degree of spatial correlation across the zones. Furthermore, it is worth noting that when we aggregate across the full length of the series, Moran's I fails to reject the null hypothesis for Taxi; it rejects the null hypothesis for TNC.

Table 4.2: Global Moran's I - Segment Level

	TNC		Taxi	
	Moran I Statistic	p-value	Moran I Statistic	p-value
Segment 1	7.9680	<0.001	4.9851	<0.001
Segment 2	3.2541	<0.001	4.1059	<0.001
Segment 3	9.1424	<0.001	6.4023	<0.001
Segment 4	6.8790	<0.001	2.5635	0.0052
Full Series	3.7491	<0.001	1.5318	0.0628

Our next level of analysis is for the major seasons. Essentially, we take all the time points corresponding to spring, summer, fall, and winter, aggregate them for all the residual series and once again perform a global Moran's test. Indeed, all four seasons exhibit spatial clustering according to the results in Table 4.3 below.

Table 4.3: Global Moran's I -Season Level

	TNC		Taxi	
	Moran I Statistic	p-value	Moran I Statistic	p-value
Spring	4.7672	<0.001	6.7718	<0.001
Summer	7.1486	<0.001	3.7837	0.0001
Autumn	8.1210	<0.001	5.8486	<0.001
Winter	5.2768	<0.001	0.4374	0.3309

All 12 months exhibit spatial clustering according to the results in Table 4.4 below.

Table 4.4: Global Moran's I -Monthly Level

	TNC		Taxi	
	Moran I Statistic	p-value	Moran I Statistic	p-value
January	4.9744	<0.001	5.2250	<0.001
February	9.7074	<0.001	10.0815	<0.001
March	9.9091	<0.001	6.7233	<0.001
April	2.1195	0.017	5.8554	<0.001
May	7.3713	<0.001	4.6212	<0.001
June	7.5752	<0.001	3.5595	<0.001
July	5.7473	<0.001	2.6634	0.004
August	4.0230	<0.001	5.0354	<0.001
September	10.0105	<0.001	3.7446	<0.001
October	6.1406	<0.001	4.6914	<0.001
November	7.3623	<0.001	3.2865	<0.001
December	4.4517	<0.001	5.6918	<0.001

4.4.2 Local Moran's I

While the global Moran's I explains the overall spatial pattern of the 'clean' trip count of TNC and Taxi in NYC, the local Moran's I is a local indicator of spatial association of a zone with its neighborhoods. It is calculated by the following equations (Anselin, 1995):

$$I_i = \frac{z_i - \bar{Z}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j}(z_j - \bar{Z})$$

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n} (z_j - \bar{Z})^2}{n - 1}$$

where

- w_{ij} : is the spatial weights of location i and j
- z_i: is the observation at location i
- n: is the number of observations

The plots below are local Moran's I maps of taxi zones. The first plot indicates the values of local Moran's I for each taxi zone. The second plot only shows local Moran's I value of zones that are significant correlated with its neighborhood (p value < 0.05). The third plot shows p-value of local Moran's I. Similar as the global Moran's I, local Moran's I are also calculated at the same four temporal levels including: full length, four segments, four seasons and twelve months

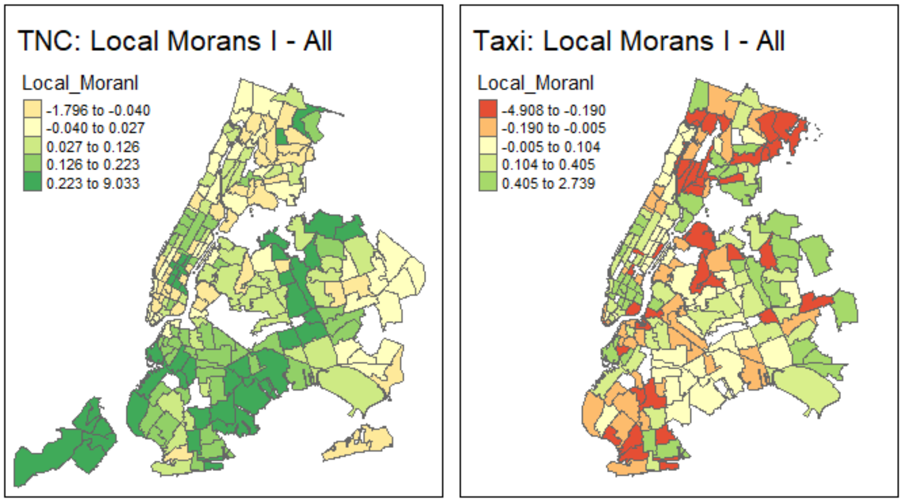
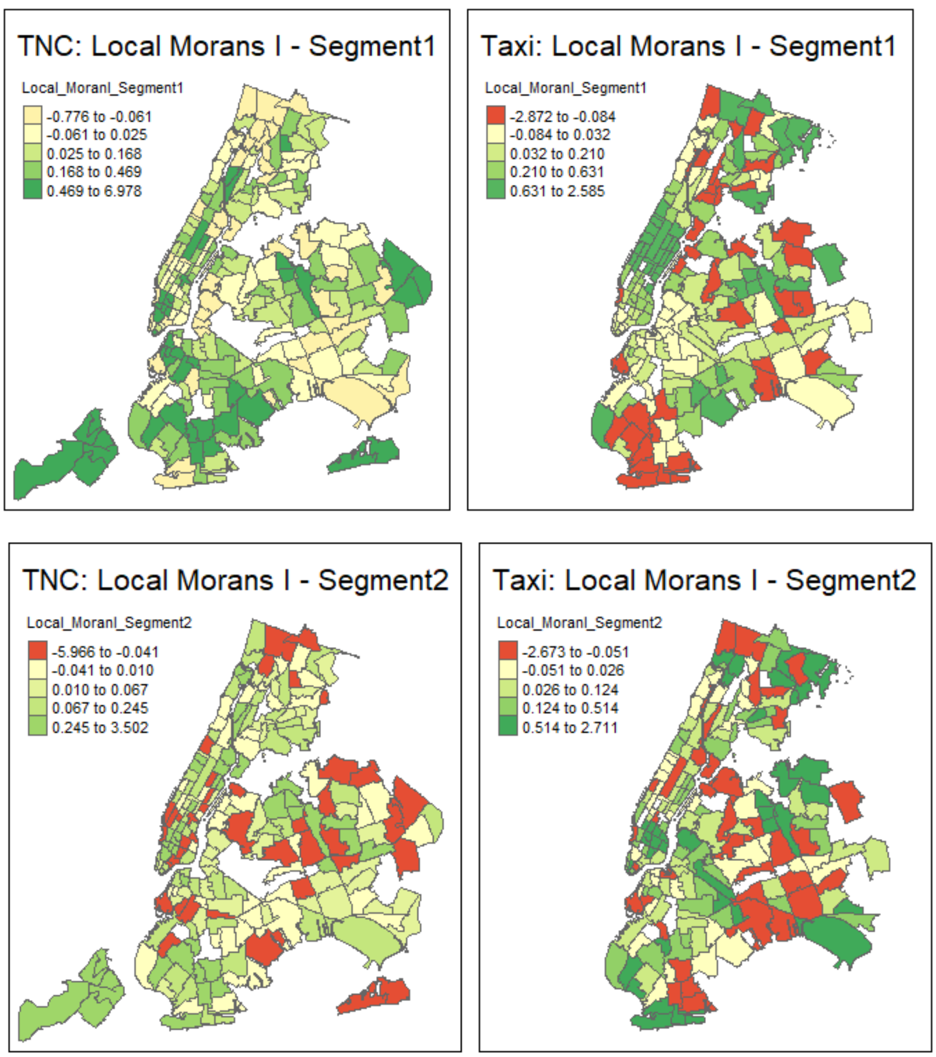


Figure 4.3 Local Moran's I of ARIMA/GARCH Residuals – Full Length



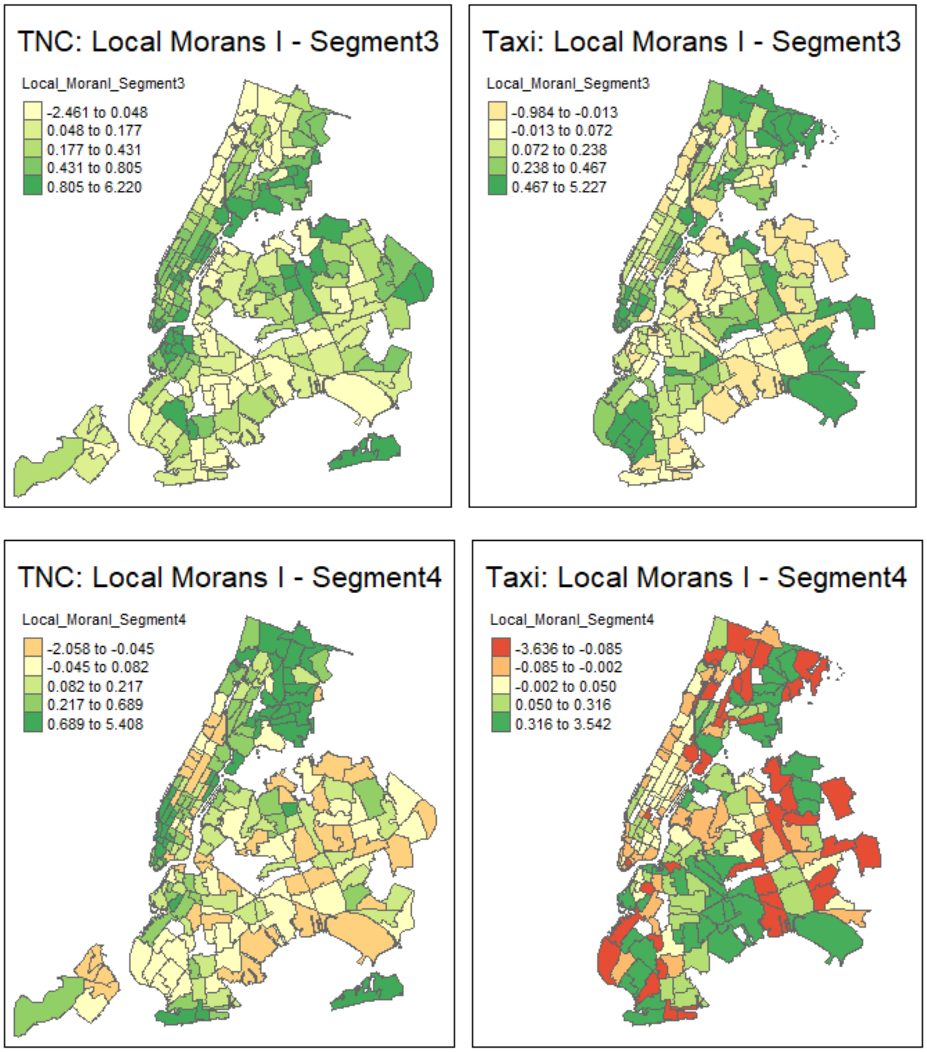
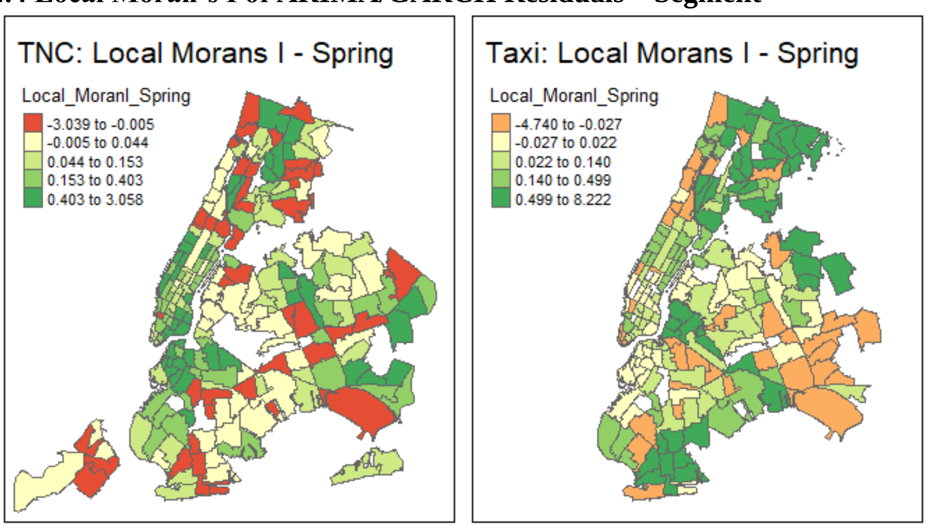


Figure 4.4 Local Moran's I of ARIMA/GARCH Residuals – Segment



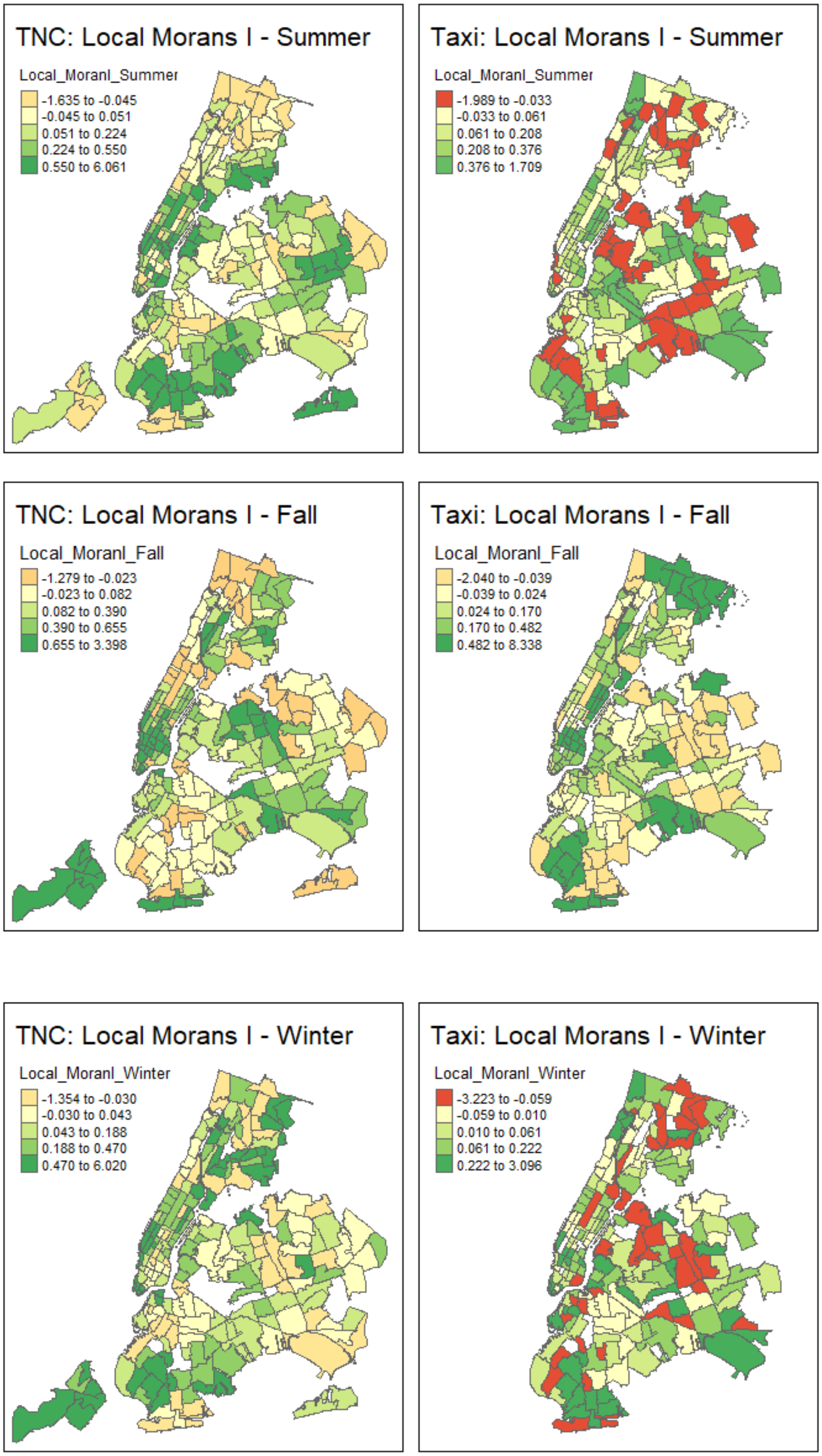
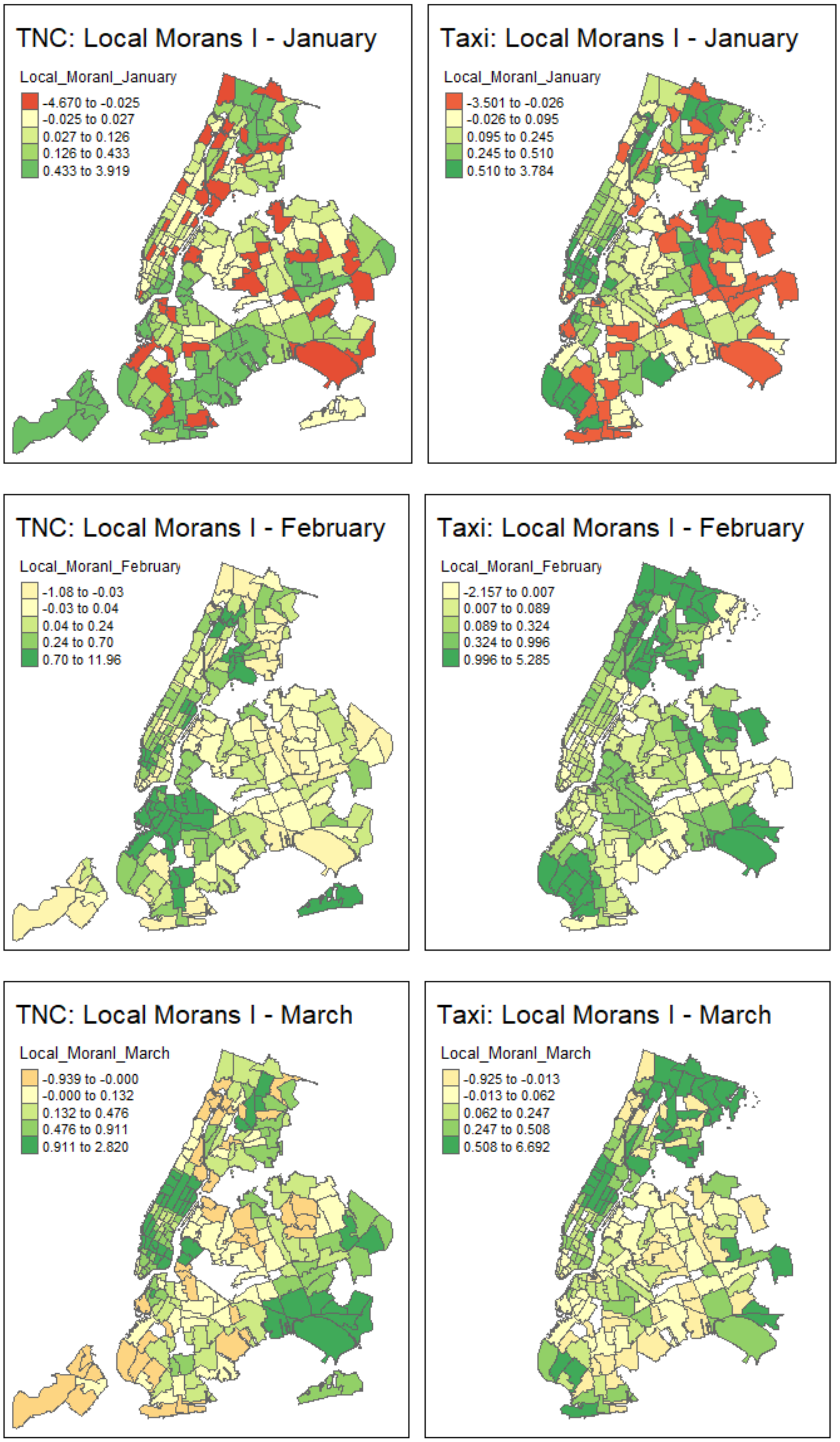
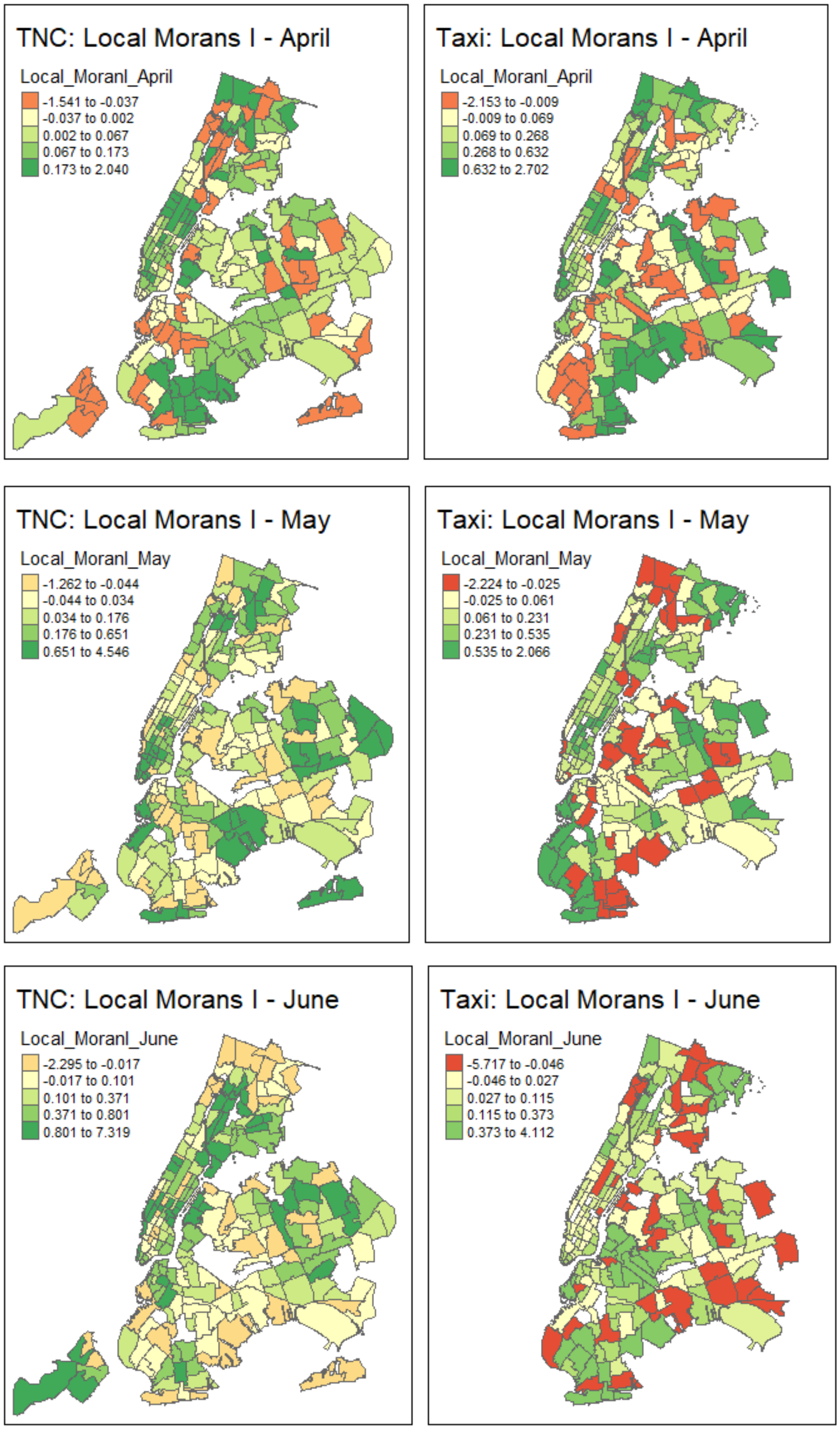
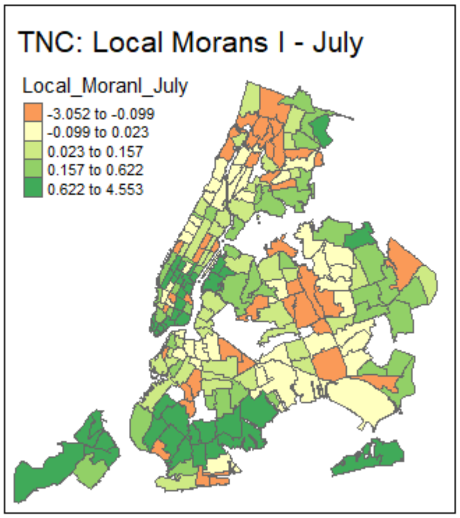
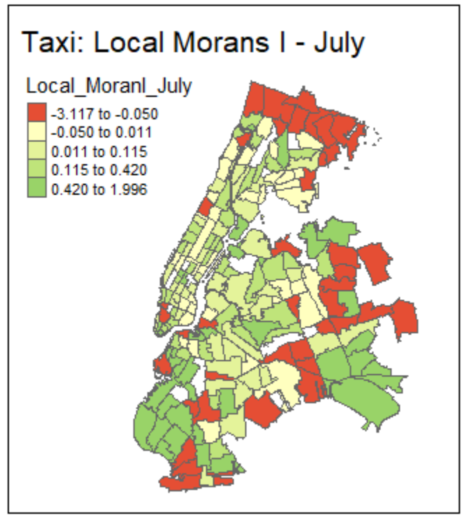


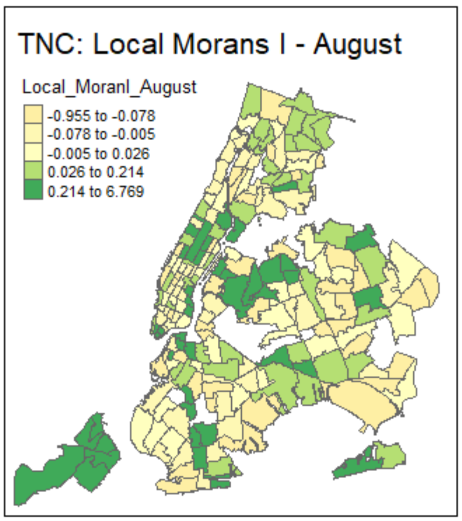
Figure 4.5 Local Moran's I of ARIMA/GARCH Residuals – Season

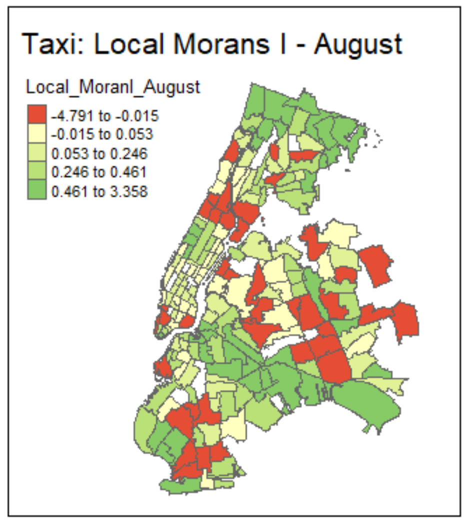

TNC: Local Morans I - January
Local_MoranI_January
-4.670 to -0.025
-0.025 to 0.027
0.027 to 0.126
0.126 to 0.433
0.433 to 3.919
Taxi: Local Morans I - January
Local_MoranI_January
-3.501 to -0.026
-0.026 to 0.095
0.095 to 0.245
0.245 to 0.510
0.510 to 3.784
TNC: Local Morans I - February
Local_MoranI_February
-1.08 to -0.03
-0.03 to 0.04
0.04 to 0.24
0.24 to 0.70
0.70 to 11.96
Taxi: Local Morans I - February
Local_MoranI_February
-2.157 to 0.007
0.007 to 0.089
0.089 to 0.324
0.324 to 0.996
0.996 to 5.285
TNC: Local Morans I - March
Local_MoranI_March
-0.939 to -0.000
-0.000 to 0.132
0.132 to 0.476
0.476 to 0.911
0.911 to 2.820
Taxi: Local Morans I - March
Local_MoranI_March
-0.925 to -0.013
-0.013 to 0.062
0.062 to 0.247
0.247 to 0.508
0.508 to 6.692

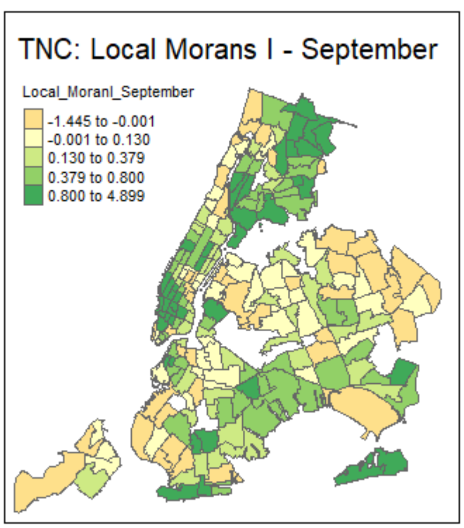

TNC: Local Morans I - April
Local_MoranI_April
-1.541 to -0.037
-0.037 to 0.002
0.002 to 0.067
0.067 to 0.173
0.173 to 2.040
Taxi: Local Morans I - April
Local_MoranI_April
-2.153 to -0.009
-0.009 to 0.069
0.069 to 0.268
0.268 to 0.632
0.632 to 2.702
TNC: Local Morans I - May
Local_MoranI_May
-1.262 to -0.044
-0.044 to 0.034
0.034 to 0.176
0.176 to 0.651
0.651 to 4.546
Taxi: Local Morans I - May
Local_MoranI_May
-2.224 to -0.025
-0.025 to 0.061
0.061 to 0.231
0.231 to 0.535
0.535 to 2.066
TNC: Local Morans I - June
Local_MoranI_June
-2.295 to -0.017
-0.017 to 0.101
0.101 to 0.371
0.371 to 0.801
0.801 to 7.319
Taxi: Local Morans I - June
Local_MoranI_June
-5.717 to -0.046
-0.046 to 0.027
0.027 to 0.115
0.115 to 0.373
0.373 to 4.112

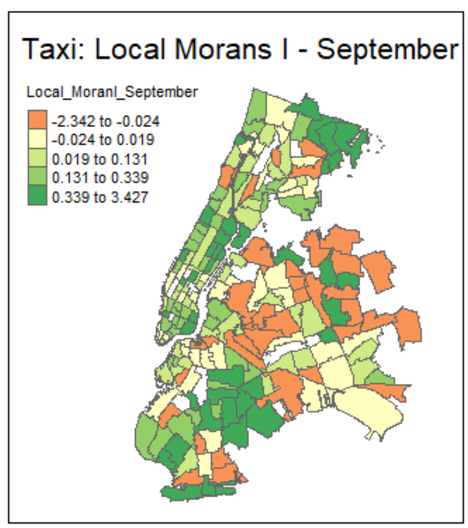

TNC: Local Morans I - July
Local_MoranI_July
-3.052 to -0.099
-0.099 to 0.023
0.023 to 0.157
0.157 to 0.622
0.622 to 4.553


Taxi: Local Morans I - July
Local_MoranI_July
-3.117 to -0.050
-0.050 to 0.011
0.011 to 0.115
0.115 to 0.420
0.420 to 1.996


TNC: Local Morans I - August
Local_MoranI_August
-0.955 to -0.078
-0.078 to -0.005
-0.005 to 0.026
0.026 to 0.214
0.214 to 6.769


Taxi: Local Morans I - August
Local_MoranI_August
-4.791 to -0.015
-0.015 to 0.053
0.053 to 0.246
0.246 to 0.461
0.461 to 3.358


TNC: Local Morans I - September
Local_MoranI_September
-1.445 to -0.001
-0.001 to 0.130
0.130 to 0.379
0.379 to 0.800
0.800 to 4.899


Taxi: Local Morans I - September
Local_MoranI_September
-2.342 to -0.024
-0.024 to 0.019
0.019 to 0.131
0.131 to 0.339
0.339 to 3.427

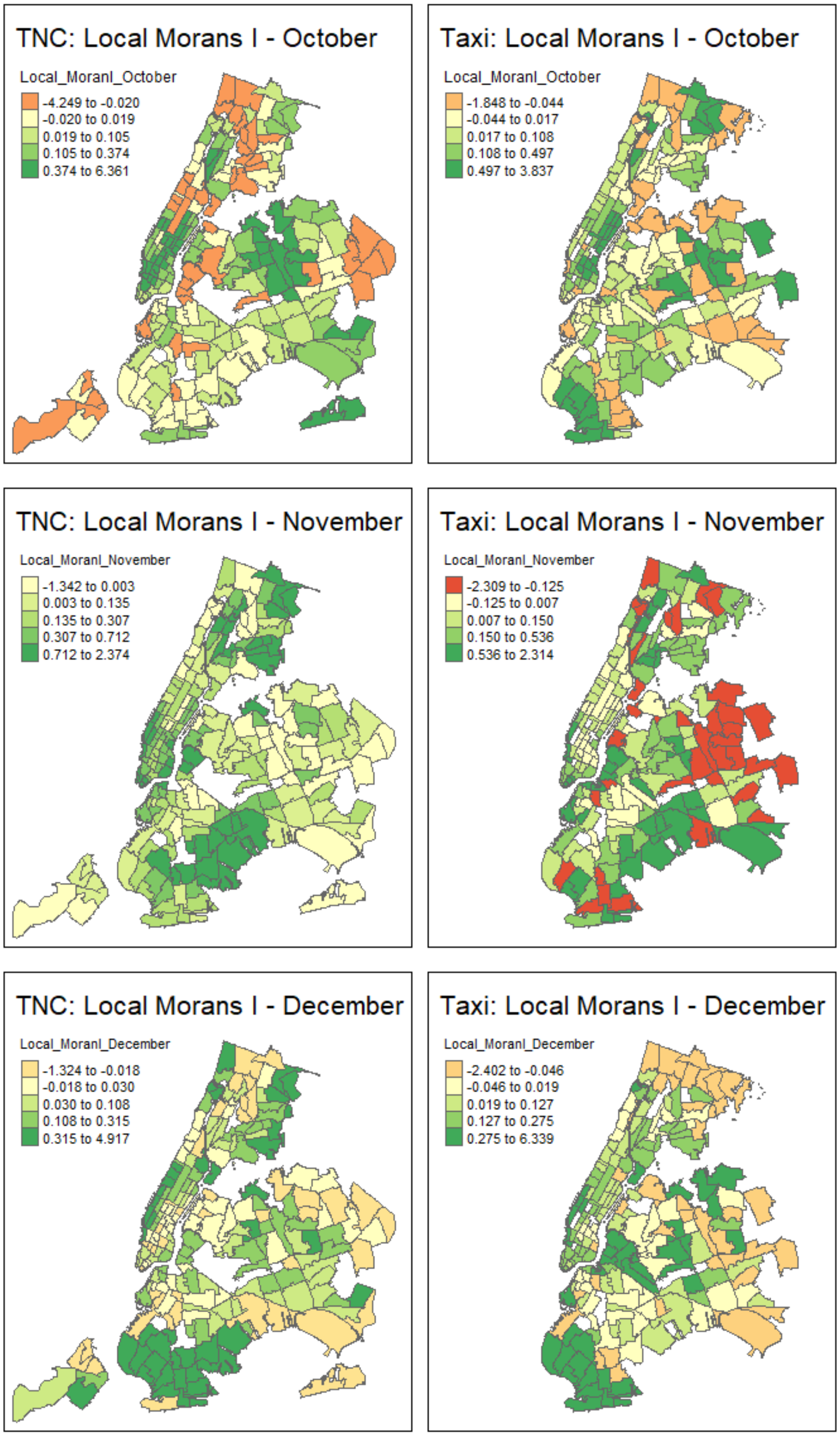


Figure 4.6 Local Moran's I of ARIMA/GARCH Residuals – Month

4.5 Multiple Linear Regression of ARIMA/GARCH Residuals on Land Use and Demographic Variables

After removing the temporal pattern from both TNC and taxi trip counts, the potential influence of other land use and demographic characteristics was removed by a multiple linear regression (MLR). This section presents a short description of census and land use data used for MLR and also the model estimation results.

4.5.1 Census and Land Use Data

The U.S. Census Bureau collects variety demographic information about American population and economy at different geographical resolutions annually. For our analysis, we focus on total number of population, number of full time employees, median age and median earning in each taxi zone in New York City from 2015 to 2017. The data are available to download at the website of U.S. Census Bureau. One must be note that the data is reported at census track level which is smaller than taxi zones. An aggregation is required to get the data at the taxi zone level. Land use data of New York City is collected by the Department of City Planning (DCP) which is a primary land use agency in New York City. DCP collects detailed land use and geographic data at the tax lot level annually including residential area, commercial area, retail area etc. The dataset also provides the census track id for each tax lot. Therefore, a similar aggregation has been performed to obtain the land use data in each taxi zone from 2015 to 2017.

4.5.2 Multiple Linear Regression on Time Series Residuals

Having established the presence of spatial auto-correlation between the aggregated residuals of the time series models, the residual series from each mode are pooled and then a multiple linear regression model is fit where we seek to account for variation in the residual series by land use and demographic covariates. We denote the model as such:

$$e = X\beta + \eta$$

where e is a vector of residuals that is created by pooling together the residual from each time series model, $\boldsymbol{X}$ is a design matrix consisting of all main effects for the demographic and land variables plus all second-order interactions between the demographic and land use variables which yields a total of covariates, β is the vector of coefficients for each covariate and $\eta \sim \mathrm{N}(0, \sigma^2 \mathrm{I_n})$ where I_n denotes the nxn identity matrix. Stepwise model selection is then performed on these models using the MASS library in R (Ripley et al., 2019). The selection criterion is AIC and both forward and backwards selection are utilized.

For the land use covariates, the sum of lot area and building area is used to create an exposure variable. That yields a proportion for eight types of land use in a respective taxi zone. More specifically, we had the following 8 land use covariates:

1. (Residential Area)/(Total Area)
2. (Commercial Area)/(Total Area)
3. (Retail Area)/(Total Area)
4. (Factory Area)/(Total Area)
5. (Storage Area)/(Total Area)
6. (Garage Area)/(Total Area)
7. (Office Area)/(Total Area)
8. (Other Area)/(Total Area)

Furthermore, we used the following 4 demographic variables:

1. (Total Population)/(# of Buildings)
2. (Fulltime Employed)/(# of Buildings)
3. Median Age
4. Median Earnings

Two different regression analysis are built on each mode. The first model incorporates only main effects while the second includes all main effects plus second order interactions. For both analyses, I use stepwise model selection based off the AIC to find a parsimonious subset of the predictors in both sets of models. Regression outputs for these models are given below-Note that stepwise selection gave the same model for taxi in both cases.

4.5.2.1. Taxi Results

Stepwise model selection reveals that only the number of fulltime employed individuals has a statistically significant correlation with the residuals of taxi services at the 5% level. Results are detailed in the table below.

Table 4.5 MLR Results of Taxi

Taxi-Best Model According to AIC				
	Term	Std. Error	T-Statistic	P-Value
Intercept	-0.003	0.001	-2.521	0.012
Fulltime Employment	-1.000e-08	4.336e-09	-2.307	0.022
Observations	25413			
Adjusted R^2	0.0002094			
Residual Std. Error	0.1491			
F-Stat (df=1;25411)	5.5322			

4.5.2.2 TNC Results

Model selection revealed a larger subset of covariates that were statistically significant for the 0.05 level in explaining variation in the TNC residuals. For the main effects, full time employment, median earnings, and median age had statistically significant relationships to the TNC residuals. In addition, interaction effects between the residential land use percentage and median age, office land use percentage and median age, and commercial percentage and fulltime employment were all found to have a statistically significant correlation with the TNC residuals.

Table 4.6 MLR Results of TNC

TNC-Best Model According to AIC				
	Term	Std. Error	T-Statistic	P-Value
Intercept	-0.005	0.004	-1.3405	0.180
ResidentialPct	0.024	0.013	1.778	0.075
FulltimeEmp	-2.311e-08	5.487e-09	-4.213	2.53e-05
Median Earnings	-2.151	8.391	-2.564	0.010
Median Age	0.001	0.0001	3.467	0.0005
ResidentialPct:Median Earnings	3.418e-07	1.900e-07	1.828	0.068
StoragePct:Median Earnings	1.022e-06	6.949e-07	1.472	0.141
ResidentialPct:Median Age	-0.0009	0.0004	-2.097	0.0360
OfficePct:MedianAge	0.0006	0.0002	2.950	0.0032
CommercialPct:FulltimeEmp	-9.396e-08	2.949e-08	-3.187	0.0014
Observations	27477			
Adjusted R^2	0.005			
Residual Std. Error	0.1234			
F-Stat (df=9;27467)	15.563			

4.6 Spatial Analysis on ARIMA/GARCH/MLR Residuals

After removing the impact of demographic and land use effects by MLR, similar spatial analysis as shown in section 4.4 is repeated on the residuals from MLR.

4.6.1 Global Moran's I on ARIMA/GARCH/MLR Residuals

The results in Table 4.7 indicate that for all 4 segments of TNC, Moran's I reject the null in favor of the alternative hypothesis which is that there is a significant degree of spatial correlation across the zones. For taxi, Moran's I tests of the first 3 segments of taxi data also reject the null hypothesis, but it fails to reject the null for segment 4. Furthermore, it is worth noting that when we aggregate across the full length of the series, Moran's I fails to reject the null hypothesis for Taxi; it rejects the null hypothesis for TNC.

Table 4.7 Global Moran's I on on ARIMA/GARCH/MLR Residuals-Segment Level

	TNC		Taxi	
	Moran I Statistic	P-value	Moran I Statistic	P-value
Segment 1	7.1904	<0.001	3.9660	<0.001
Segment 2	2.5985	0.0047	3.1412	<0.001
Segment 3	6.9465	<0.001	5.8728	<0.001
Segment 4	1.8741	0.0305	1.2383	0.1078
Full Series	1.7523	0.0399	1.1995	0.1152

All four seasons exhibit spatial clustering for TNC according to the results in Table 4.8 below. However, Winter season of taxi data does not exhibit spatial autocorrelation.

Table 4.8 Global Moran's I on ARIMA/GARCH/MLR Residuals -Season Level

	TNC		Taxi	
	Moran I Statistic	P-value	Moran I Statistic	P-value
Spring	4.3515	<0.001	6.5305	<0.001
Summer	7.0102	<0.001	3.8943	<0.001
Autumn	7.7366	<0.001	5.8456	<0.001
Winter	5.2898	<0.001	0.4810	0.3152

All 12 months exhibit spatial clustering according to the results in Table 4.9 below.

Table 4.9 Global Moran's I on ARIMA/GARCH/MLR Residuals -Monthly Level

	TNC		Taxi	
	Moran I Statistic	P-value	Moran I Statistic	P-value
January	4.9848	<0.001	5.1341	<0.001
February	9.2391	<0.001	9.9803	<0.001
March	9.8670	<0.001	6.6852	<0.001
April	2.2448	0.0124	5.8475	<0.001
May	6.9221	<0.001	4.7753	<0.001
June	8.1056	<0.001	3.6554	<0.001
July	5.4489	<0.001	2.6596	0.0039
August	4.0695	<0.001	5.0425	<0.001
September	9.8920	<0.001	3.7427	<0.001
October	5.8578	<0.001	4.6852	<0.001
November	7.5311	<0.001	3.2887	<0.001
December	4.5969	<0.001	5.6927	<0.001

4.6.2 Local Moran's I on ARIMA/GARCH/MLR Residuals

The plots below show the local Moran's I values calculated using the residuals from MLR

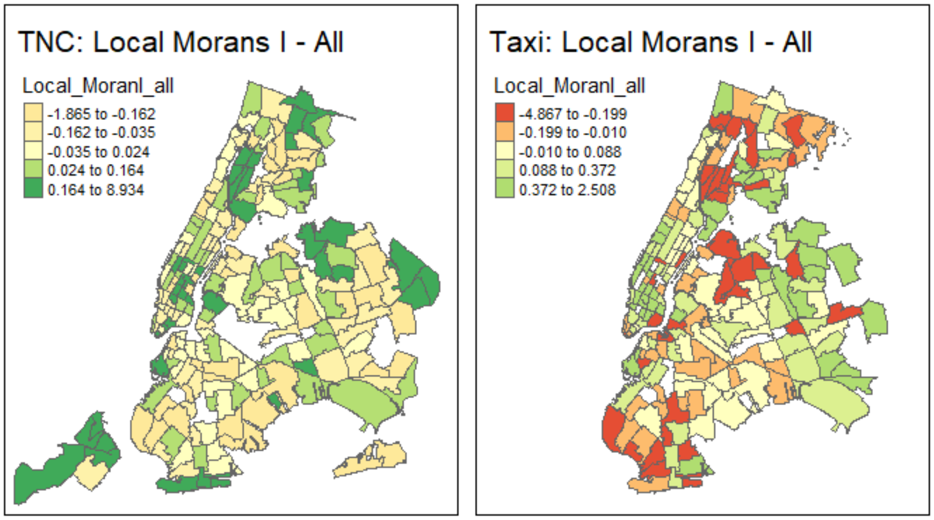
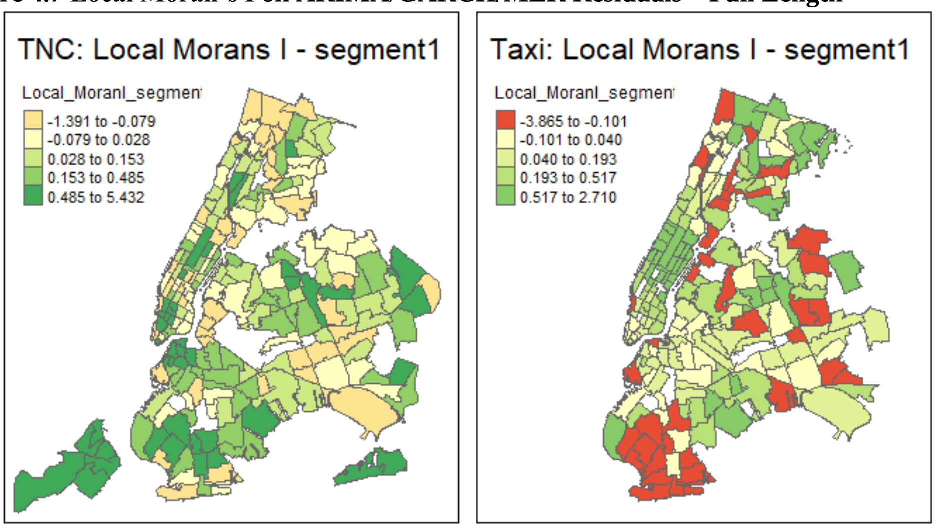


Figure 4.7 Local Moran's I on ARIMA/GARCH/MLR Residuals – Full Length



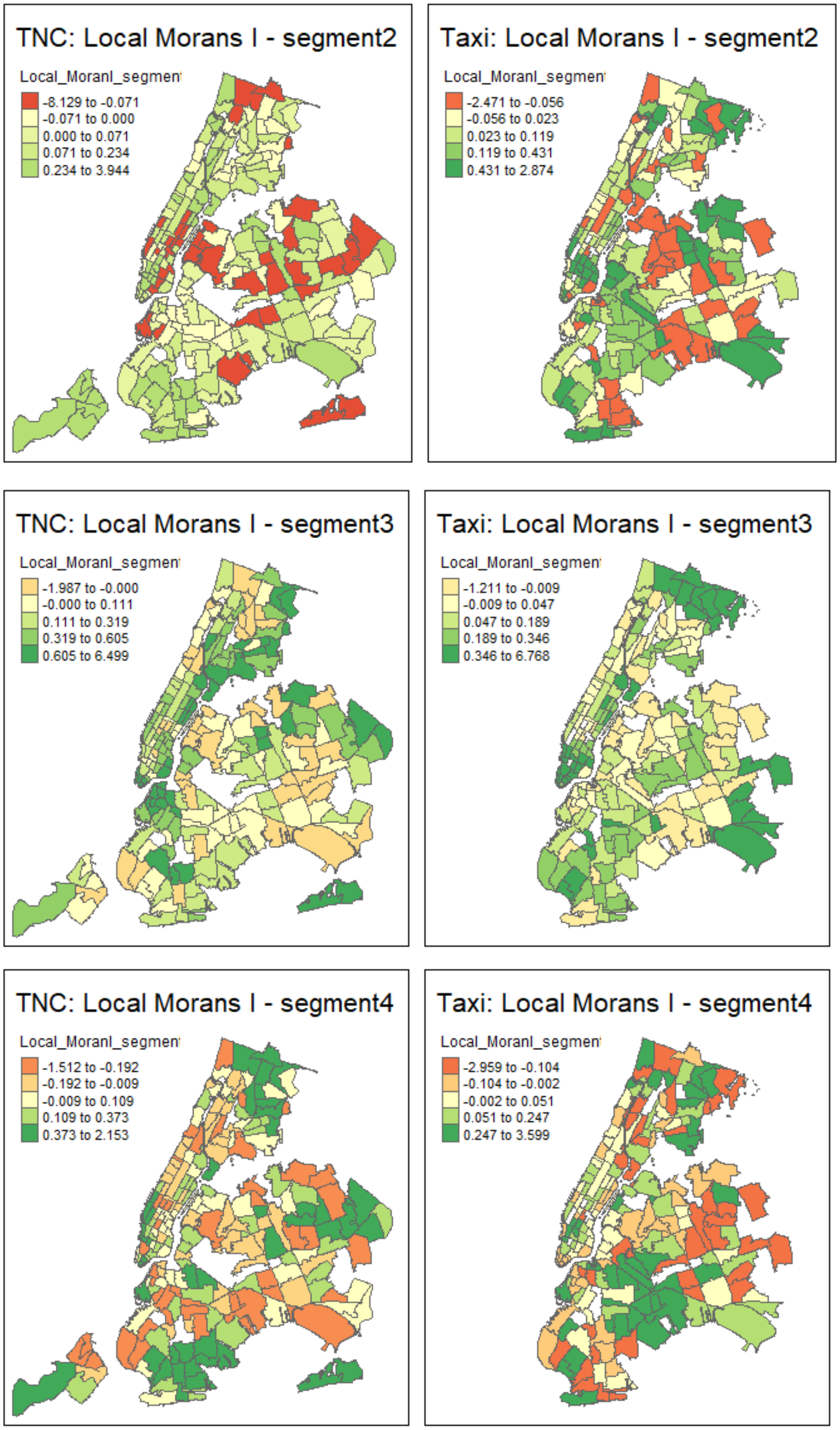
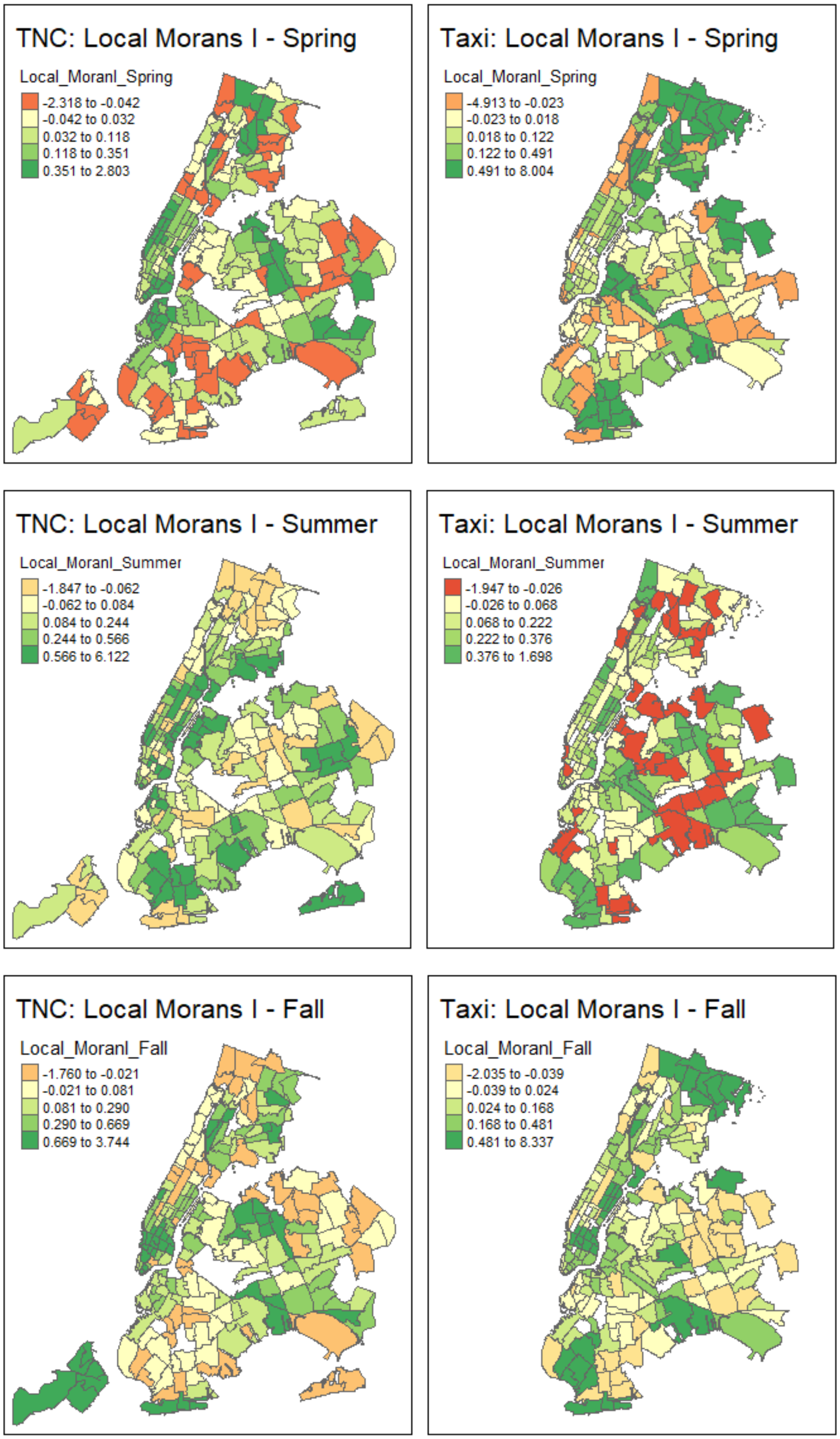


Figure 4.8 Local Moran's I on ARIMA/GARCH/MLR Residuals – Segment


TNC: Local Morans I - Spring
Local_MoranI_Spring
-2.318 to -0.042
-0.042 to 0.032
0.032 to 0.118
0.118 to 0.351
0.351 to 2.803
Taxi: Local Morans I - Spring
Local_MoranI_Spring
-4.913 to -0.023
-0.023 to 0.018
0.018 to 0.122
0.122 to 0.491
0.491 to 8.004
TNC: Local Morans I - Summer
Local_MoranI_Summer
-1.847 to -0.062
-0.062 to 0.084
0.084 to 0.244
0.244 to 0.566
0.566 to 6.122
Taxi: Local Morans I - Summer
Local_MoranI_Summer
-1.947 to -0.026
-0.026 to 0.068
0.068 to 0.222
0.222 to 0.376
0.376 to 1.698
TNC: Local Morans I - Fall
Local_MoranI_Fall
-1.760 to -0.021
-0.021 to 0.081
0.081 to 0.290
0.290 to 0.669
0.669 to 3.744
Taxi: Local Morans I - Fall
Local_MoranI_Fall
-2.035 to -0.039
-0.039 to 0.024
0.024 to 0.168
0.168 to 0.481
0.481 to 8.337

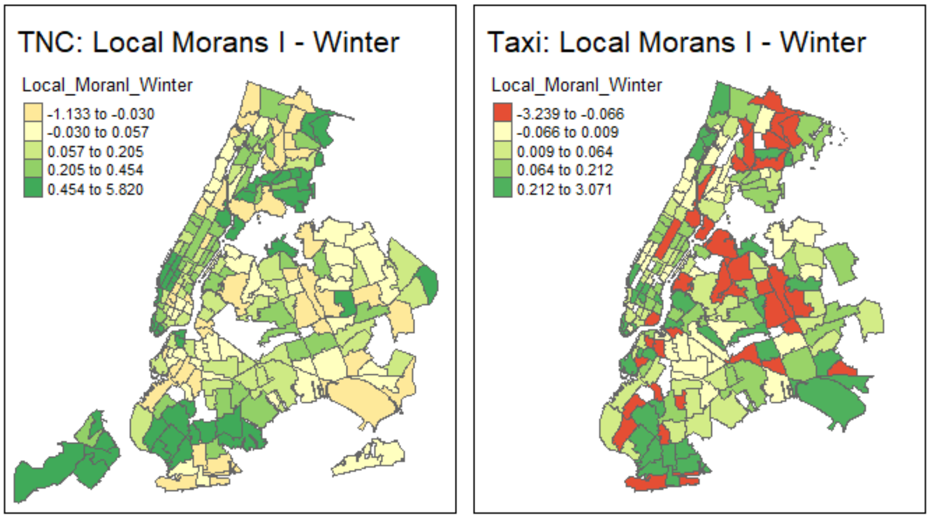
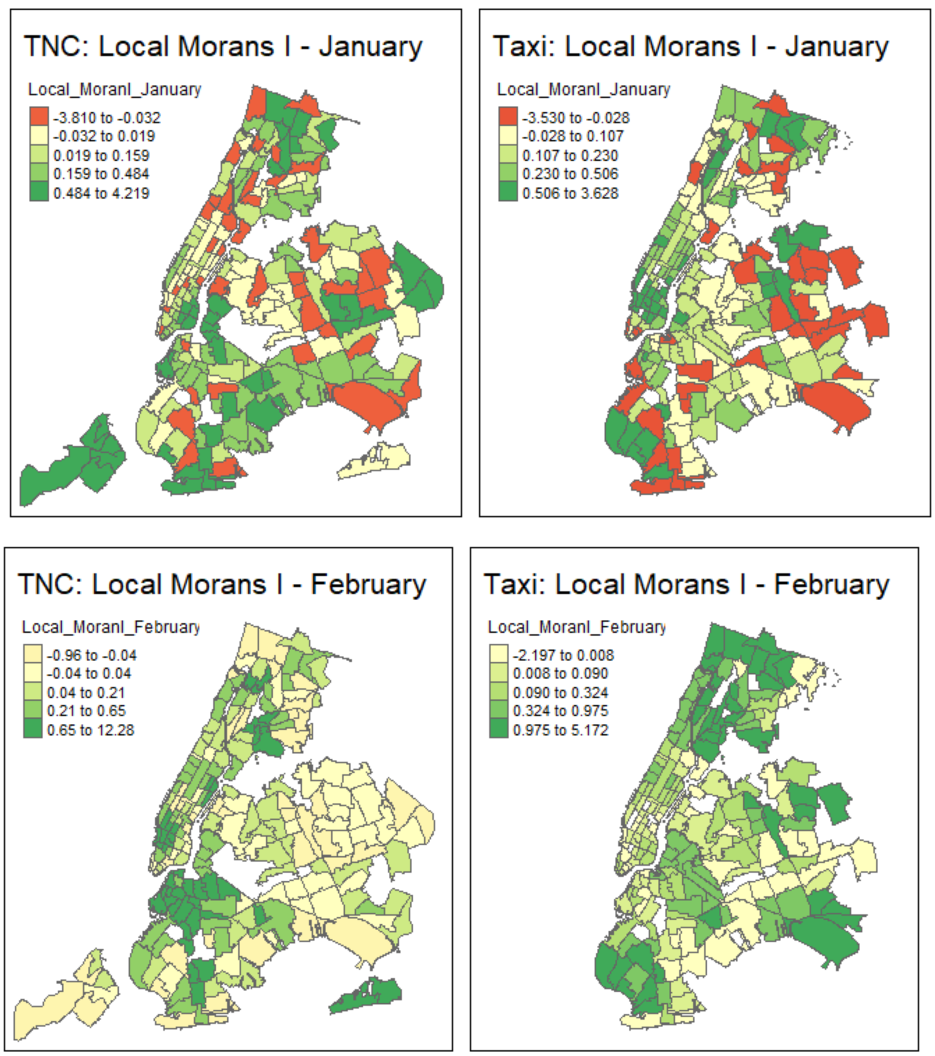
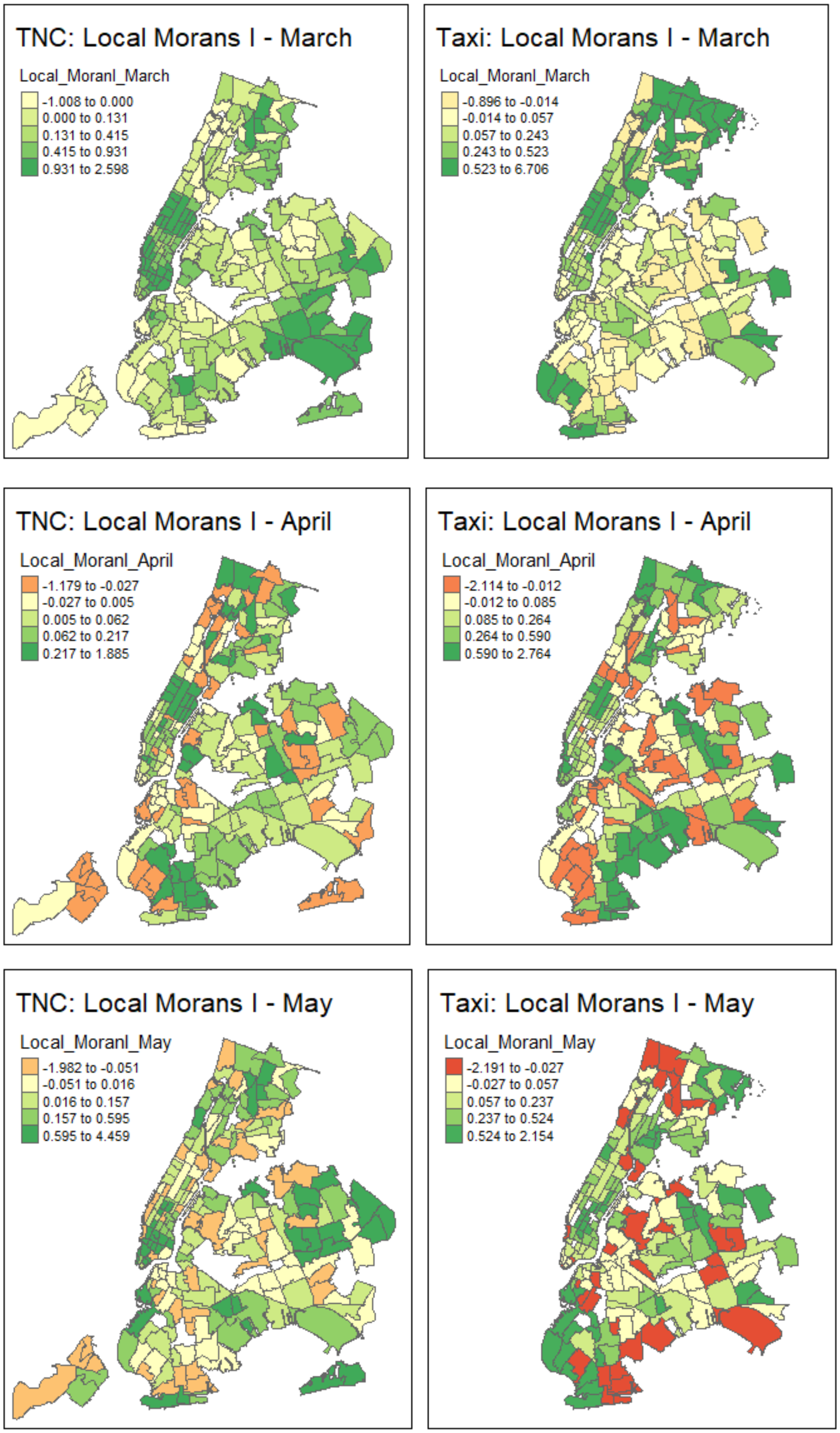
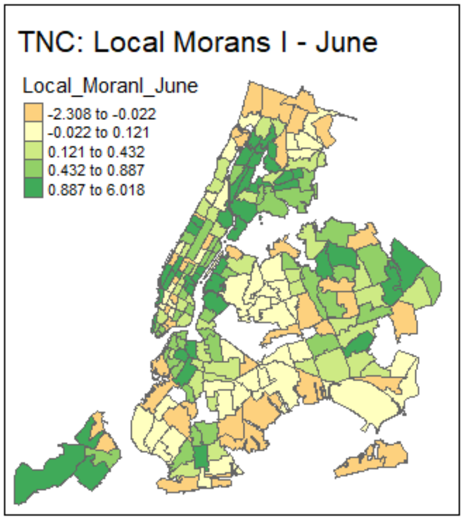
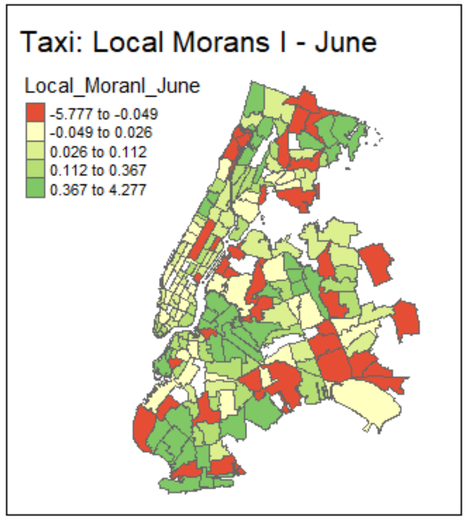


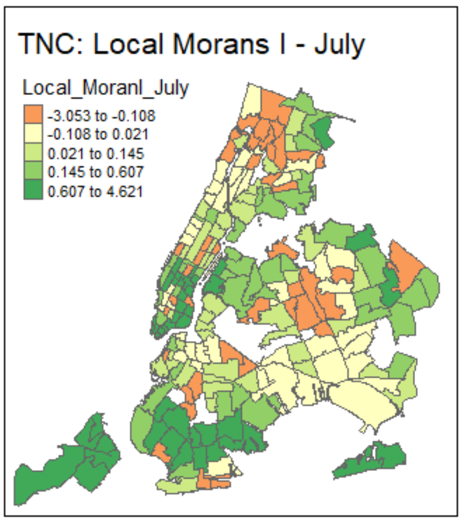
Figure 4.9 Local Moran's I on ARIMA/GARCH/MLR Residuals – Season

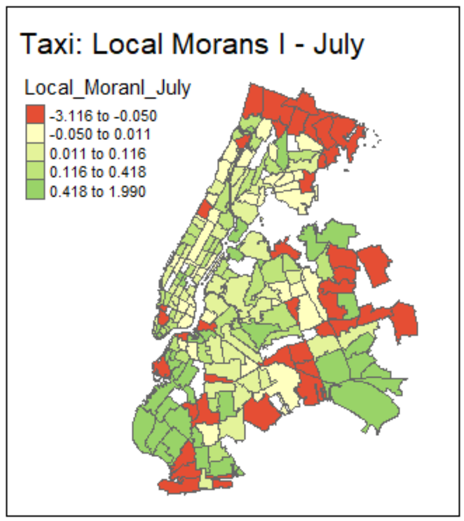


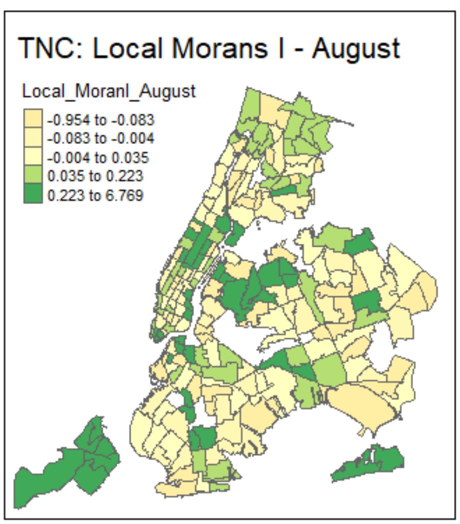

TNC: Local Morans I - March
Local_MoranI_March
-1.008 to 0.000
0.000 to 0.131
0.131 to 0.415
0.415 to 0.931
0.931 to 2.598
Taxi: Local Morans I - March
Local_MoranI_March
-0.896 to -0.014
-0.014 to 0.057
0.057 to 0.243
0.243 to 0.523
0.523 to 6.706
TNC: Local Morans I - April
Local_MoranI_April
-1.179 to -0.027
-0.027 to 0.005
0.005 to 0.062
0.062 to 0.217
0.217 to 1.885
Taxi: Local Morans I - April
Local_MoranI_April
-2.114 to -0.012
-0.012 to 0.085
0.085 to 0.264
0.264 to 0.590
0.590 to 2.764
TNC: Local Morans I - May
Local_MoranI_May
-1.982 to -0.051
-0.051 to 0.016
0.016 to 0.157
0.157 to 0.595
0.595 to 4.459
Taxi: Local Morans I - May
Local_MoranI_May
-2.191 to -0.027
-0.027 to 0.057
0.057 to 0.237
0.237 to 0.524
0.524 to 2.154

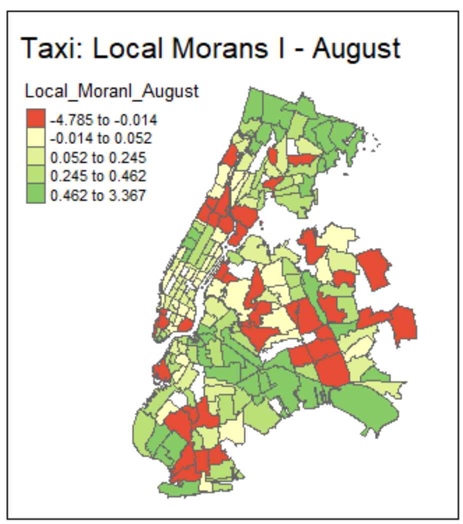

TNC: Local Morans I - June
Local_MoranI_June
-2.308 to -0.022
-0.022 to 0.121
0.121 to 0.432
0.432 to 0.887
0.887 to 6.018

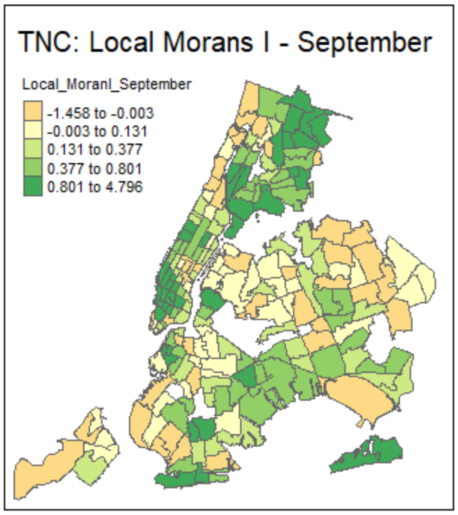

Taxi: Local Morans I - June
Local_MoranI_June
-5.777 to -0.049
-0.049 to 0.026
0.026 to 0.112
0.112 to 0.367
0.367 to 4.277

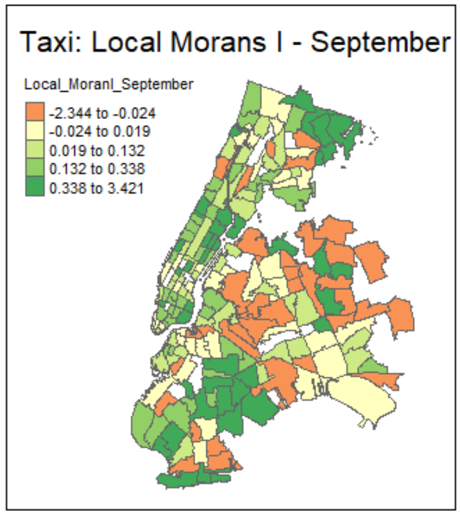

TNC: Local Morans I - July
Local_MoranI_July
-3.053 to -0.108
-0.108 to 0.021
0.021 to 0.145
0.145 to 0.607
0.607 to 4.621

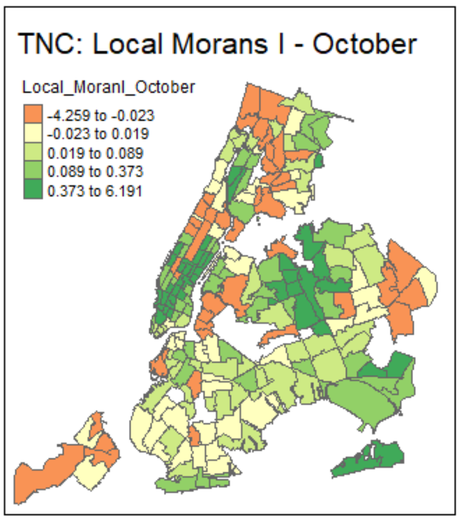

Taxi: Local Morans I - July
Local_MoranI_July
-3.116 to -0.050
-0.050 to 0.011
0.011 to 0.116
0.116 to 0.418
0.418 to 1.990

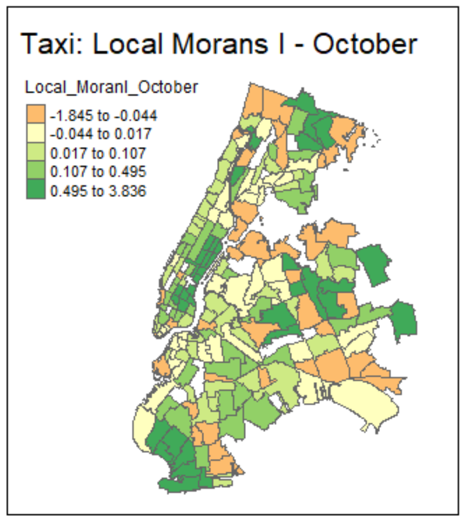

TNC: Local Morans I - August
Local_MoranI_August
-0.954 to -0.083
-0.083 to -0.004
-0.004 to 0.035
0.035 to 0.223
0.223 to 6.769

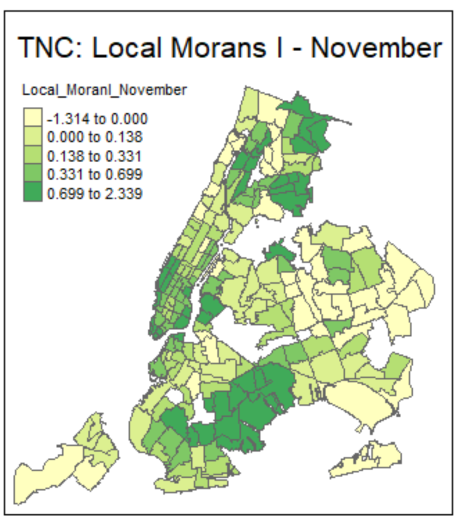

Taxi: Local Morans I - August
Local_MoranI_August
-4.785 to -0.014
-0.014 to 0.052
0.052 to 0.245
0.245 to 0.462
0.462 to 3.367

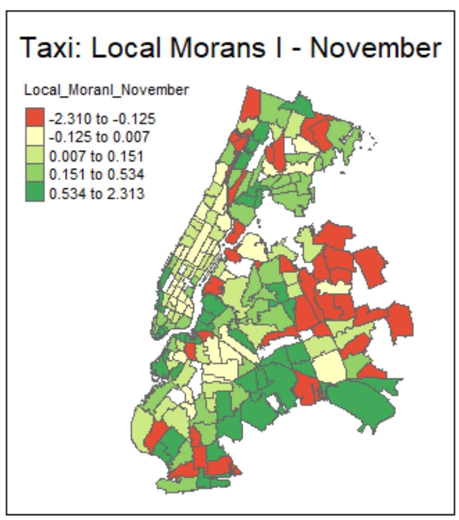

TNC: Local Morans I - September
Local_MoranI_September
-1.458 to -0.003
-0.003 to 0.131
0.131 to 0.377
0.377 to 0.801
0.801 to 4.796


Taxi: Local Morans I - September
Local_MoranI_September
-2.344 to -0.024
-0.024 to 0.019
0.019 to 0.132
0.132 to 0.338
0.338 to 3.421


TNC: Local Morans I - October
Local_MoranI_October
-4.259 to -0.023
-0.023 to 0.019
0.019 to 0.089
0.089 to 0.373
0.373 to 6.191


Taxi: Local Morans I - October
Local_MoranI_October
-1.845 to -0.044
-0.044 to 0.017
0.017 to 0.107
0.107 to 0.495
0.495 to 3.836


TNC: Local Morans I - November
Local_MoranI_November
-1.314 to 0.000
0.000 to 0.138
0.138 to 0.331
0.331 to 0.699
0.699 to 2.339


Taxi: Local Morans I - November
Local_MoranI_November
-2.310 to -0.125
-0.125 to 0.007
0.007 to 0.151
0.151 to 0.534
0.534 to 2.313

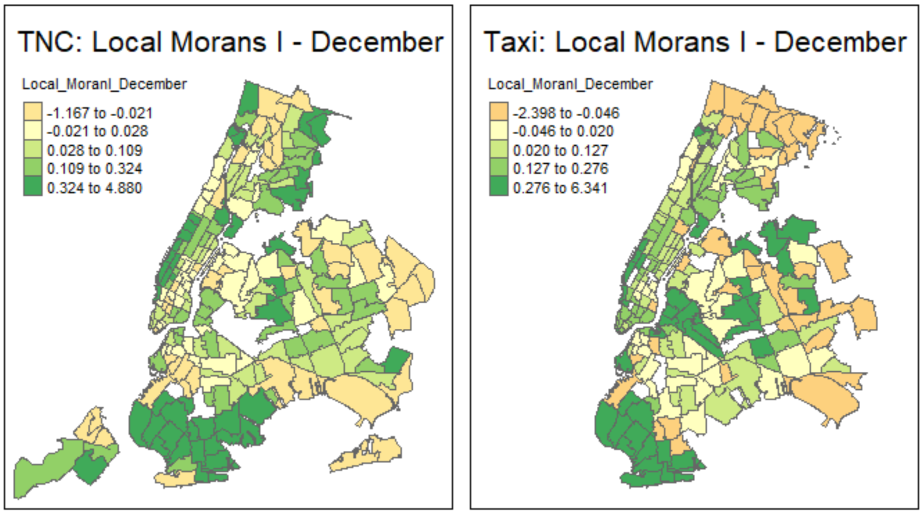


Figure 4.10 Local Moran's I on ARIMA/GARCH/MLR Residuals – Month

Chapter 5. Summary of R Code used in this Project

All code for the analyses described in Chapter 3 and 4 of this report can be found on the Github repositories linked below. The first link corresponds to the models used in chapter 3 while the second corresponds to the analysis performed in chapter 4.

1. https://github.com/pato6664/ALR-Models
2. https://github.com/pato6664/ARIMA-GARCH_SpatialAnalysis

Chapter 6. Summary and Conclusions

The summary and conclusions of each analysis has been detailed the individual chapters. The research carried out in this project has both methodological and substantive contributions. Much more can be done to better understand these travel models with a goal towards using the information for planning purposes.

6.1 Directions for Future Research

Many interesting lines of inquiry may be investigated in future research. First, Citi Bike's role in the marketplace of shared ride modes still seems unclear and needs closer scrutiny. While it does have weekly seasonal dynamics like the other modes, it has no appreciable increasing trend and it has no significant cross-correlation with the other modes on a daily level. Second, there are a significant number of outliers in the data set and a careful look into imputing them will be useful. Third, residuals from the models in Section 4 and Section 5 indicate some departure from normality and conditional homoscedasticity. Research into building more sophisticated models that can adequately incorporate time-varying volatility and heavy-tailed behavior will be attractive. Finally, disaggregating the data for NYC into areal units (boroughs or taxi zones) would enable us to study spatial and temporal dynamics in a richer modeling framework.

References

- Aitchison, J., 1986. The Statistical Analysis of Compositional Data. London: Chapman and Hall.
- Anselin, L. Local Indicators of Spatial Association—LISA, Geographical Analysis 27(2): 93–115, 1995.
- Box, G. & Cox, D., 1964. An Analysis of Transformations. Journal of The Royal Statistical Society. Series B 26(2), pp. 211-252.
- Bivand. R. et al. 2019. Package 'spdep'. https://cran.r-project.org/web/packages/spdep/spdep.pdf
- Brodeur, A. & Nield, K., 2018. An empirical analysis of taxi, Lyft and Uber rides: Evidence from weather shocks in NYC. Journal of Economic Behavior & Organization, Volume 152, pp. 1-16.
- Citi Bike [Data File], 2017. Citi Bike Trip History Data. https://www.citibikenyc.com/system-data
- Contreras, S. D. & Paz, A., 2018. The effects of ride-hailing companies on the taxicab industry in Las Vegas, Nevada. Transportation Research Part A: Policy and Practice, Volume 115, pp. 63-70.
- Conway, M. W., Salon, D. & King, D. A., 2018. Trends in Taxi Use and the Advent of Ridehailing, 1995–2017: Evidence from the US National Household Travel Survey. Urban Science, p. 79.
- DeMay, D., 2018. Report: Goodbye taxis, Seattleites take Uber, Lyft 3.5 times more often. https://www.seattlepi.com/local/transportation/article/Goodbye-taxis-Seattleites-take-Uber-way-more-12866867.php
- Durbin, J. & Koopman, S., 2012. Time Series Analysis by State Space Methods. 2 ed. Oxford: Oxford Unviersity Press.
- Egozcue, J., Pawlowsky-Glahn, V. & Mateu-Figueras, G., 2003. Isometric Logratio Transformations for Compositional Data Analysis. Mathematical Geology., Vol. 35, pp. 279-300.
- Erhardt, G., Graehler, M. J. & Mucci, R. A., 2019. Understanding the Recent Transit Ridership Decline in Major US Cities: Service Cuts or Emerging Modes?. Washington D.C., 98th Annual Meeting of the Transportation Research Board.
- Gerte, R. et al., 2019. Understanding the Relationships Between Demand for Shared Ride Modes: A Case Study Using Open Data From New York City. Washington D.C., 98th Annual Meeting of Transportation Research Board.
- Gessner, K., 2019. Rideshare: With IPOs looming, Uber leads market share, but Lyft has gained ground. https://blog.secondmeasure.com/datapoints/rideshare-industry-overview/
- Gimond, M. 2019. Intro to GIS and Spatial Analysis. Retrieved March 11, 2019, from https://mgimond.github.io/Spatial/index.html
- Hall, J. D., Palsson, C. & Price, J., 2018. Is Uber a substitute or complement for public transit?. Journal of Urban Economics, Volume 108, pp. 36-50.
- Hoffman, K., Sundararajan, A. & Ipeirotis, P. G., 2016. Ridesharing and the Use of Public Transportation. https://pdfs.semanticscholar.org/1461/e4b58fc0a1ddce2d033bb0ebc99361df5de6.pdf
- Jin, S. T., Kong, H. & Sui, D. Z., 2019. Uber, Public Transit, and Urban Transportation Equity: A Case Study in New York City. The Professional Geographer.
- Metropolitan Transit Authority (MTA) [Data File], 2017. Turnstile Data. http://web.mta.info/developers/turnstile.html
- NYC Taxi & Limousine Commission (TLC) [Data File], 2017. TLC Trip Data. http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml
- Pfaff, B., 2008. VAR, SVAR, and SVEC Models: Implementation Within R Package. Journal of Statistical Software, 27 (4).
- Ravishanker, N., Dey, D. & Iyengar, M., 2001. Compositional Time Series Analysis of Mortality Proportions. Communication in Statistics-Theory and Methods, pp. 2281-2291.
- Rayens, W. S. & Srinivasan, C., 1991. Estimation in Compositional Data Analysis. Journal of Chemometrics, 5(4), pp. 361-374.

- Rayle, L. et al., 2016. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transportation Policy, Volume 45, pp. 168-178.
- Ripley. B., Venables. B., Bates. D. M., Hornik. K., Gebhardt. A., and Firth., D. 2019. Package "Mass". https://cran.r-project.org/web/packages/MASS/MASS.pdf
- Schaller, B., 2018. The New Automobility: Lyft, Uber and the Future of American Cities, New York City: Schaller Consulting.
- Serhiyenko, V., 2014. Dynamic Compositional Modeling of Pedestrian Crash Counts on Urban Roads in Connecticut. Accident Analysis & Prevention, Volume 64, pp. 78-85.
- Shumway, R. H. & Stoffer, D. S., 2017. Time Series Analysis and Its Applications. 4 ed. Cham: Springer.
- Siddiqui, F., 2018. As ride hailing booms in D.C., it's not just eating into the taxi market - it's increasing vehicle trips. https://www.washingtonpost.com/local/trafficandcommuting/as-ride-hailing-booms-in-dc-its-not-just-eating-into-the-taxi-market--its-increasing-vehicle-trips/2018/04/23/d1990fde-4707-11e8-827e-190efaf1f1ee_story.html?noredirect=on&utm_term=.a2e517ba5b86
- Smith, T. & Brundson, T., 1989. The Time Series Analysis of Compositional Data. Proc. Amer. Statist. Assoc, pp. 26-32.
- Toman, P., Zhang, J., Ravishanker, N. & Konduri, K., 2019. Dynamic Predictive Models for Ridesourcing Services in New York City Using Daily Compositional Data (under review).
- Warerkar, T., 2017. Uber surpasses yellow cabs in average daily ridership in NYC: For the first time ever, Uber has topped the city's yellow cabs in daily ridership figures. https://ny.curbed.com/2017/10/13/16468716/uber-yellow-cab-nyc-surpass-ridership

APPENDICES

A. Brief Review of Statistical Methods

A.1. Dynamic Linear Model (DLM)

Dynamic linear modeling refers to a broad class of models that are readily generalizable to all manner of time series phenomenon. Indeed, the dynamic linear model is itself a subset of the even broader class of modeling frameworks known as state-space models. Two key characteristics define the state space modeling paradigm which is that we have a latent process denoted $\boldsymbol{\Theta_t}$ which is commonly called the state process. A key assumption of the state space model is that the state process is assumed to be Markovian, thus, the future and past observation are independent of the present. In addition, the observations, denoted as $\boldsymbol{Y_t}$, are independent given the state process $\boldsymbol{\Theta_t}$. This second feature implies then that any sort of dependence amongst observations is transmitted over time through the underlying state. The field of State Space Modeling or Dynamic Linear Modeling is a rich one and interested readers can find further expositions in Shumway and Stoffer (2017), Durbin and Koopman (2012), and numerous other texts on the subject.

A.2. Compositional Data Analysis

Compositional data can be broadly defined as data where all elements are non-negative and the sum of all data elements is equal to one. One can readily extend this idea of compositional data analysis to time series analysis where at every time point t, the sum of all the component time series are positive and sum to one. More formally, the G-variate compositional time series of positive value components can be defined as, $\vec{X} = \left(X_{t,1}, \ldots, X_{t,G}\right)$ for $t = 1, \ldots, T$, where the structure is defined by the $g = G - 1$ components such that, X_t lies in the g-dimensional simplex:

$$\boldsymbol{S^g = \left[\left(X_{t,1}, \ldots, X_{t,g}\right): X_{t,1} > 0, \ldots, X_{t,g} > 0; X_{t,1} + \cdots + X_{t,g} = 1\right]}$$

In order to cope with issues such as non-normality, compositional data analysis is often performed via some sort of suitable transformation of the g-dimensional simplex into the g-dimensional Euclidean space $\mathbb{R}^{g}$. One of the quintessential data transformation approaches of compositional data analysis was the Additive Log Ratio (ALR) transformation and the Centered Log Ratio transformation (CLR), see Aitchison (1986). Rayens and Srinivasan (1991) generalized the ideas of compositional data analysis to include the Box-Cox transformation (1964), and Egozcue et al. (2003) expanded the topic through the development of the Isometric Log Ratio (ILR) transformation. Aitchison (1986), Smith and Brunsdon (1989), Ravishanker et al. (2001), and Serhiyenko et al. (2014) all discuss the topic of applying compositional methods in the context of time series analysis, the last paper dealing with transportation safety. The primary topic of interest in these papers is the use of some form of the generalized Box-Cox transformation, including the ALR transformation for compositional time series and the subsequent use of standard time series methods to analyze real, vector-valued data.

B. Additional Plots of Local Moran's I

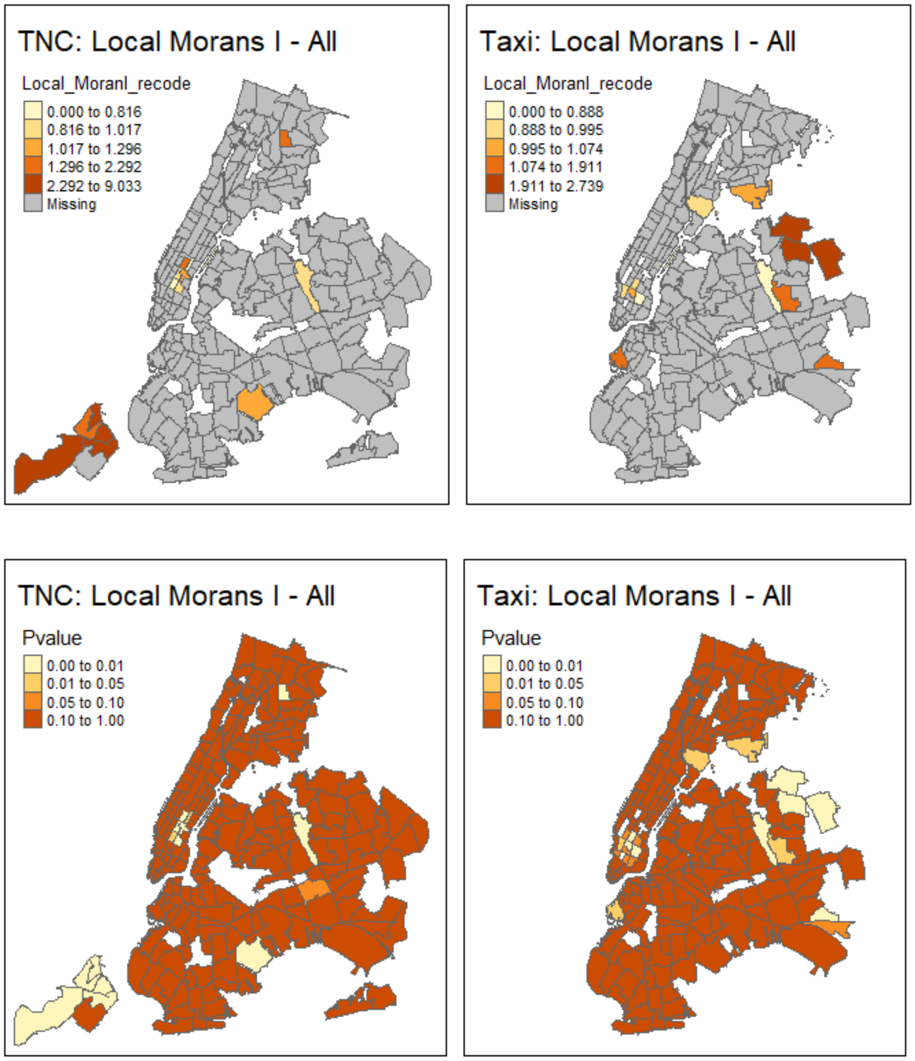
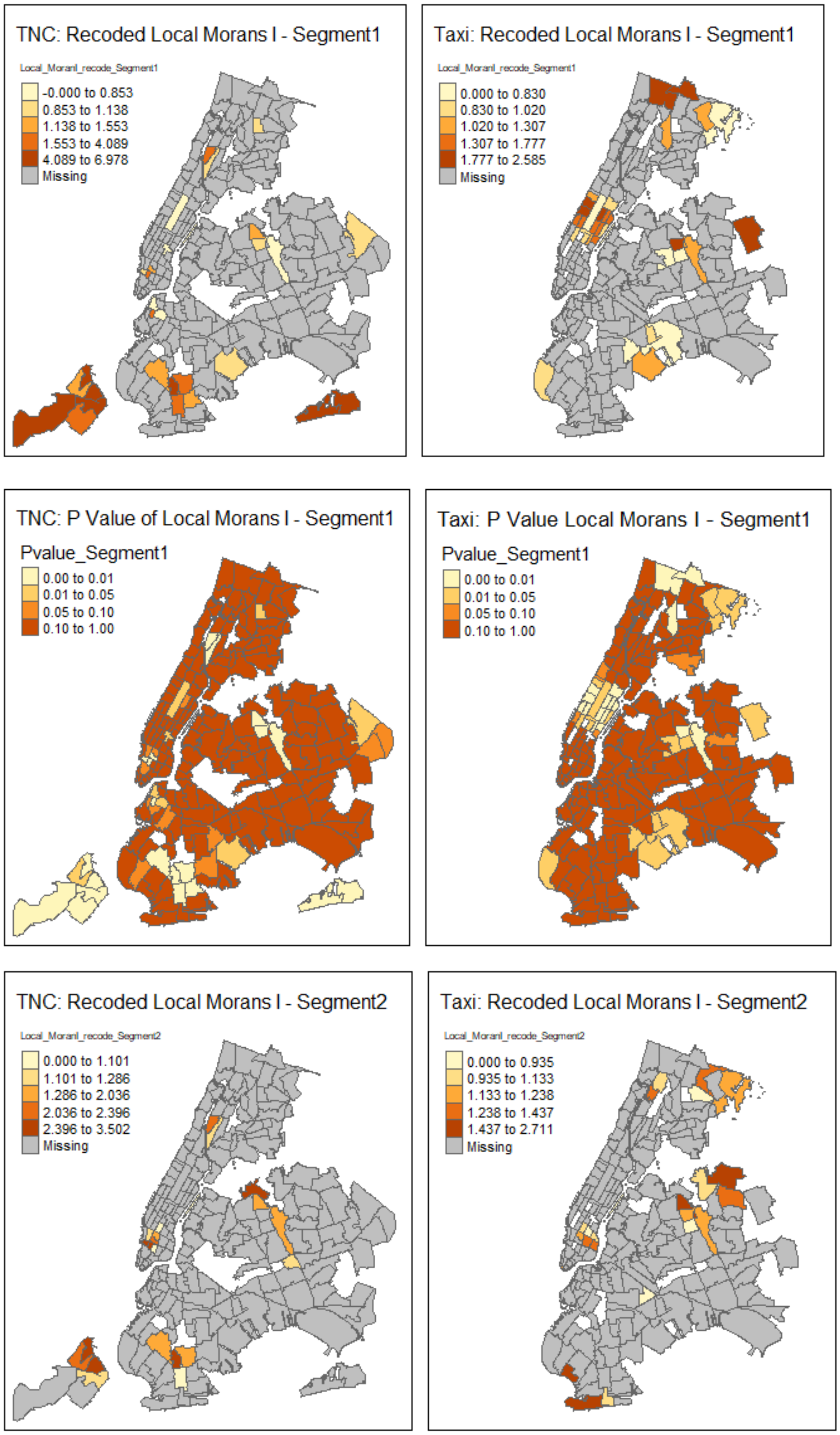
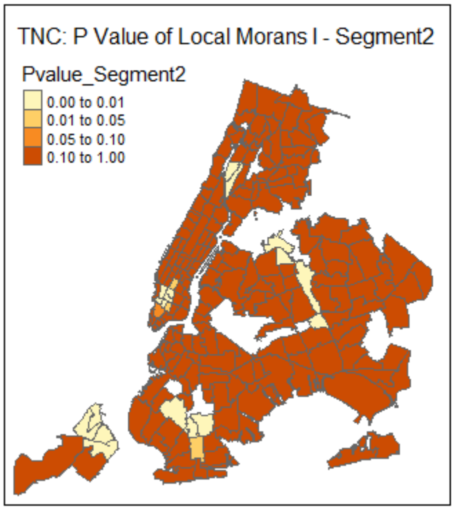
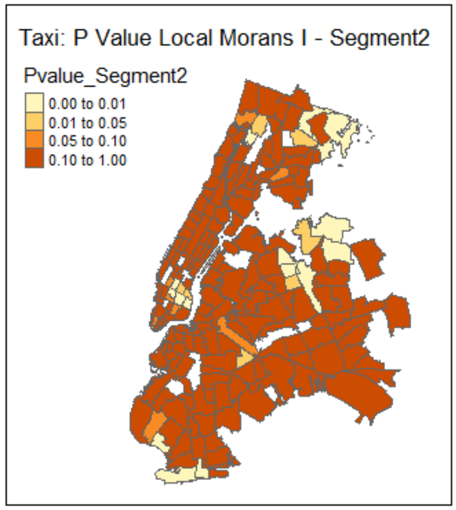
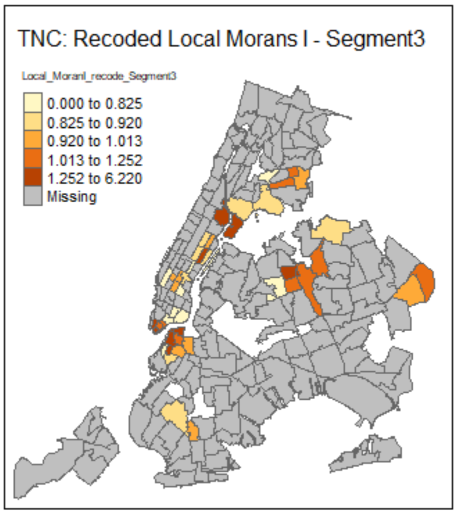


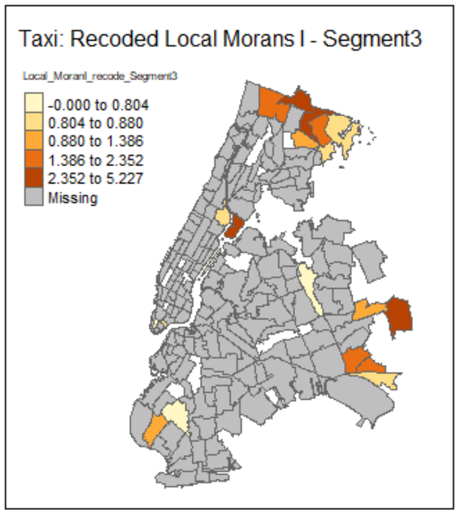
Figure B. 1 Recoded Local Moran's I and P-value of ARIMA/GARCH Residuals – Full Length

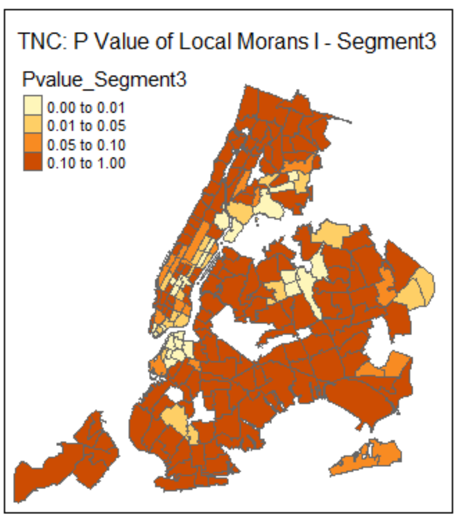

TNC: Recoded Local Morans I - Segment1
Local_MoranI_recode_Segment1
-0.000 to 0.853
0.853 to 1.138
1.138 to 1.553
1.553 to 4.089
4.089 to 6.978
Missing
Taxi: Recoded Local Morans I - Segment1
Local_MoranI_recode_Segment1
0.000 to 0.830
0.830 to 1.020
1.020 to 1.307
1.307 to 1.777
1.777 to 2.585
Missing
TNC: P Value of Local Morans I - Segment1
Pvalue_Segment1
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - Segment1
Pvalue_Segment1
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
TNC: Recoded Local Morans I - Segment2
Local_MoranI_recode_Segment2
0.000 to 1.101
1.101 to 1.286
1.286 to 2.036
2.036 to 2.396
2.396 to 3.502
Missing
Taxi: Recoded Local Morans I - Segment2
Local_MoranI_recode_Segment2
0.000 to 0.935
0.935 to 1.133
1.133 to 1.238
1.238 to 1.437
1.437 to 2.711
Missing

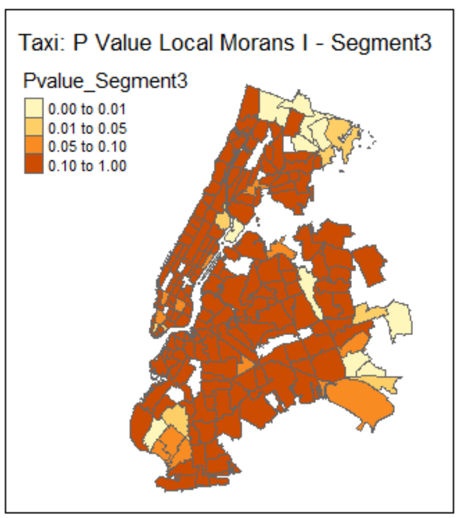

TNC: P Value of Local Morans I - Segment2
Pvalue_Segment2
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - Segment2
Pvalue_Segment2
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


TNC: Recoded Local Morans I - Segment3
Local_MoranI_recode_Segment3
0.000 to 0.825
0.825 to 0.920
0.920 to 1.013
1.013 to 1.252
1.252 to 6.220
Missing


Taxi: Recoded Local Morans I - Segment3
Local_MoranI_recode_Segment3
-0.000 to 0.804
0.804 to 0.880
0.880 to 1.386
1.386 to 2.352
2.352 to 5.227
Missing


TNC: P Value of Local Morans I - Segment3
Pvalue_Segment3
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - Segment3
Pvalue_Segment3
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

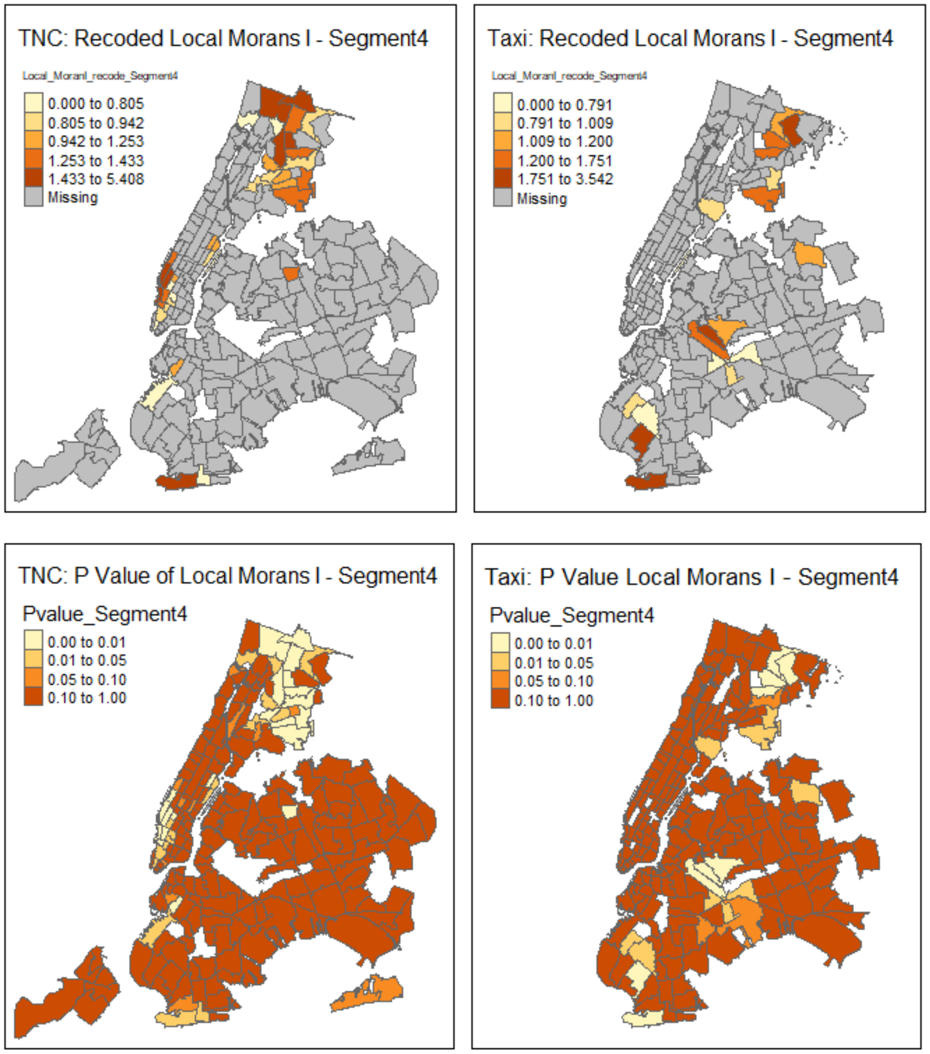
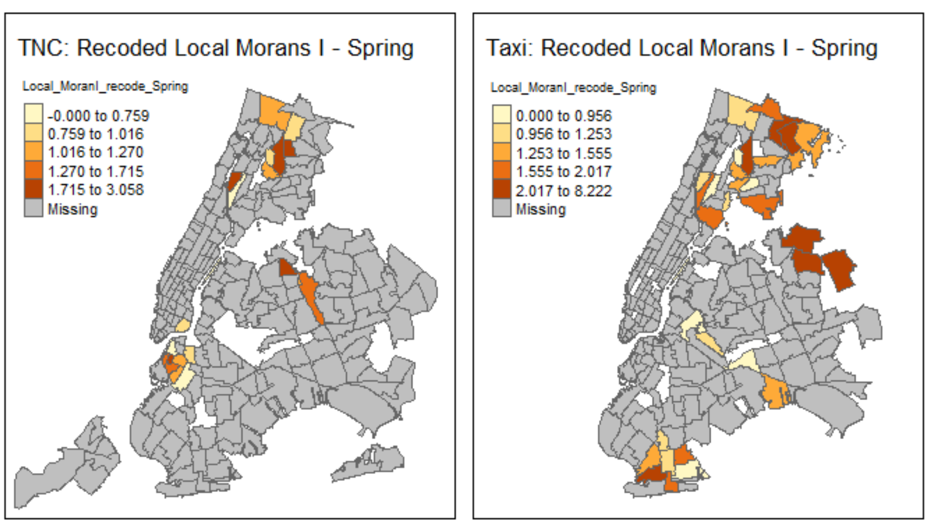
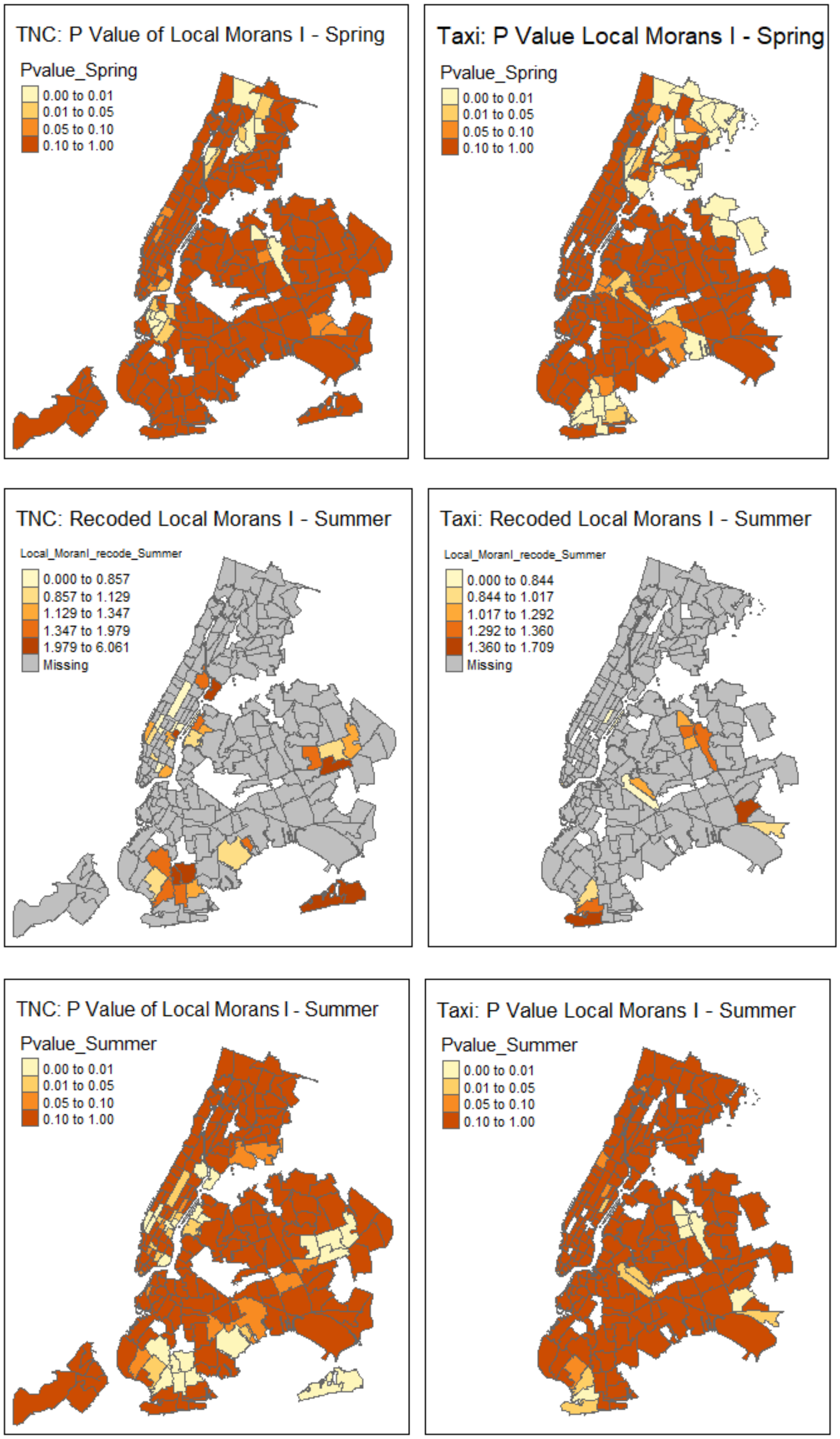
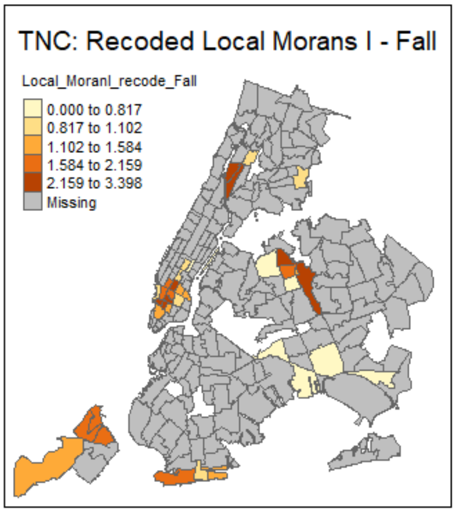
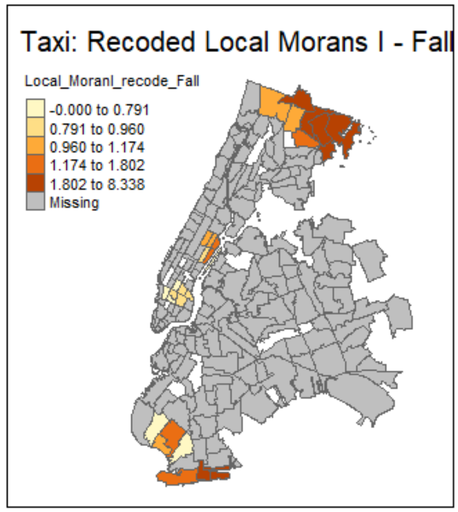


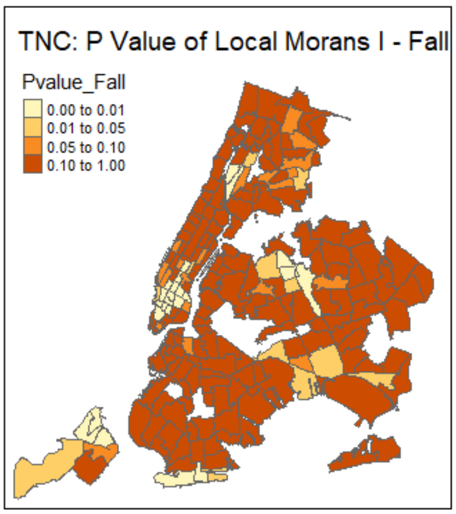
Figure B. 2 Recoded Local Moran's I and P-value of ARIMA/GARCH Residuals – Segment

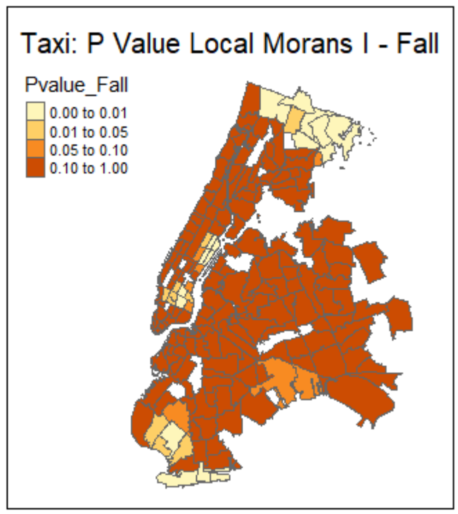


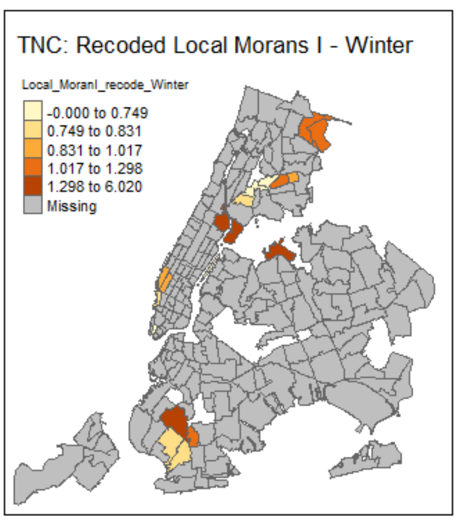

TNC: P Value of Local Morans I - Spring
Pvalue_Spring
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - Spring
Pvalue_Spring
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
TNC: Recoded Local Morans I - Summer
Local_MoranI_recode_Summer
0.000 to 0.857
0.857 to 1.129
1.129 to 1.347
1.347 to 1.979
1.979 to 6.061
Missing
Taxi: Recoded Local Morans I - Summer
Local_MoranI_recode_Summer
0.000 to 0.844
0.844 to 1.017
1.017 to 1.292
1.292 to 1.360
1.360 to 1.709
Missing
TNC: P Value of Local Morans I - Summer
Pvalue_Summer
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - Summer
Pvalue_Summer
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

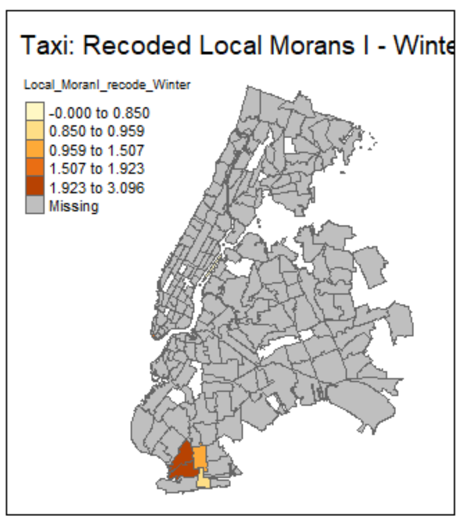

TNC: Recoded Local Morans I - Fall
Local_MoranI_recode_Fall
0.000 to 0.817
0.817 to 1.102
1.102 to 1.584
1.584 to 2.159
2.159 to 3.398
Missing


Taxi: Recoded Local Morans I - Fall
Local_MoranI_recode_Fall
-0.000 to 0.791
0.791 to 0.960
0.960 to 1.174
1.174 to 1.802
1.802 to 8.338
Missing


TNC: P Value of Local Morans I - Fall
Pvalue_Fall
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - Fall
Pvalue_Fall
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


TNC: Recoded Local Morans I - Winter
Local_MoranI_recode_Winter
-0.000 to 0.749
0.749 to 0.831
0.831 to 1.017
1.017 to 1.298
1.298 to 6.020
Missing


Taxi: Recoded Local Morans I - Winte
Local_MoranI_recode_Winter
-0.000 to 0.850
0.850 to 0.959
0.959 to 1.507
1.507 to 1.923
1.923 to 3.096
Missing

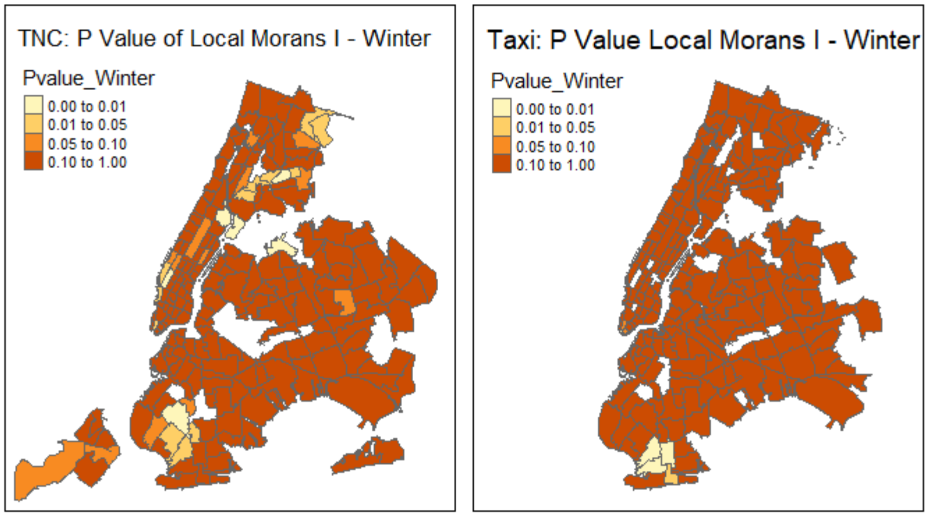
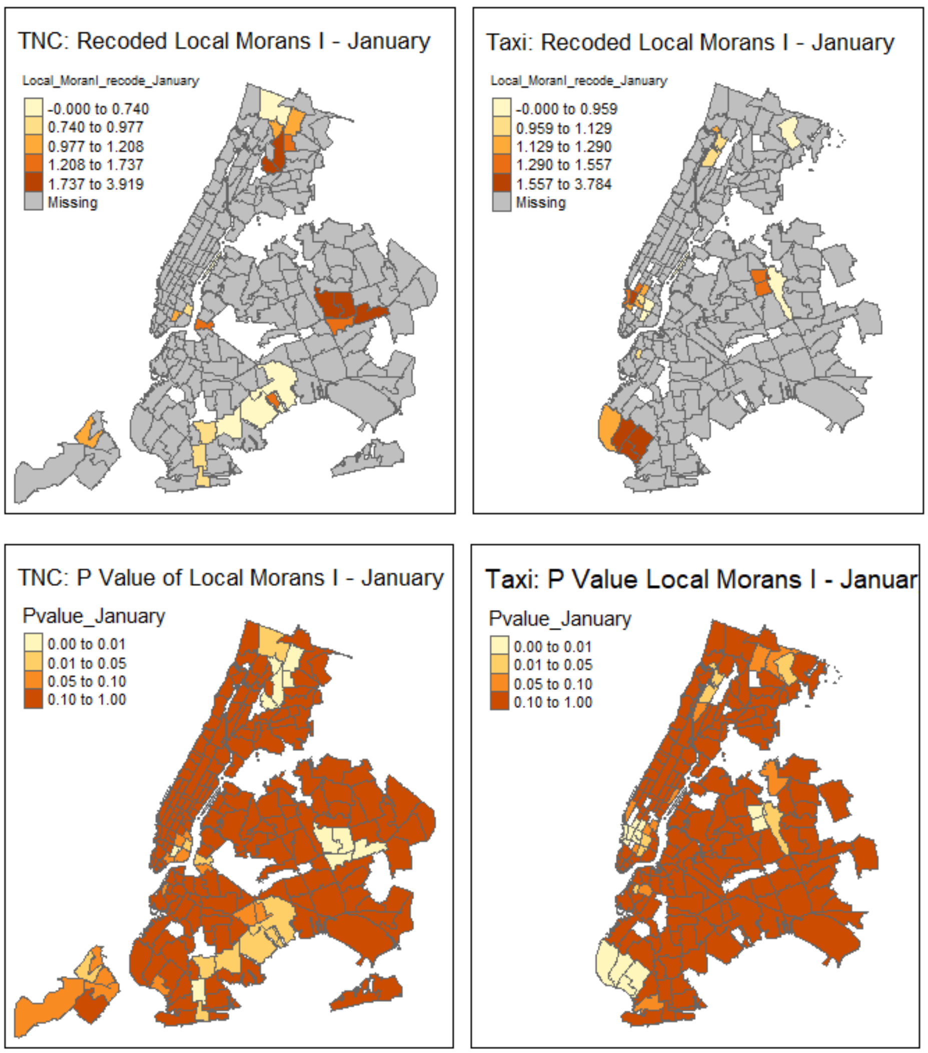
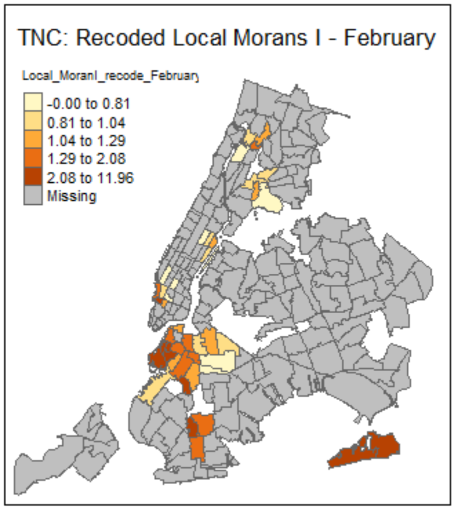
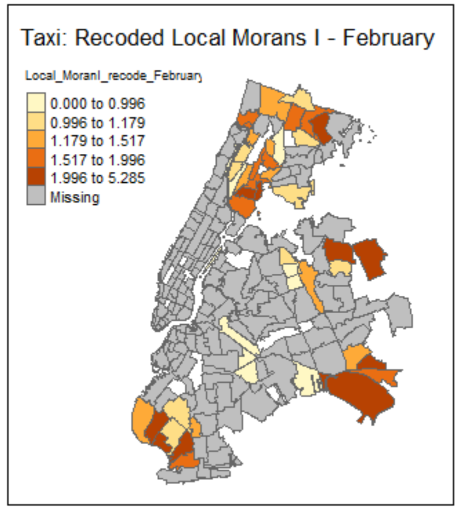
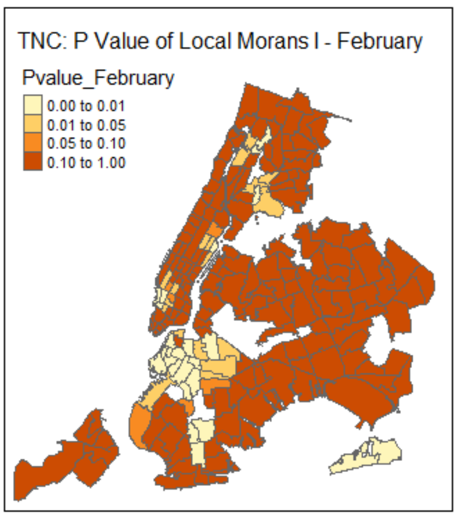


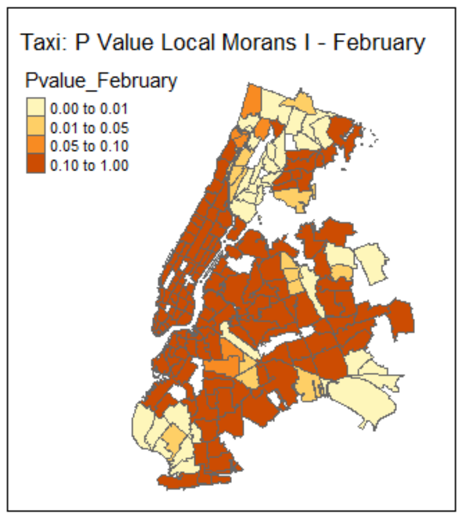
Figure B. 3 Recoded Local Moran's I and P-value of ARIMA/GARCH Residuals – Season

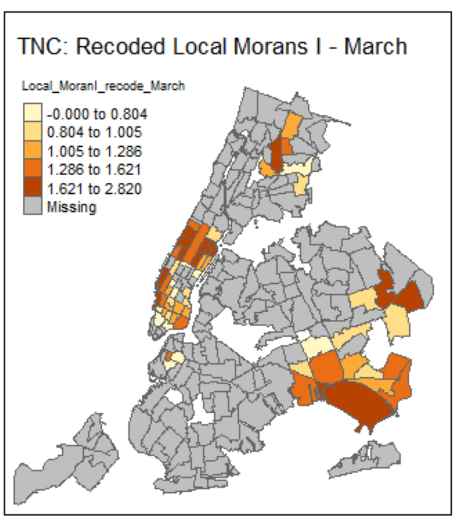


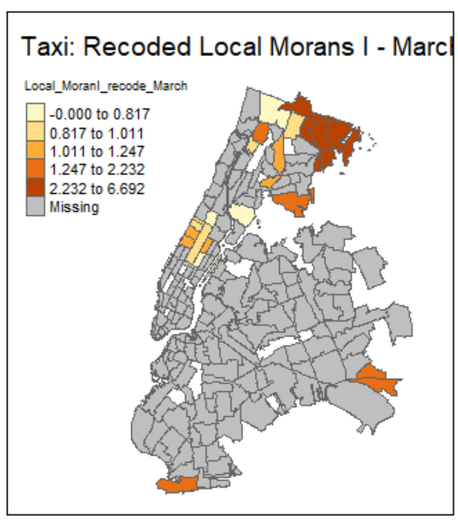

TNC: Recoded Local Morans I - February
Local_MoranI_recode_Februar
-0.00 to 0.81
0.81 to 1.04
1.04 to 1.29
1.29 to 2.08
2.08 to 11.96
Missing

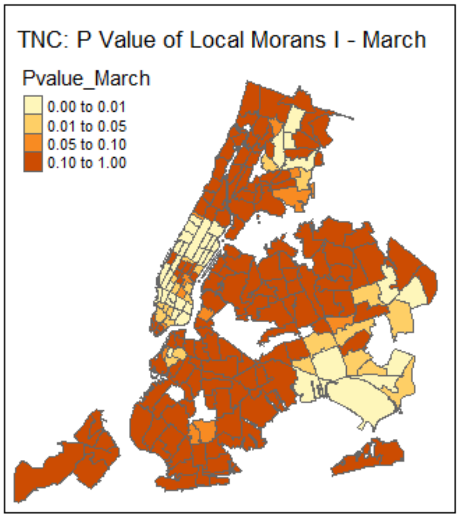

Taxi: Recoded Local Morans I - February
Local_MoranI_recode_Februar
0.000 to 0.996
0.996 to 1.179
1.179 to 1.517
1.517 to 1.996
1.996 to 5.285
Missing

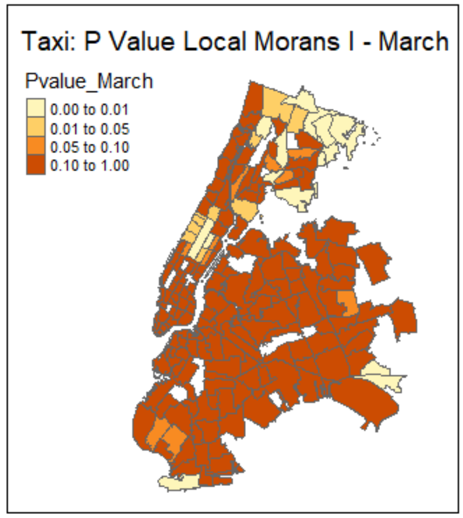

TNC: P Value of Local Morans I - February
Pvalue_February
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

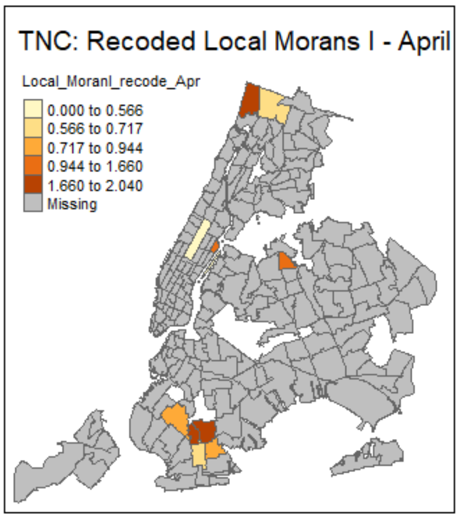

Taxi: P Value Local Morans I - February
Pvalue_February
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

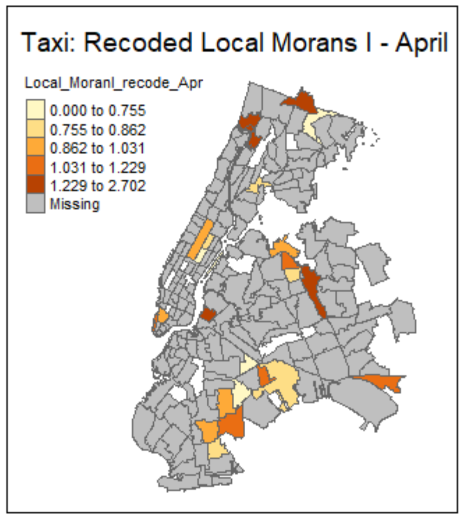

TNC: Recoded Local Morans I - March
Local_MoranI_recode_March
-0.000 to 0.804
0.804 to 1.005
1.005 to 1.286
1.286 to 1.621
1.621 to 2.820
Missing

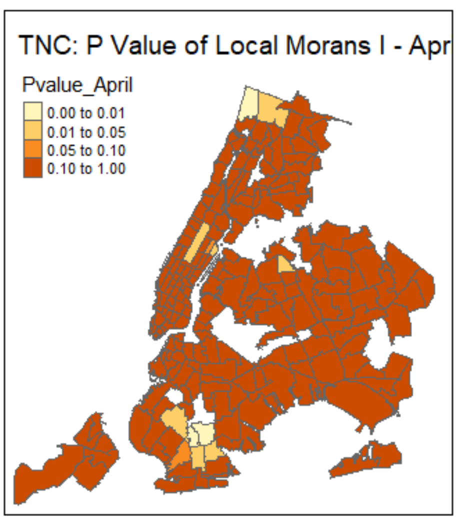

Taxi: Recoded Local Morans I - Marc
Local_MoranI_recode_March
-0.000 to 0.817
0.817 to 1.011
1.011 to 1.247
1.247 to 2.232
2.232 to 6.692
Missing

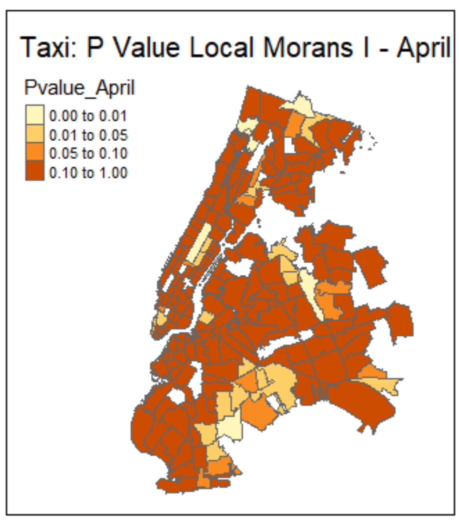

TNC: P Value of Local Morans I - March
Pvalue_March
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

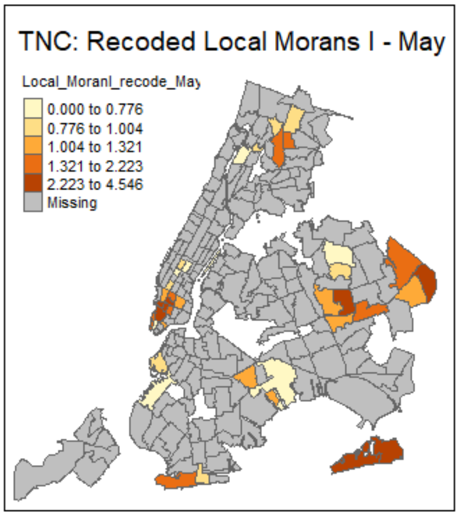

Taxi: P Value Local Morans I - March
Pvalue_March
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

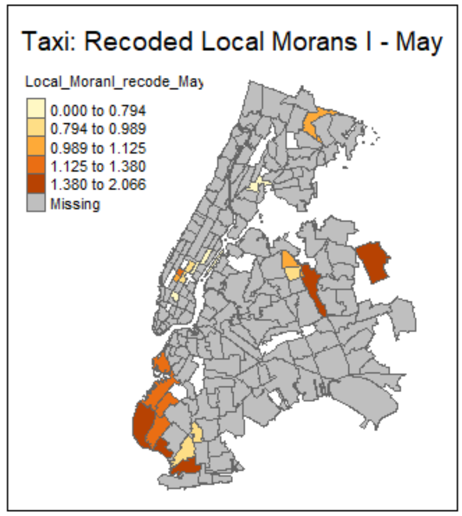

TNC: Recoded Local Morans I - April
Local_MoranI_recode_Apr
0.000 to 0.566
0.566 to 0.717
0.717 to 0.944
0.944 to 1.660
1.660 to 2.040
Missing

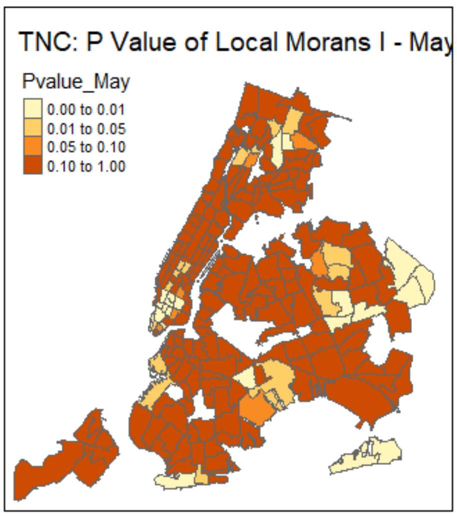

Taxi: Recoded Local Morans I - April
Local_MoranI_recode_Apr
0.000 to 0.755
0.755 to 0.862
0.862 to 1.031
1.031 to 1.229
1.229 to 2.702
Missing

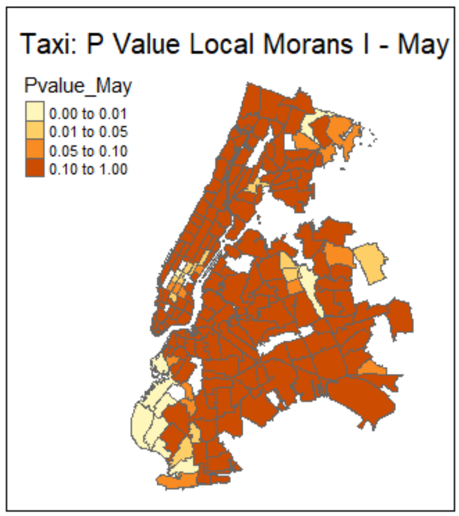

TNC: P Value of Local Morans I - Apr
Pvalue_April
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

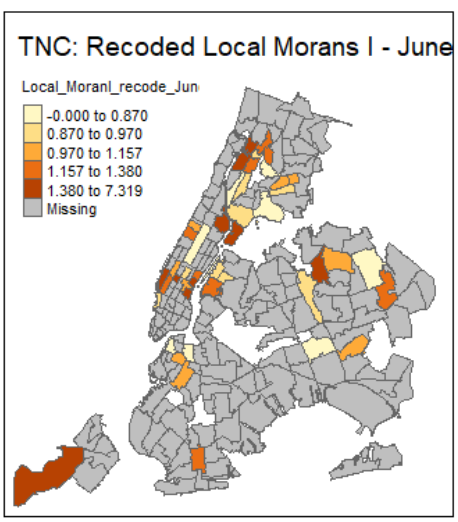

Taxi: P Value Local Morans I - April
Pvalue_April
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

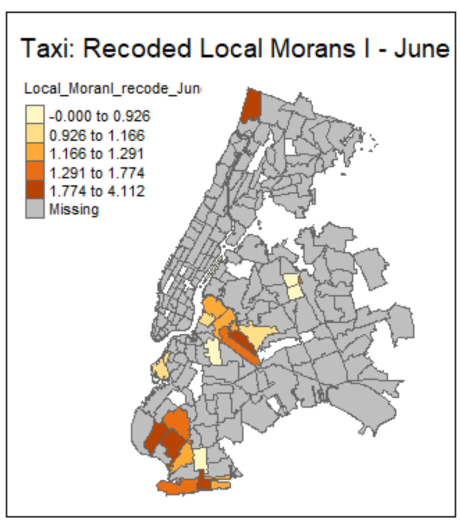

TNC: Recoded Local Morans I - May
Local_MoranI_recode_May
0.000 to 0.776
0.776 to 1.004
1.004 to 1.321
1.321 to 2.223
2.223 to 4.546
Missing

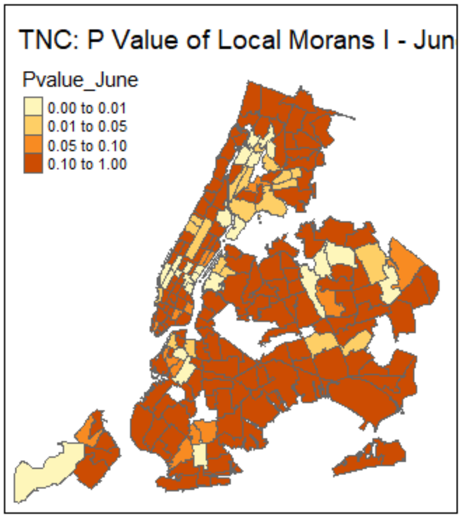

Taxi: Recoded Local Morans I - May
Local_MoranI_recode_May
0.000 to 0.794
0.794 to 0.989
0.989 to 1.125
1.125 to 1.380
1.380 to 2.066
Missing

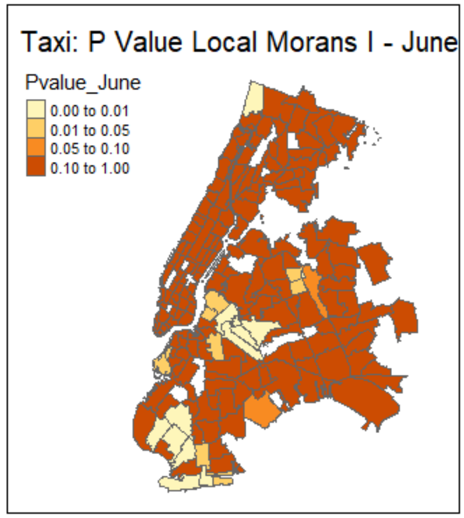

TNC: P Value of Local Morans I - May
Pvalue_May
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

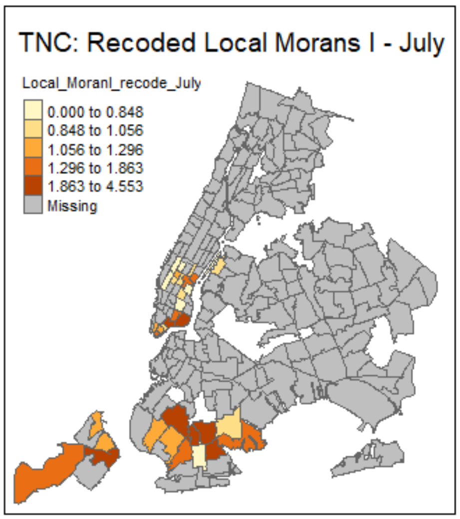

Taxi: P Value Local Morans I - May
Pvalue_May
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

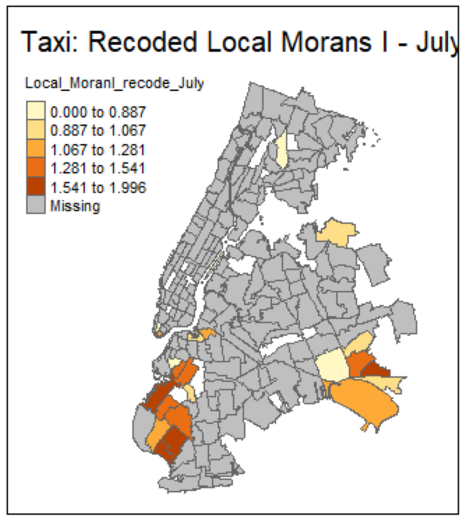

TNC: Recoded Local Morans I - June
Local_MoranI_recode_Jun
-0.000 to 0.870
0.870 to 0.970
0.970 to 1.157
1.157 to 1.380
1.380 to 7.319
Missing

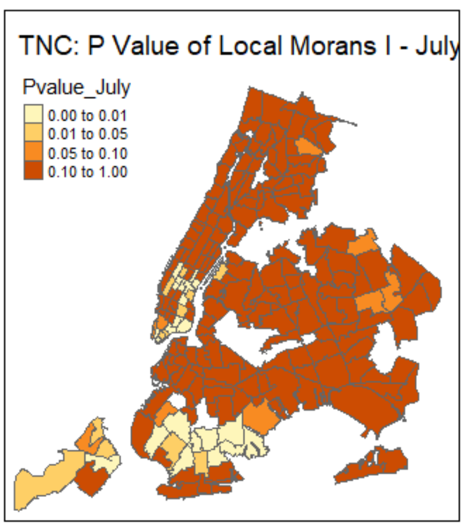

Taxi: Recoded Local Morans I - June
Local_MoranI_recode_Jun
-0.000 to 0.926
0.926 to 1.166
1.166 to 1.291
1.291 to 1.774
1.774 to 4.112
Missing

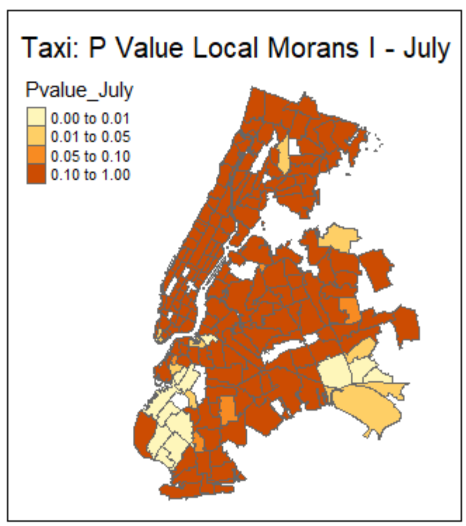

TNC: P Value of Local Morans I - Jun
Pvalue_June
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

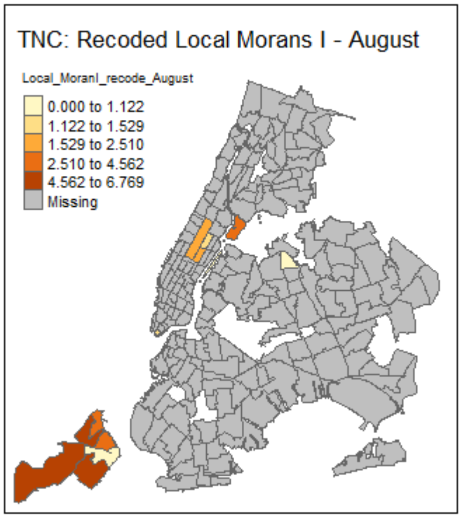

Taxi: P Value Local Morans I - June
Pvalue_June
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

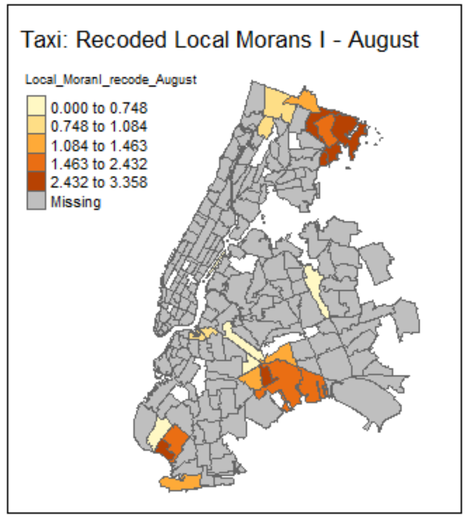

TNC: Recoded Local Morans I - July
Local_MoranI_recode_July
0.000 to 0.848
0.848 to 1.056
1.056 to 1.296
1.296 to 1.863
1.863 to 4.553
Missing

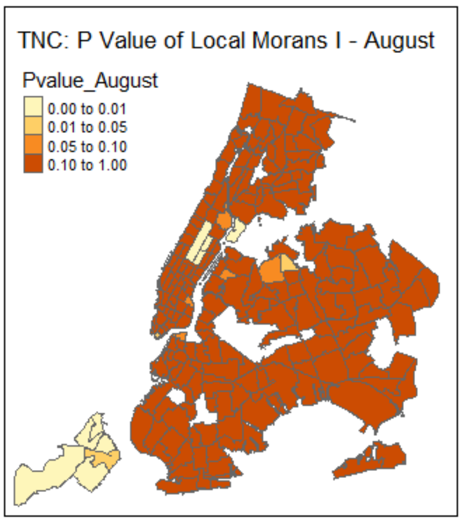

Taxi: Recoded Local Morans I - July
Local_MoranI_recode_July
0.000 to 0.887
0.887 to 1.067
1.067 to 1.281
1.281 to 1.541
1.541 to 1.996
Missing

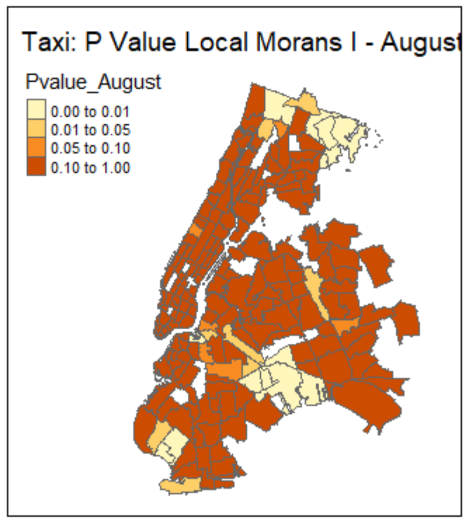

TNC: P Value of Local Morans I - July
Pvalue_July
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

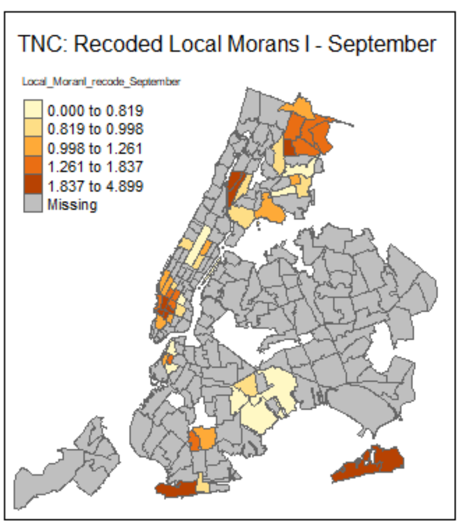

Taxi: P Value Local Morans I - July
Pvalue_July
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

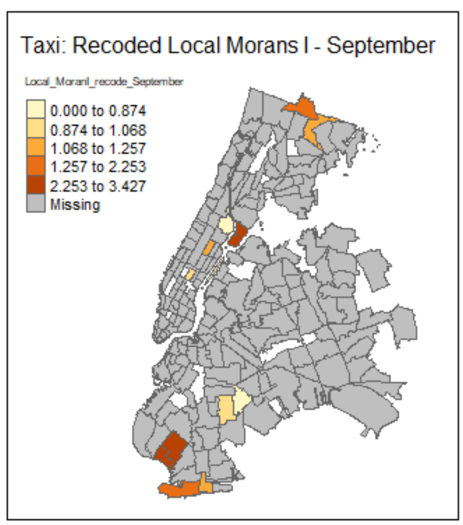

TNC: Recoded Local Morans I - August
Local_MoranI_recode_August
0.000 to 1.122
1.122 to 1.529
1.529 to 2.510
2.510 to 4.562
4.562 to 6.769
Missing

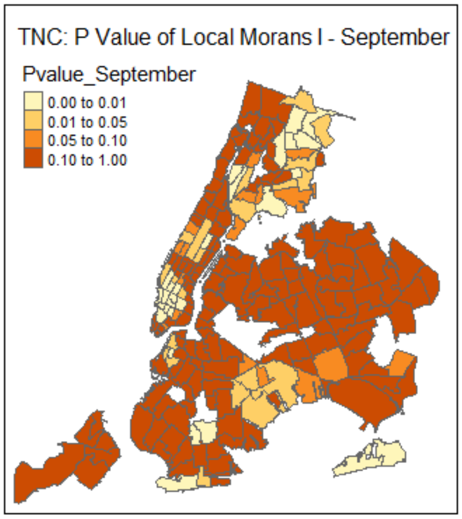

Taxi: Recoded Local Morans I - August
Local_MoranI_recode_August
0.000 to 0.748
0.748 to 1.084
1.084 to 1.463
1.463 to 2.432
2.432 to 3.358
Missing

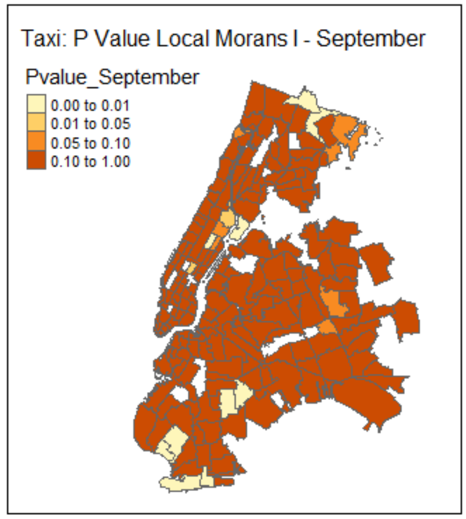

TNC: P Value of Local Morans I - August
Pvalue_August
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

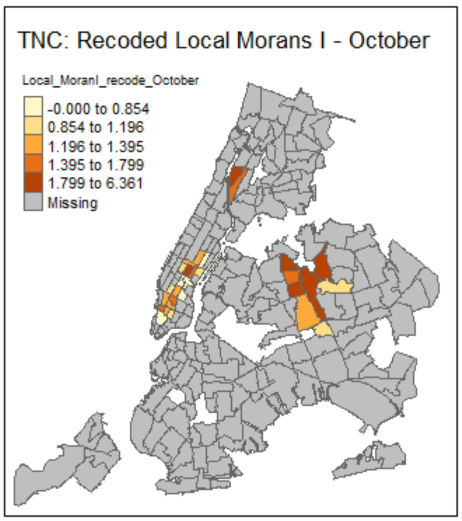

Taxi: P Value Local Morans I - August
Pvalue_August
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

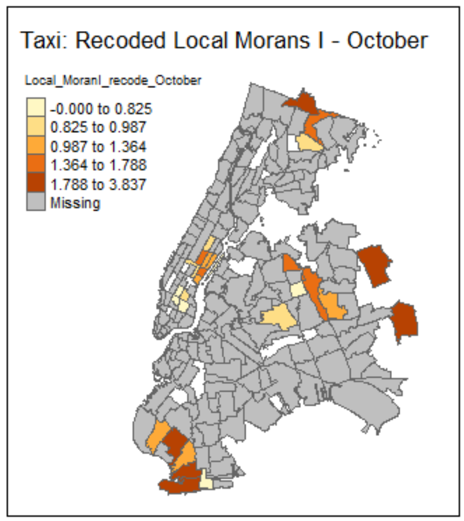

TNC: Recoded Local Morans I - September
Local_MoranI_recode_September
0.000 to 0.819
0.819 to 0.998
0.998 to 1.261
1.261 to 1.837
1.837 to 4.899
Missing

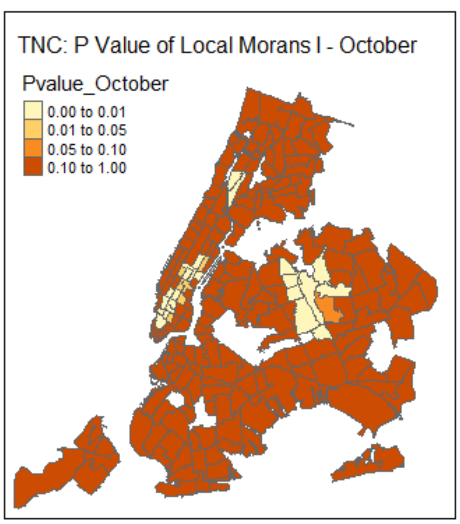

Taxi: Recoded Local Morans I - September
Local_MoranI_recode_September
0.000 to 0.874
0.874 to 1.068
1.068 to 1.257
1.257 to 2.253
2.253 to 3.427
Missing

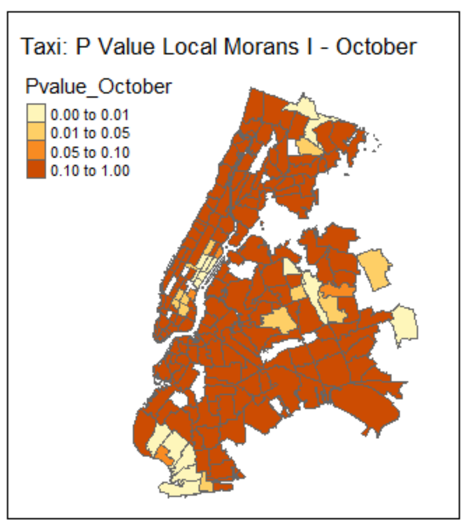

TNC: P Value of Local Morans I - September
Pvalue_September
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

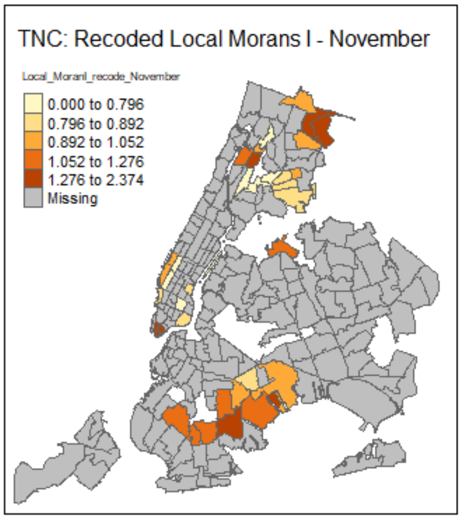

Taxi: P Value Local Morans I - September
Pvalue_September
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

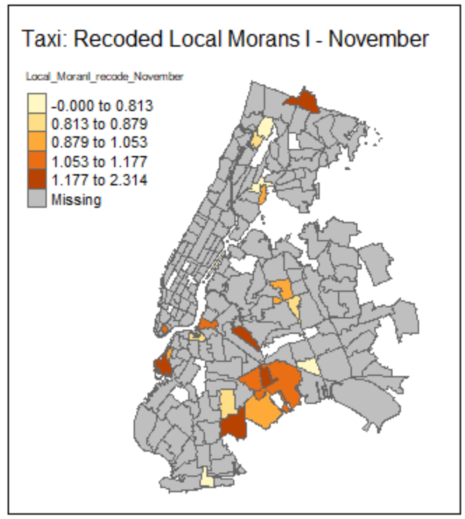

TNC: Recoded Local Morans I - October
Local_MoranI_recode_October
-0.000 to 0.854
0.854 to 1.196
1.196 to 1.395
1.395 to 1.799
1.799 to 6.361
Missing

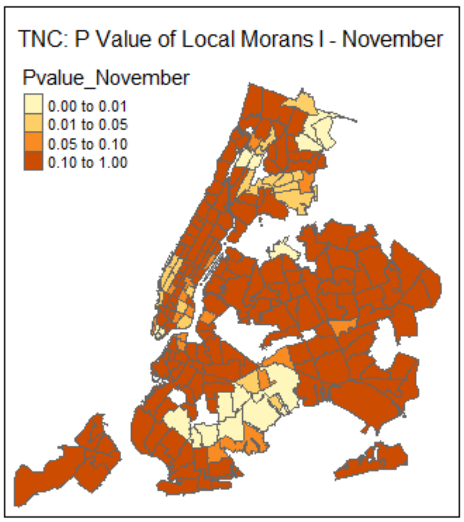

Taxi: Recoded Local Morans I - October
Local_MoranI_recode_October
-0.000 to 0.825
0.825 to 0.987
0.987 to 1.364
1.364 to 1.788
1.788 to 3.837
Missing

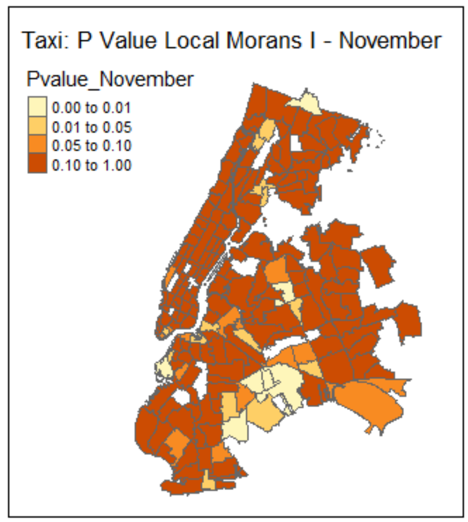

TNC: P Value of Local Morans I - October
Pvalue_October
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

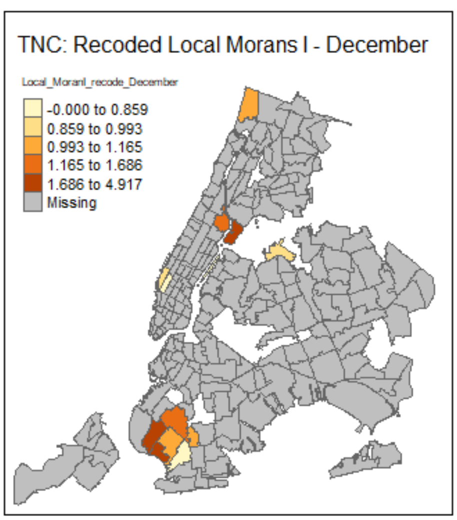

Taxi: P Value Local Morans I - October
Pvalue_October
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

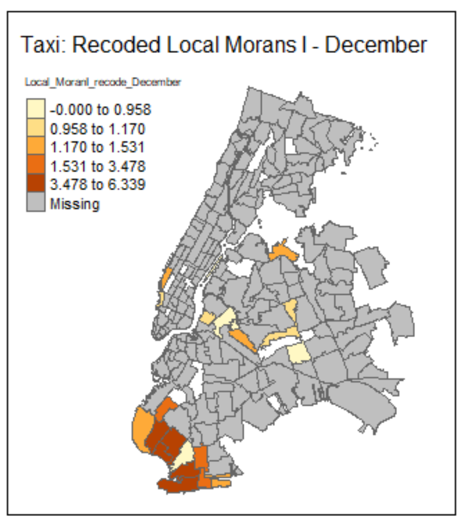

TNC: Recoded Local Morans I - November
Local_MoranI_recode_November
0.000 to 0.796
0.796 to 0.892
0.892 to 1.052
1.052 to 1.276
1.276 to 2.374
Missing


Taxi: Recoded Local Morans I - November
Local_MoranI_recode_November
-0.000 to 0.813
0.813 to 0.879
0.879 to 1.053
1.053 to 1.177
1.177 to 2.314
Missing


TNC: P Value of Local Morans I - November
Pvalue_November
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - November
Pvalue_November
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


TNC: Recoded Local Morans I - December
Local_MoranI_recode_December
-0.000 to 0.859
0.859 to 0.993
0.993 to 1.165
1.165 to 1.686
1.686 to 4.917
Missing


Taxi: Recoded Local Morans I - December
Local_MoranI_recode_December
-0.000 to 0.958
0.958 to 1.170
1.170 to 1.531
1.531 to 3.478
3.478 to 6.339
Missing

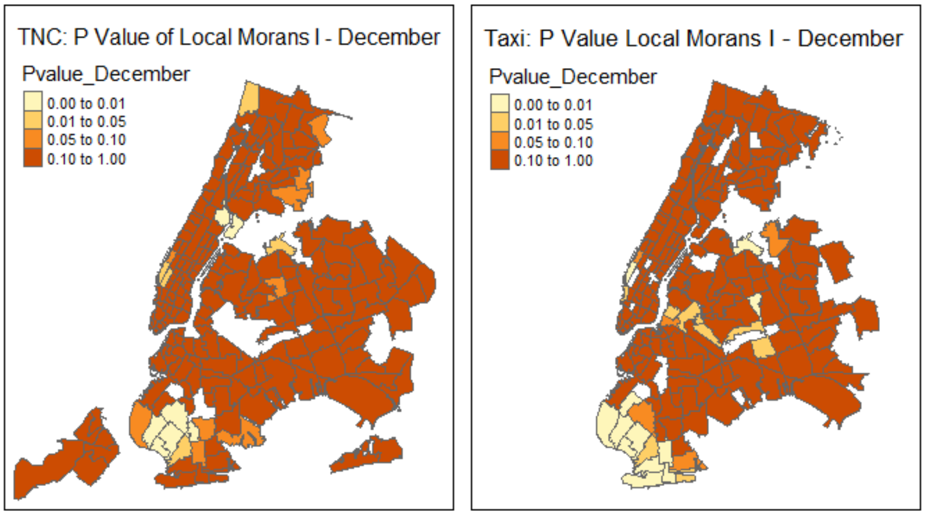


Figure B. 4 Recoded Local Moran's I and P-value of ARIMA/GARCH Residuals – Month

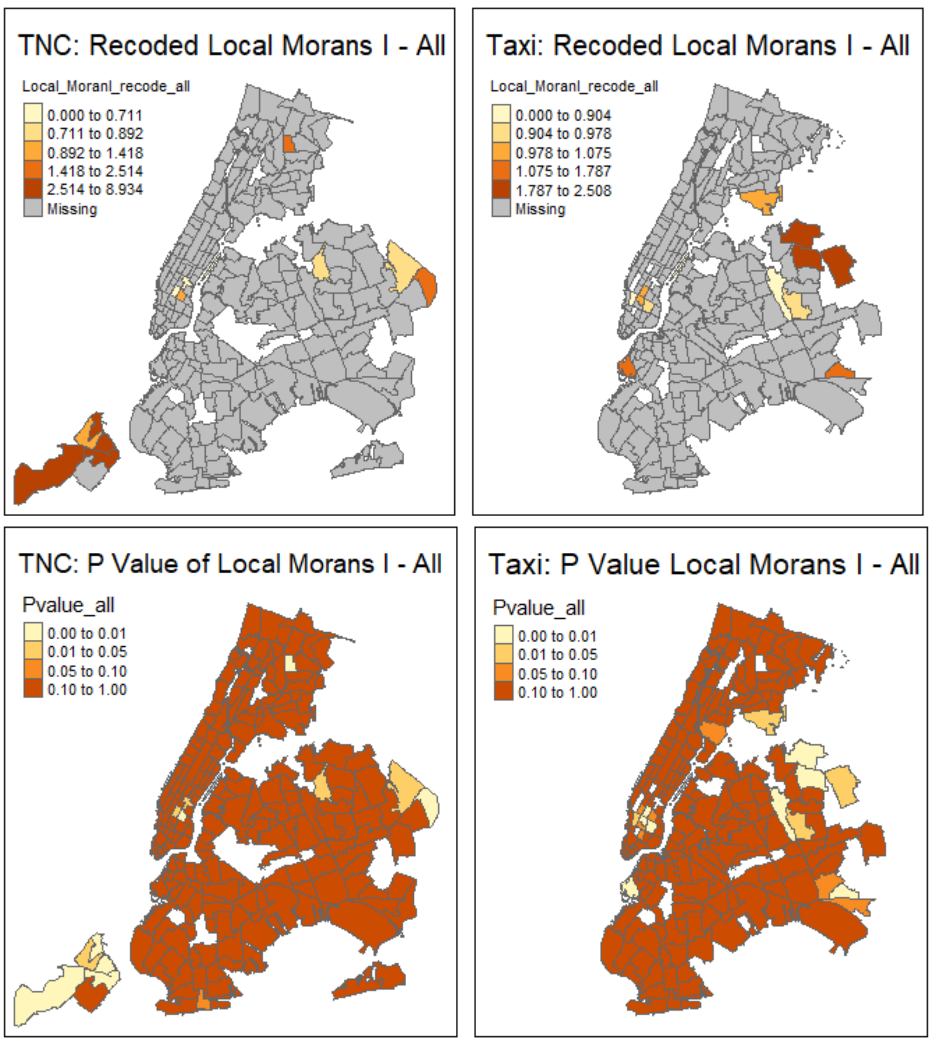
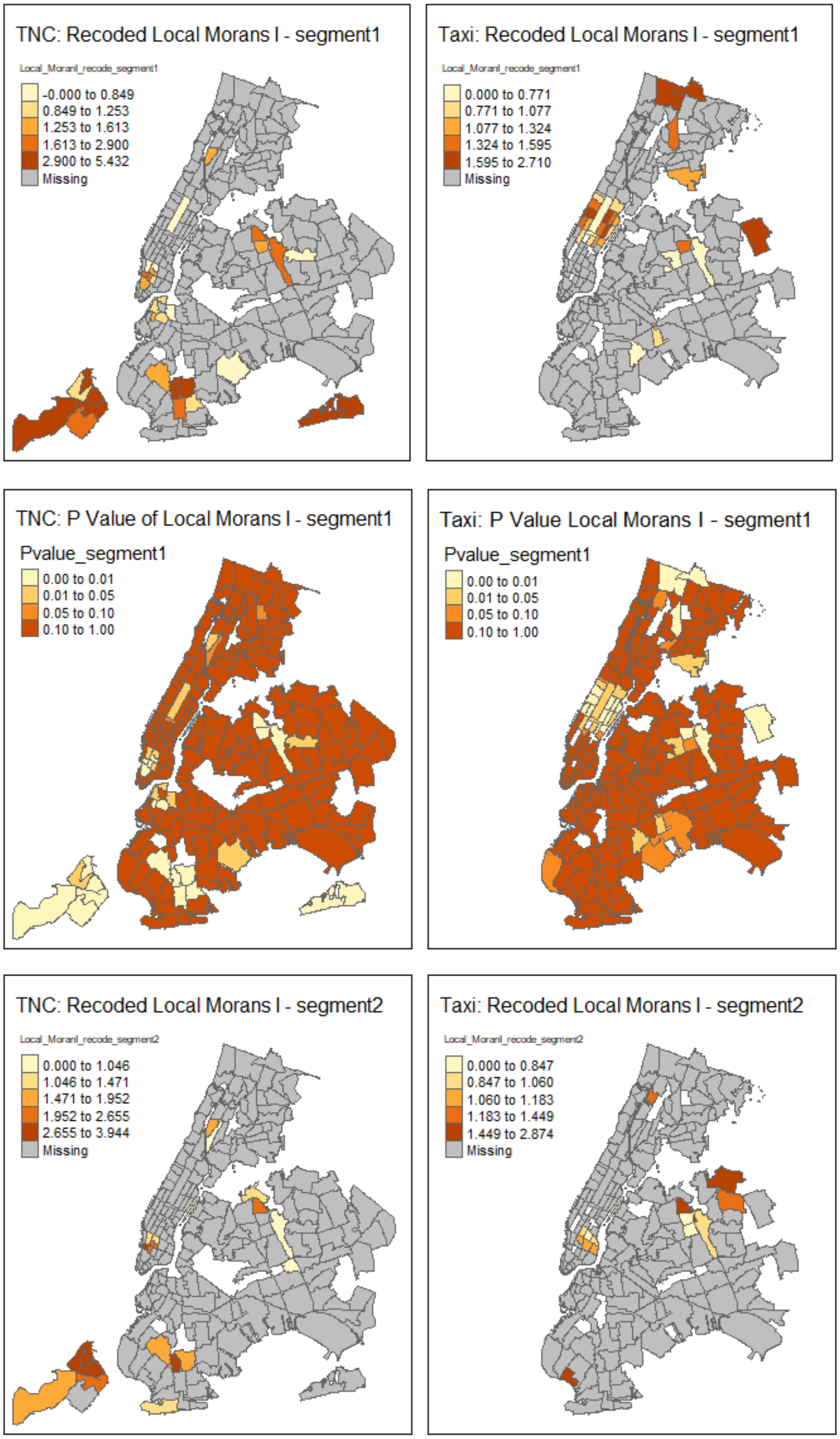
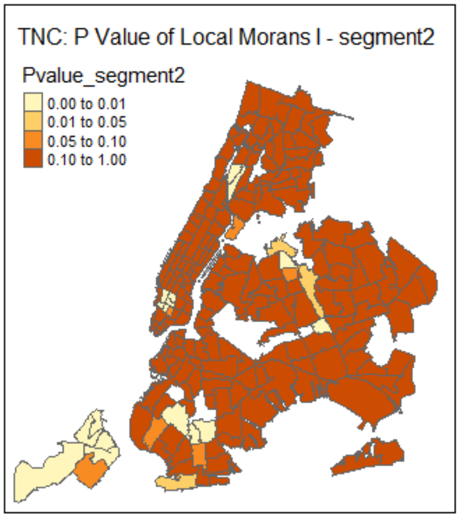
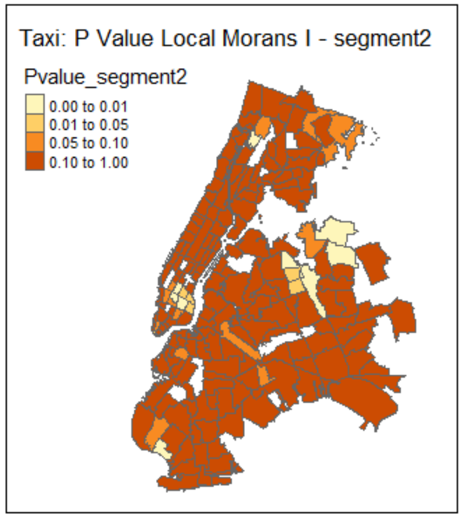
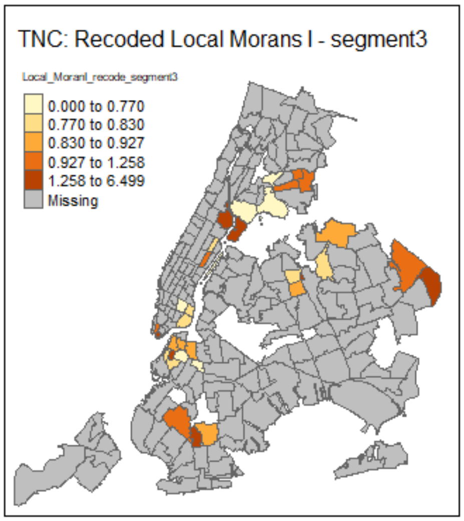


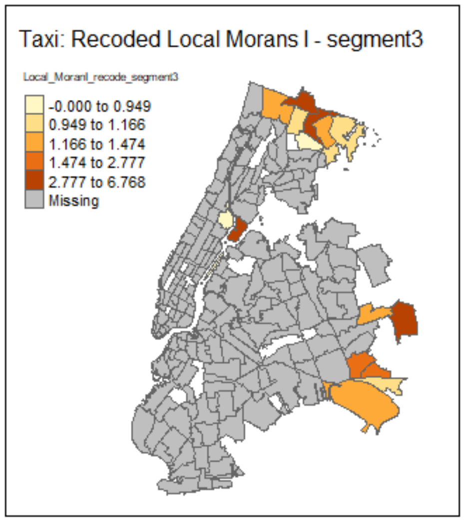
Figure B. 5 Recoded Local Moran's I and P-Value on ARIMA/GARCH/MLR Residuals – Full Length

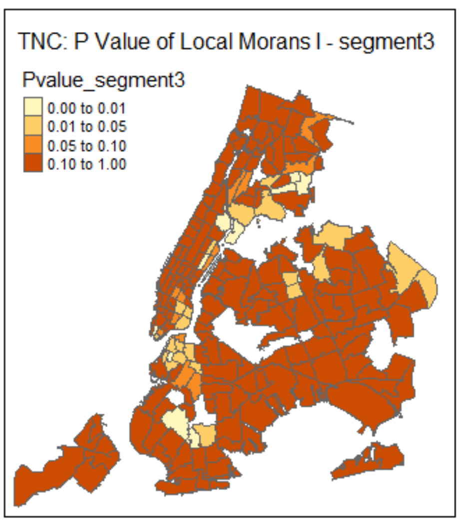

TNC: Recoded Local Morans I - segment1
Local_MoranI_recode_segment1
-0.000 to 0.849
0.849 to 1.253
1.253 to 1.613
1.613 to 2.900
2.900 to 5.432
Missing
Taxi: Recoded Local Morans I - segment1
Local_MoranI_recode_segment1
0.000 to 0.771
0.771 to 1.077
1.077 to 1.324
1.324 to 1.595
1.595 to 2.710
Missing
TNC: P Value of Local Morans I - segment1
Pvalue_segment1
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - segment1
Pvalue_segment1
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
TNC: Recoded Local Morans I - segment2
Local_MoranI_recode_segment2
0.000 to 1.046
1.046 to 1.471
1.471 to 1.952
1.952 to 2.655
2.655 to 3.944
Missing
Taxi: Recoded Local Morans I - segment2
Local_MoranI_recode_segment2
0.000 to 0.847
0.847 to 1.060
1.060 to 1.183
1.183 to 1.449
1.449 to 2.874
Missing

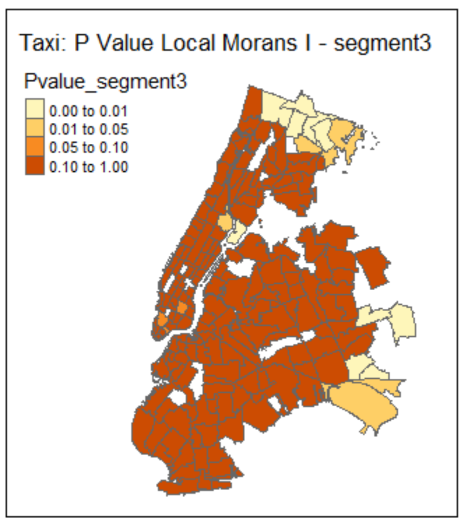

TNC: P Value of Local Morans I - segment2
Pvalue_segment2
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - segment2
Pvalue_segment2
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


TNC: Recoded Local Morans I - segment3
Local_MoranI_recode_segment3
0.000 to 0.770
0.770 to 0.830
0.830 to 0.927
0.927 to 1.258
1.258 to 6.499
Missing


Taxi: Recoded Local Morans I - segment3
Local_MoranI_recode_segment3
-0.000 to 0.949
0.949 to 1.166
1.166 to 1.474
1.474 to 2.777
2.777 to 6.768
Missing


TNC: P Value of Local Morans I - segment3
Pvalue_segment3
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - segment3
Pvalue_segment3
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

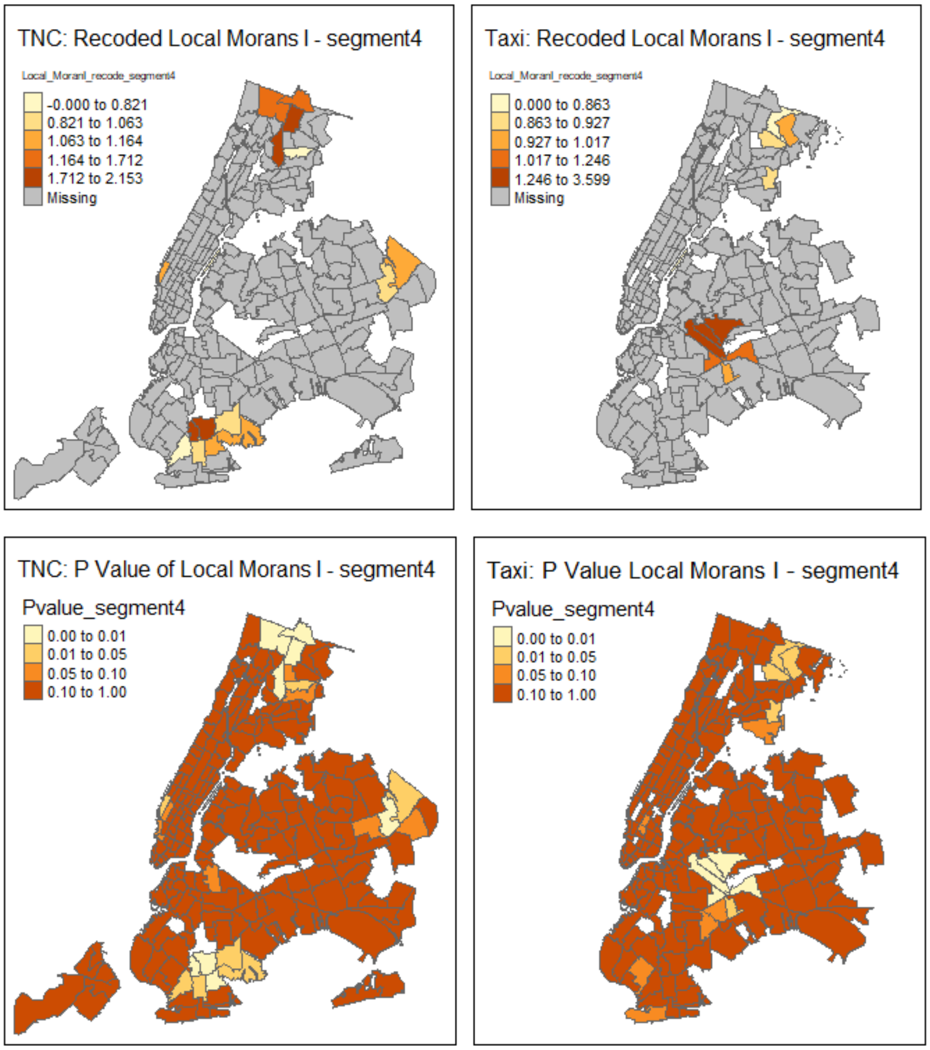
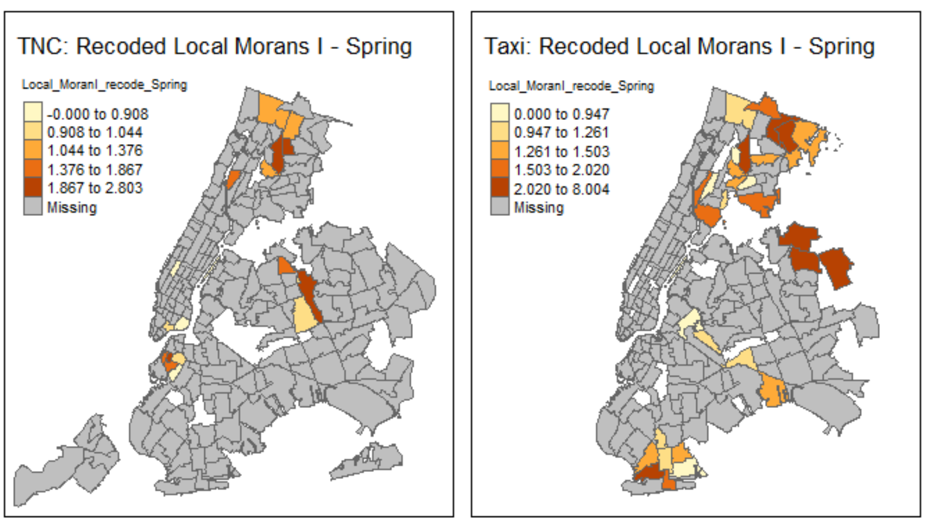
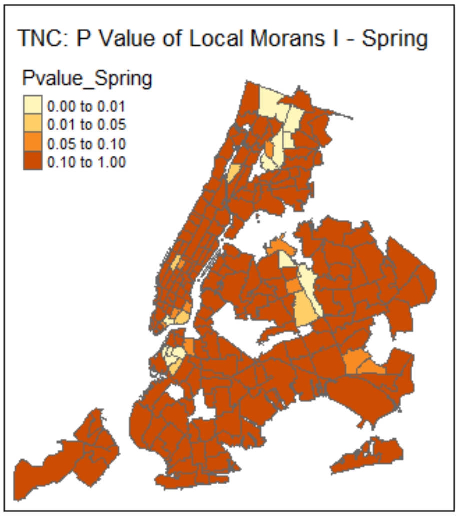
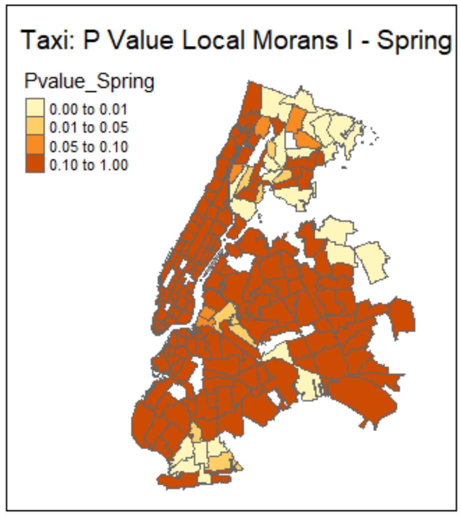
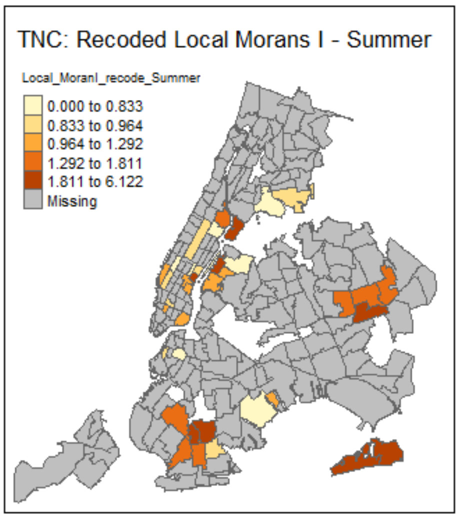


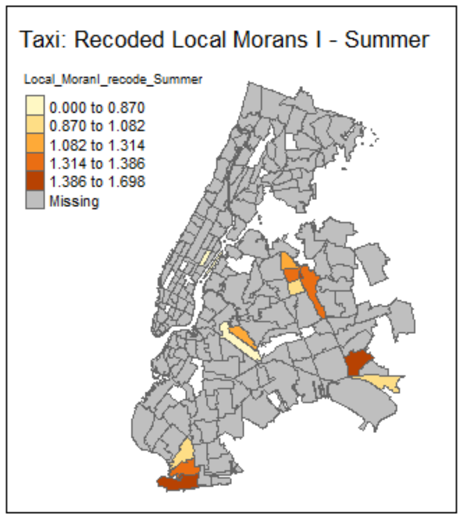
Figure B. 6 Recoded Local Moran's I and P-Value on ARIMA/GARCH/MLR Residuals – Segment

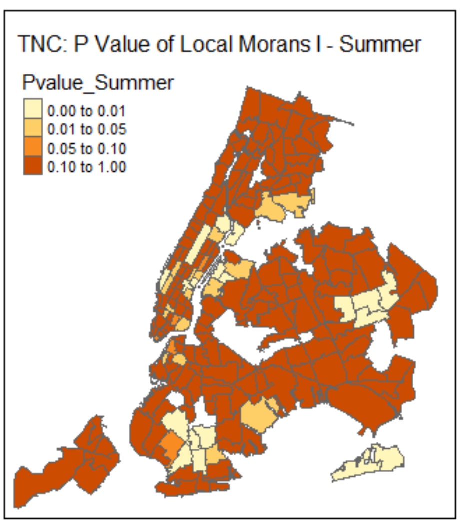


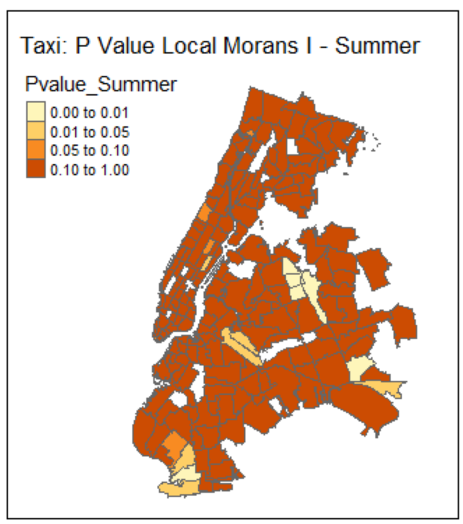

TNC: P Value of Local Morans I - Spring
Pvalue_Spring
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

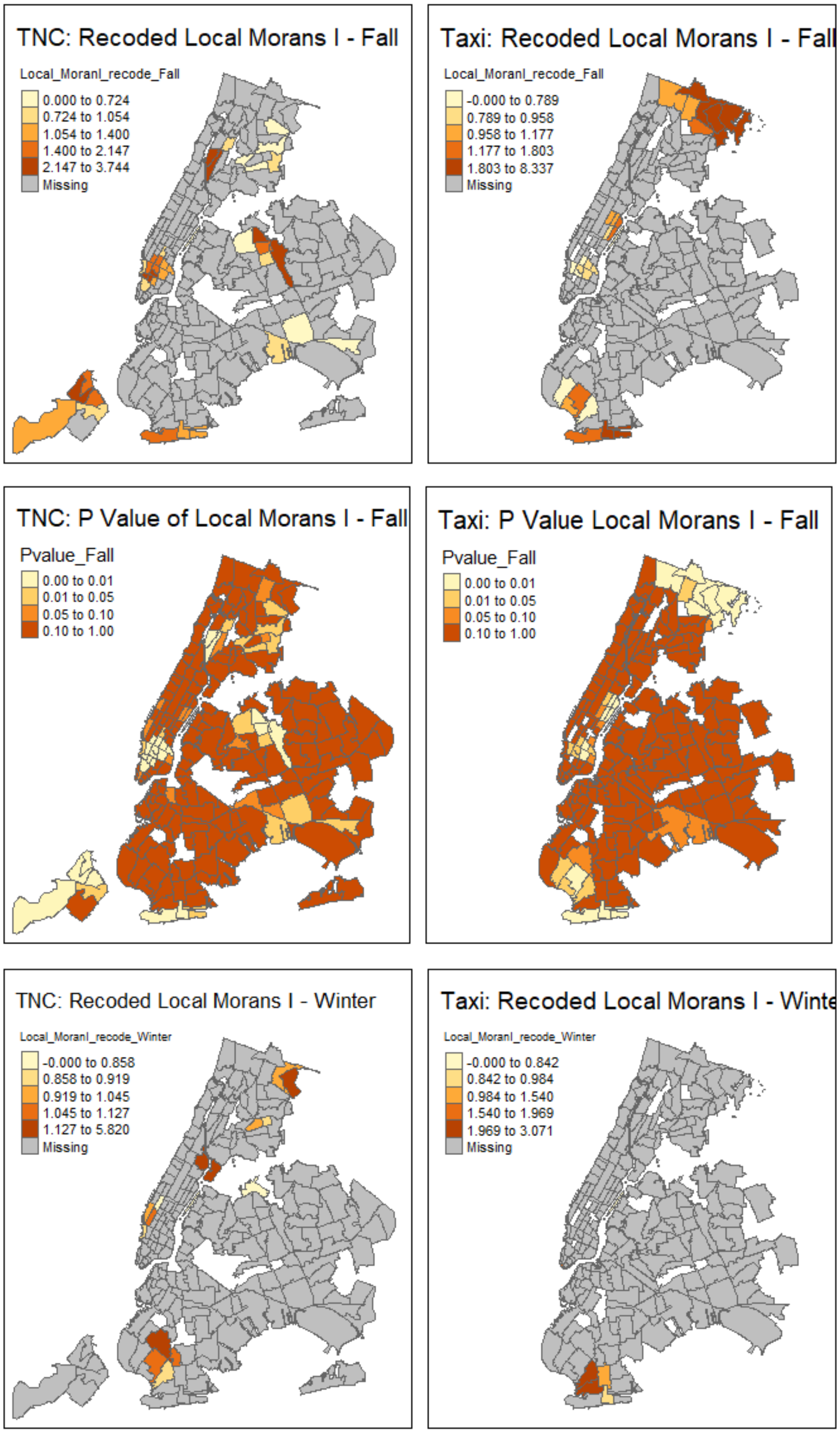

Taxi: P Value Local Morans I - Spring
Pvalue_Spring
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


TNC: Recoded Local Morans I - Summer
Local_MoranI_recode_Summer
0.000 to 0.833
0.833 to 0.964
0.964 to 1.292
1.292 to 1.811
1.811 to 6.122
Missing


Taxi: Recoded Local Morans I - Summer
Local_MoranI_recode_Summer
0.000 to 0.870
0.870 to 1.082
1.082 to 1.314
1.314 to 1.386
1.386 to 1.698
Missing


TNC: P Value of Local Morans I - Summer
Pvalue_Summer
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - Summer
Pvalue_Summer
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


TNC: Recoded Local Morans I - Fall
Local_MoranI_recode_Fall
0.000 to 0.724
0.724 to 1.054
1.054 to 1.400
1.400 to 2.147
2.147 to 3.744
Missing
Taxi: Recoded Local Morans I - Fall
Local_MoranI_recode_Fall
-0.000 to 0.789
0.789 to 0.958
0.958 to 1.177
1.177 to 1.803
1.803 to 8.337
Missing
TNC: P Value of Local Morans I - Fall
Pvalue_Fall
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - Fall
Pvalue_Fall
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
TNC: Recoded Local Morans I - Winter
Local_MoranI_recode_Winter
-0.000 to 0.858
0.858 to 0.919
0.919 to 1.045
1.045 to 1.127
1.127 to 5.820
Missing
Taxi: Recoded Local Morans I - Winte
Local_MoranI_recode_Winter
-0.000 to 0.842
0.842 to 0.984
0.984 to 1.540
1.540 to 1.969
1.969 to 3.071
Missing

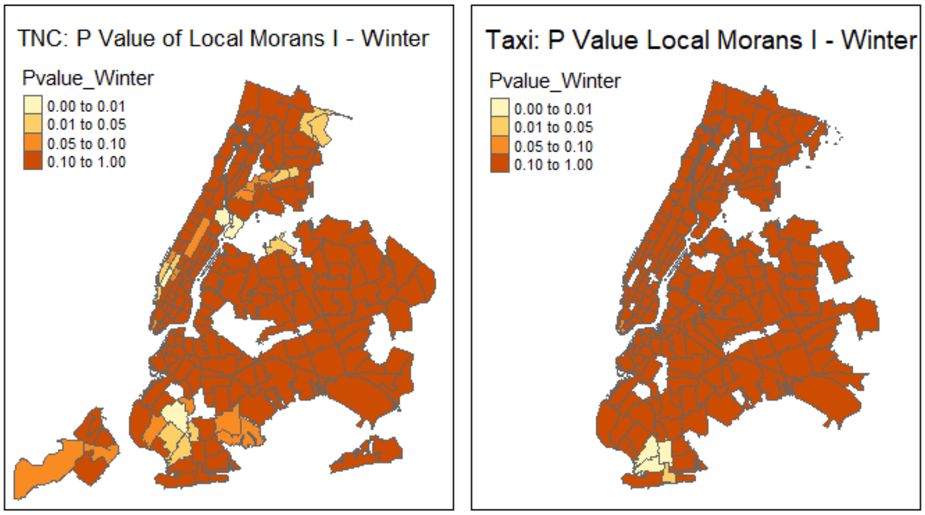
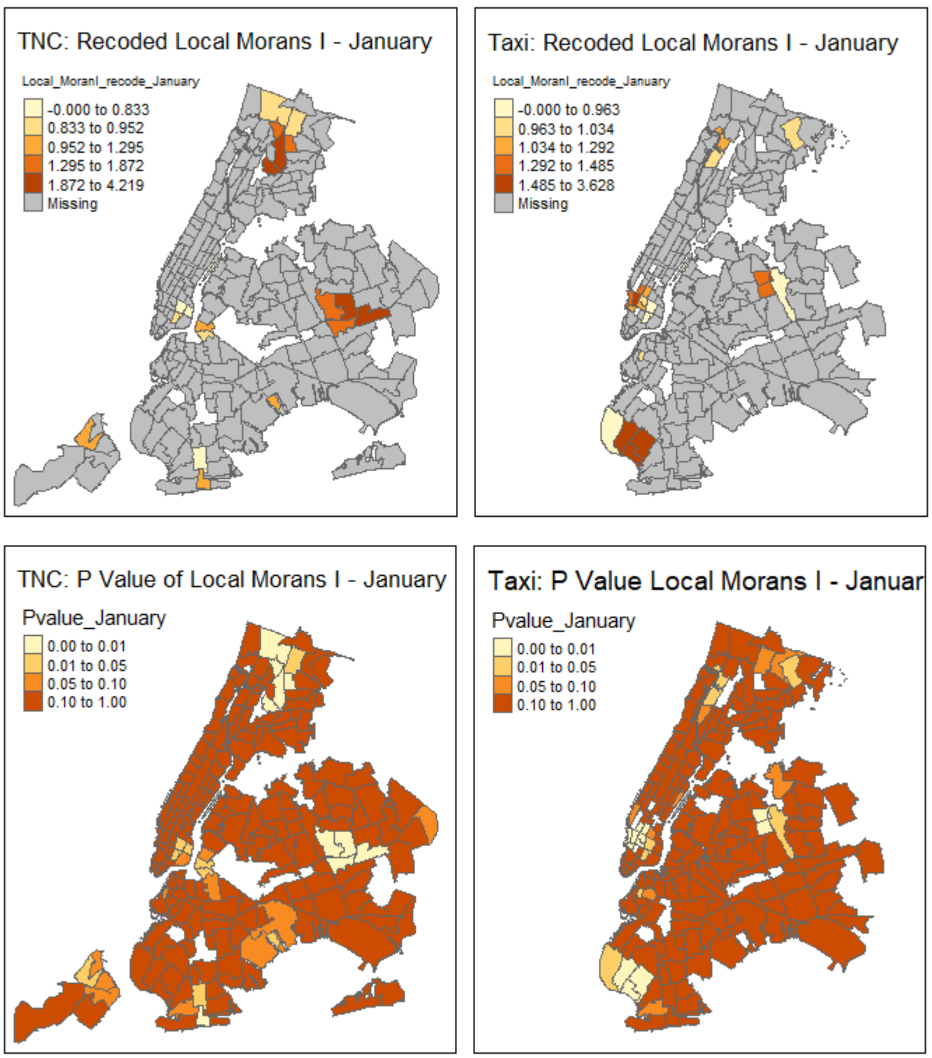
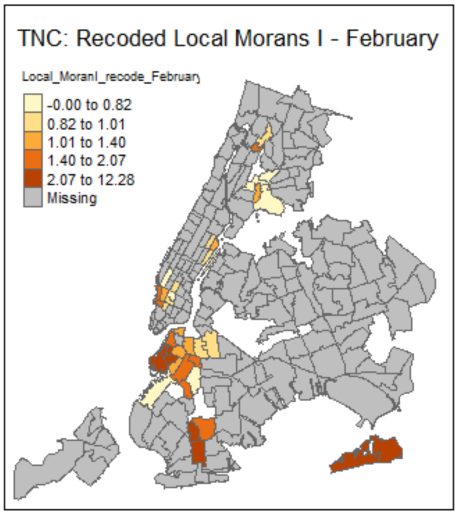
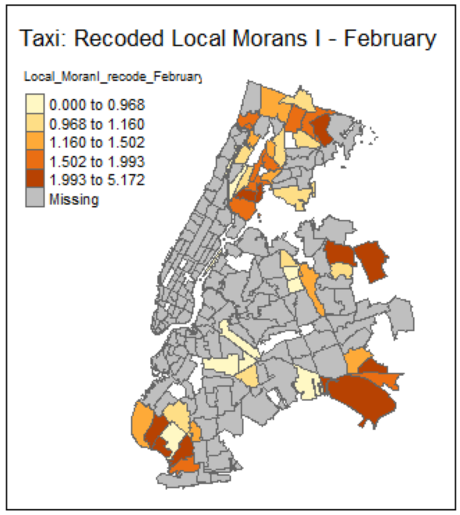
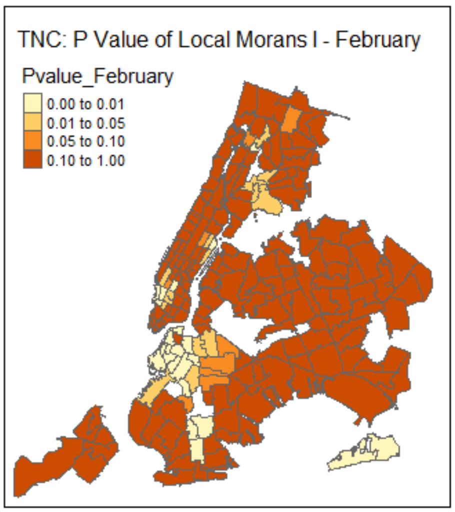


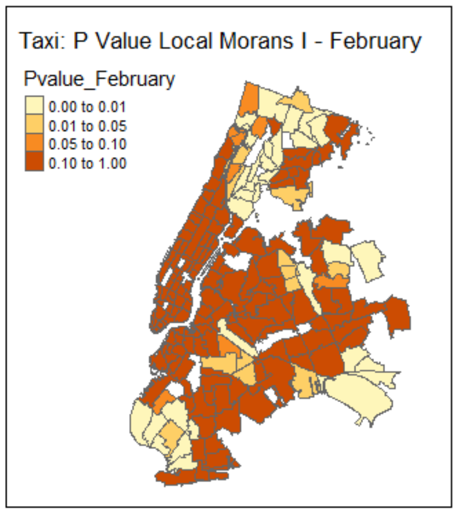
Figure B. 7 Recoded Local Moran's I and P-Value on ARIMA/GARCH/MLR Residuals – Season

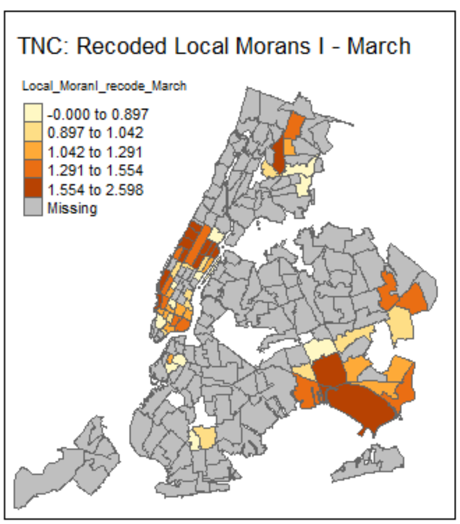


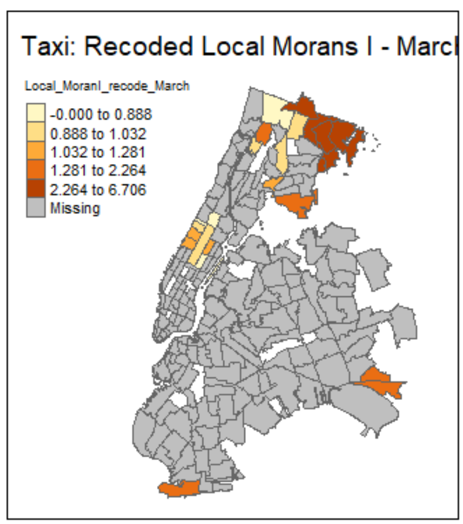

TNC: Recoded Local Morans I - February
Local_MoranI_recode_February
-0.00 to 0.82
0.82 to 1.01
1.01 to 1.40
1.40 to 2.07
2.07 to 12.28
Missing

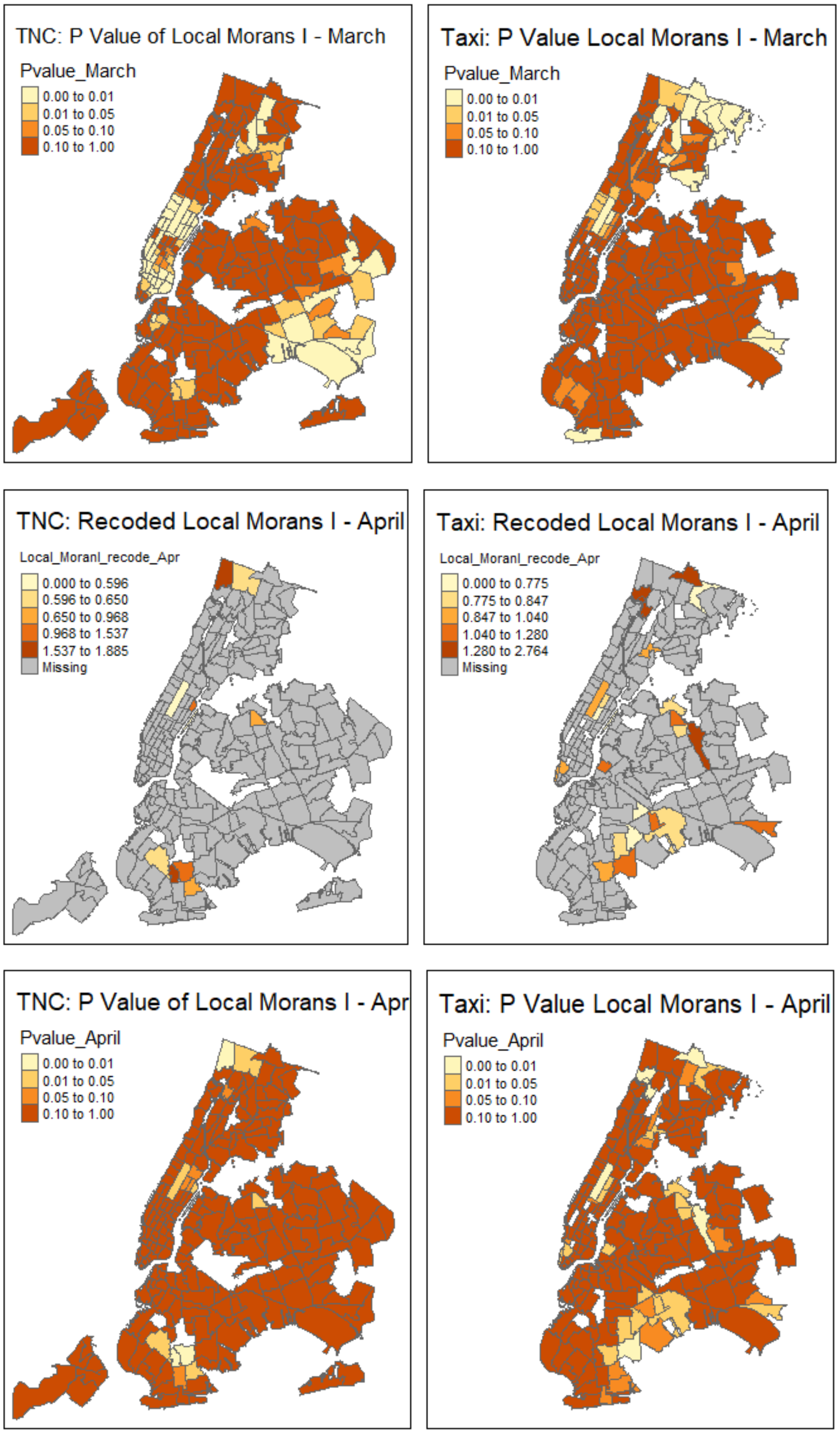

Taxi: Recoded Local Morans I - February
Local_MoranI_recode_February
0.000 to 0.968
0.968 to 1.160
1.160 to 1.502
1.502 to 1.993
1.993 to 5.172
Missing

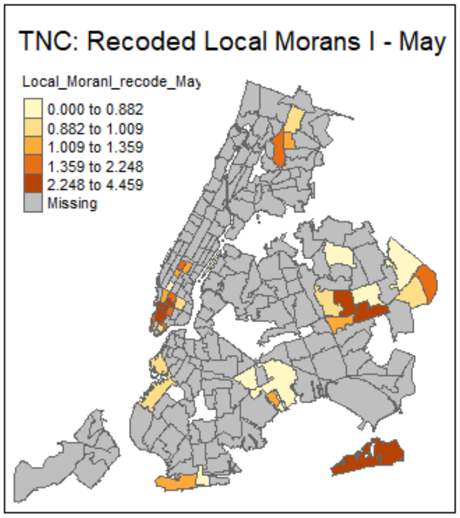

TNC: P Value of Local Morans I - February
Pvalue_February
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

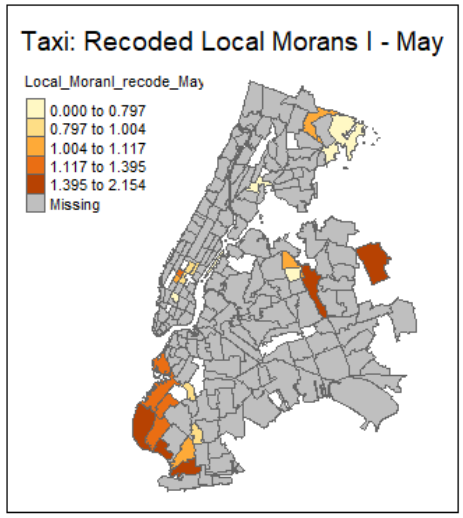

Taxi: P Value Local Morans I - February
Pvalue_February
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

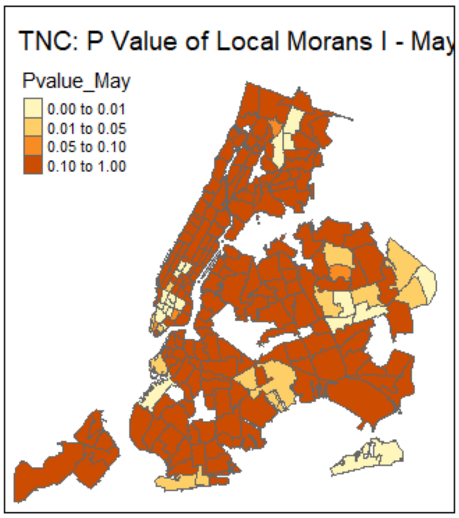

TNC: Recoded Local Morans I - March
Local_MoranI_recode_March
-0.000 to 0.897
0.897 to 1.042
1.042 to 1.291
1.291 to 1.554
1.554 to 2.598
Missing

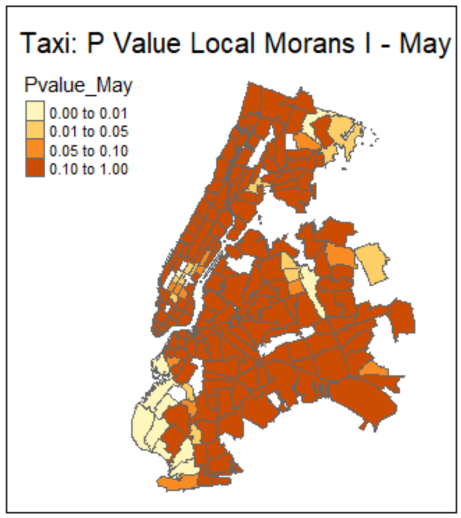

Taxi: Recoded Local Morans I - Marc
Local_MoranI_recode_March
-0.000 to 0.888
0.888 to 1.032
1.032 to 1.281
1.281 to 2.264
2.264 to 6.706
Missing

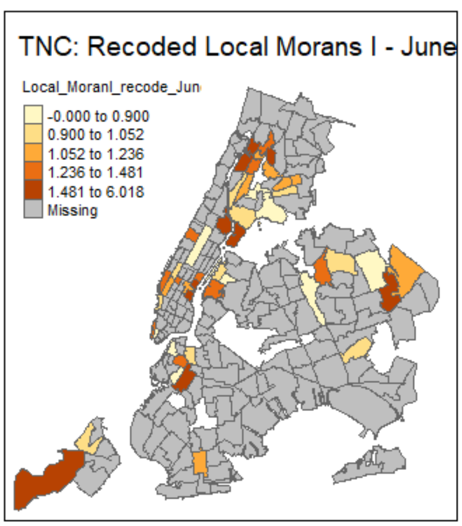

TNC: P Value of Local Morans I - March
Pvalue_March
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - March
Pvalue_March
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
TNC: Recoded Local Morans I - April
Local_MoranI_recode_Apr
0.000 to 0.596
0.596 to 0.650
0.650 to 0.968
0.968 to 1.537
1.537 to 1.885
Missing
Taxi: Recoded Local Morans I - April
Local_MoranI_recode_Apr
0.000 to 0.775
0.775 to 0.847
0.847 to 1.040
1.040 to 1.280
1.280 to 2.764
Missing
TNC: P Value of Local Morans I - Apr
Pvalue_April
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - April
Pvalue_April
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

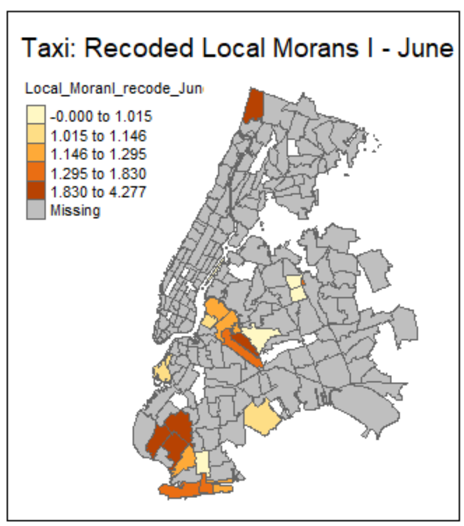

TNC: Recoded Local Morans I - May
Local_MoranI_recode_May
0.000 to 0.882
0.882 to 1.009
1.009 to 1.359
1.359 to 2.248
2.248 to 4.459
Missing

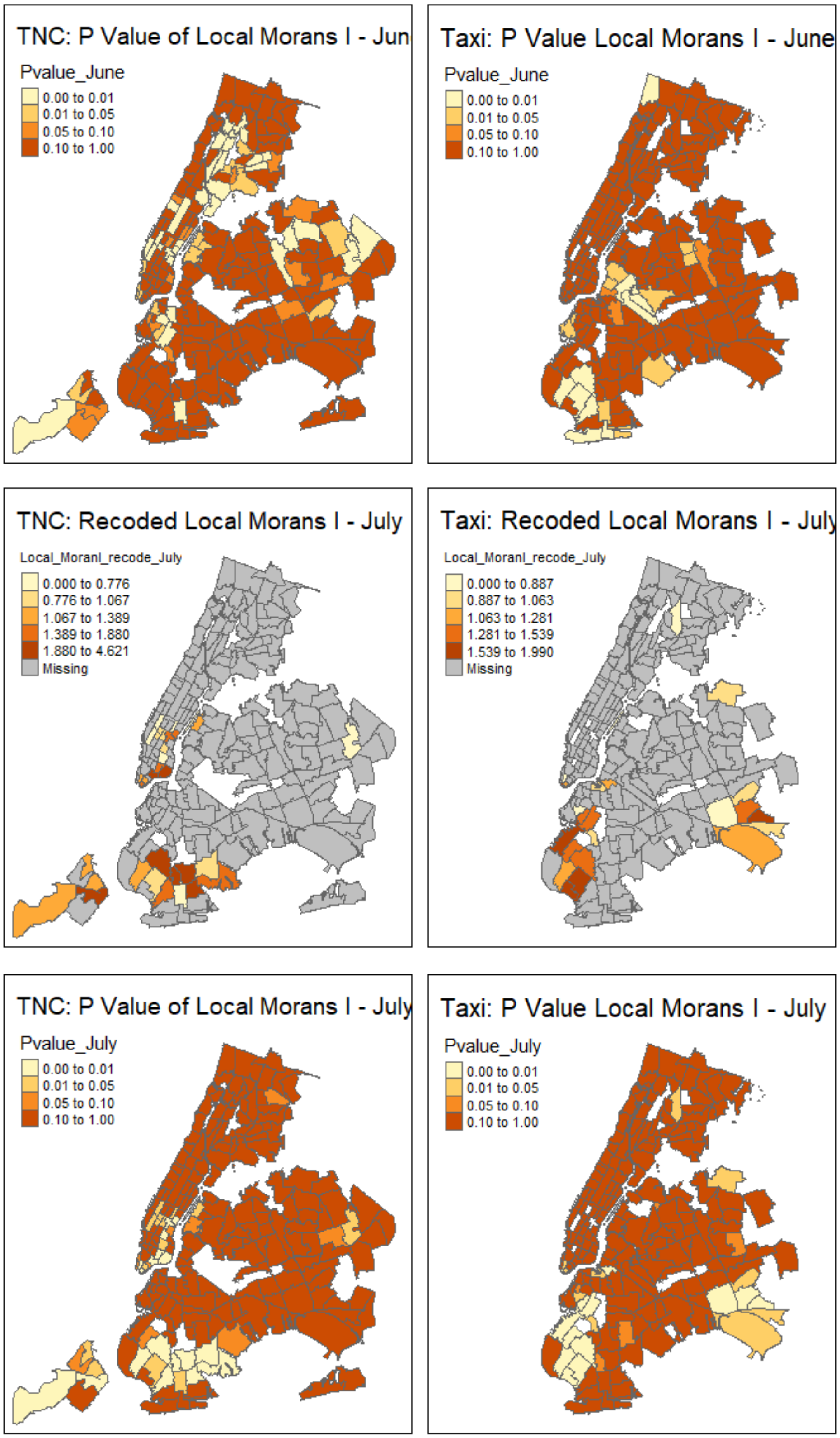

Taxi: Recoded Local Morans I - May
Local_MoranI_recode_May
0.000 to 0.797
0.797 to 1.004
1.004 to 1.117
1.117 to 1.395
1.395 to 2.154
Missing

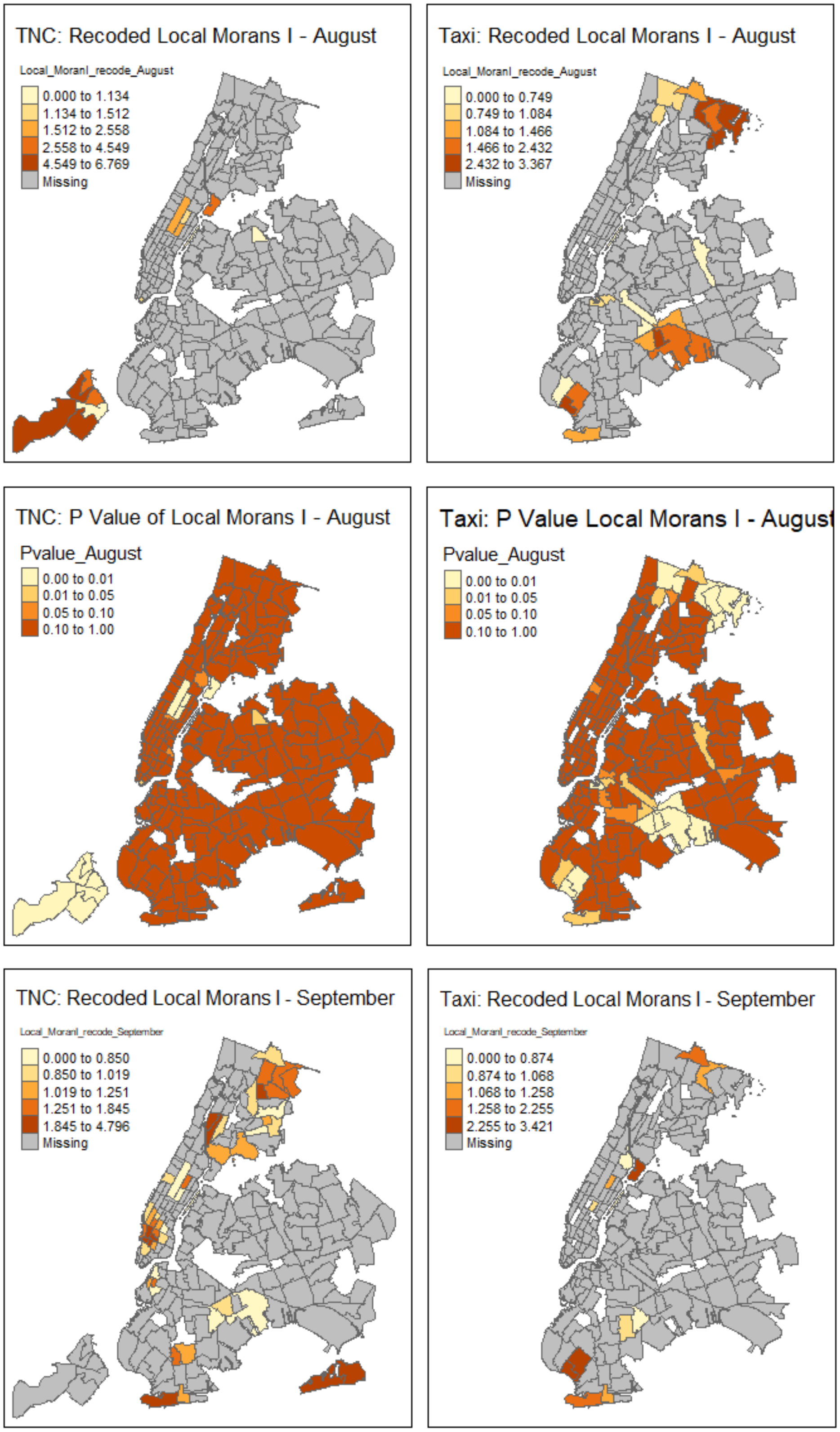

TNC: P Value of Local Morans I - May
Pvalue_May
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

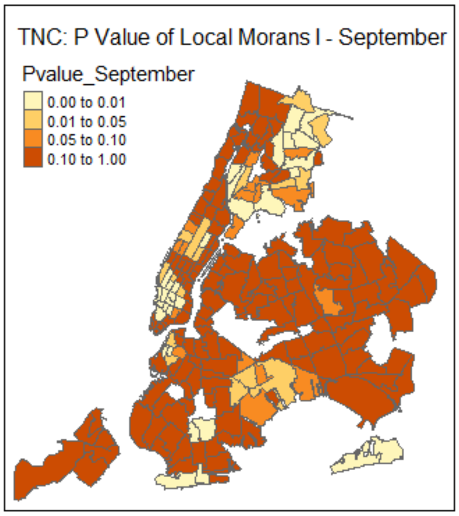

Taxi: P Value Local Morans I - May
Pvalue_May
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

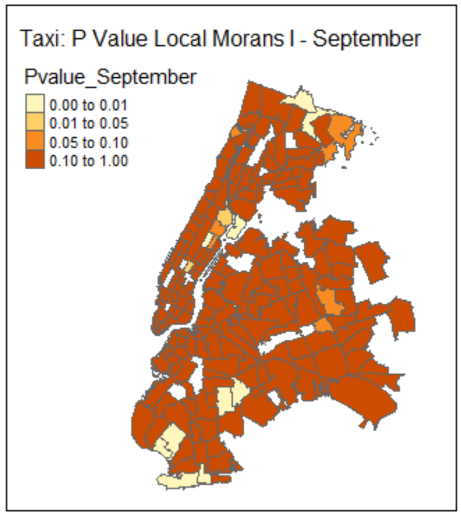

TNC: Recoded Local Morans I - June
Local_MoranI_recode_Jun
-0.000 to 0.900
0.900 to 1.052
1.052 to 1.236
1.236 to 1.481
1.481 to 6.018
Missing

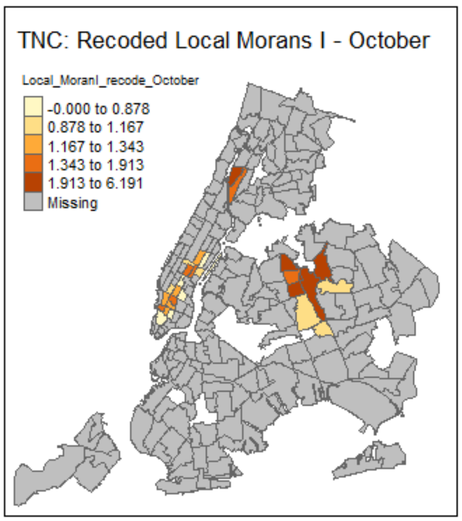

Taxi: Recoded Local Morans I - June
Local_MoranI_recode_Jun
-0.000 to 1.015
1.015 to 1.146
1.146 to 1.295
1.295 to 1.830
1.830 to 4.277
Missing

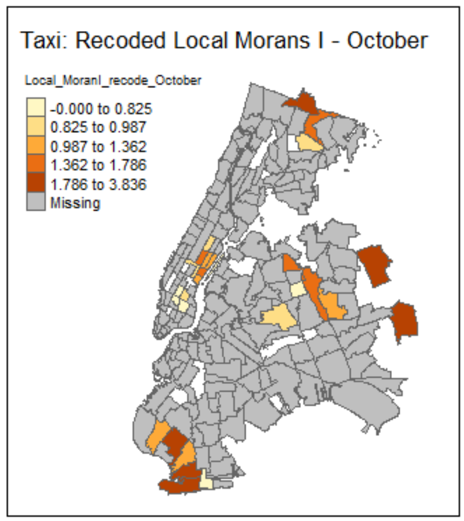

TNC: P Value of Local Morans I - Jun
Pvalue_June
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - June
Pvalue_June
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
TNC: Recoded Local Morans I - July
Local_MoranI_recode_July
0.000 to 0.776
0.776 to 1.067
1.067 to 1.389
1.389 to 1.880
1.880 to 4.621
Missing
Taxi: Recoded Local Morans I - July
Local_MoranI_recode_July
0.000 to 0.887
0.887 to 1.063
1.063 to 1.281
1.281 to 1.539
1.539 to 1.990
Missing
TNC: P Value of Local Morans I - July
Pvalue_July
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - July
Pvalue_July
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

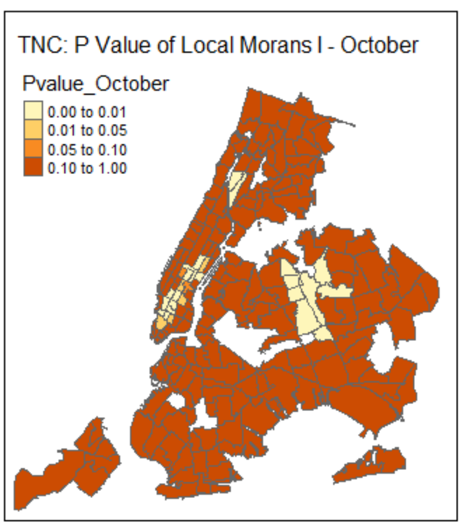

TNC: Recoded Local Morans I - August
Local_MoranI_recode_August
0.000 to 1.134
1.134 to 1.512
1.512 to 2.558
2.558 to 4.549
4.549 to 6.769
Missing
Taxi: Recoded Local Morans I - August
Local_MoranI_recode_August
0.000 to 0.749
0.749 to 1.084
1.084 to 1.466
1.466 to 2.432
2.432 to 3.367
Missing
TNC: P Value of Local Morans I - August
Pvalue_August
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
Taxi: P Value Local Morans I - August
Pvalue_August
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00
TNC: Recoded Local Morans I - September
Local_MoranI_recode_September
0.000 to 0.850
0.850 to 1.019
1.019 to 1.251
1.251 to 1.845
1.845 to 4.796
Missing
Taxi: Recoded Local Morans I - September
Local_MoranI_recode_September
0.000 to 0.874
0.874 to 1.068
1.068 to 1.258
1.258 to 2.255
2.255 to 3.421
Missing

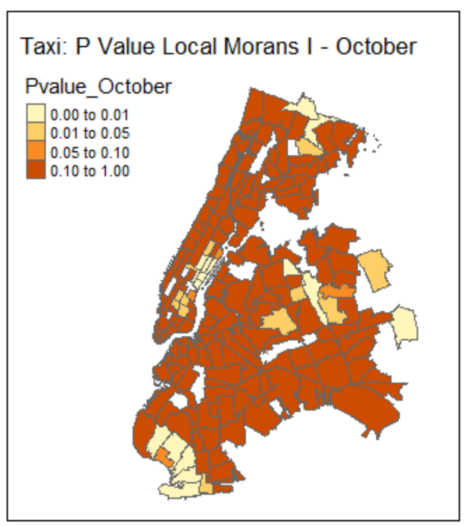

TNC: P Value of Local Morans I - September
Pvalue_September
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

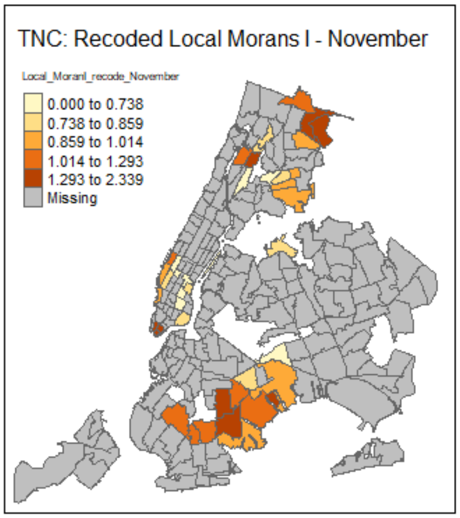

Taxi: P Value Local Morans I - September
Pvalue_September
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

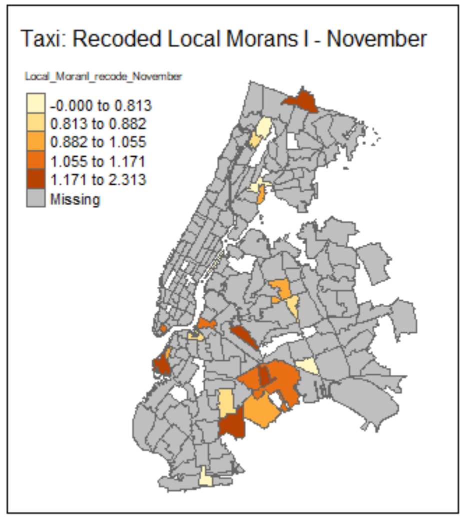

TNC: Recoded Local Morans I - October
Local_MoranI_recode_October
-0.000 to 0.878
0.878 to 1.167
1.167 to 1.343
1.343 to 1.913
1.913 to 6.191
Missing

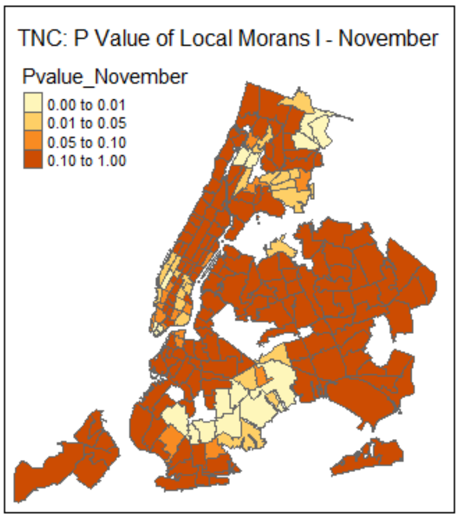

Taxi: Recoded Local Morans I - October
Local_MoranI_recode_October
-0.000 to 0.825
0.825 to 0.987
0.987 to 1.362
1.362 to 1.786
1.786 to 3.836
Missing

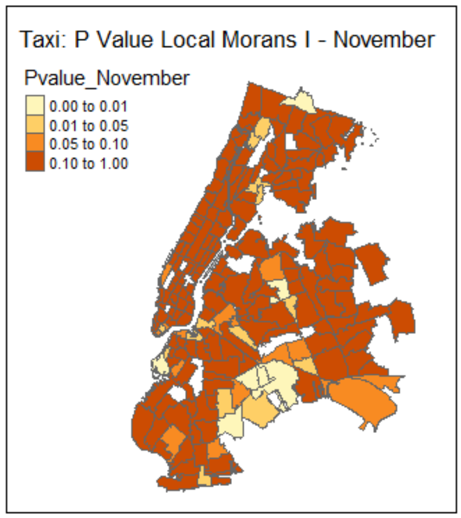

TNC: P Value of Local Morans I - October
Pvalue_October
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

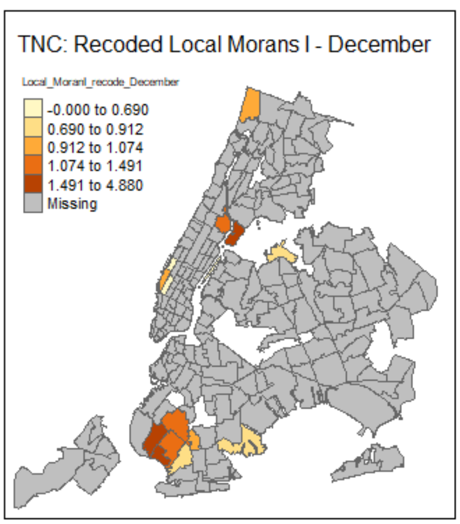

Taxi: P Value Local Morans I - October
Pvalue_October
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00

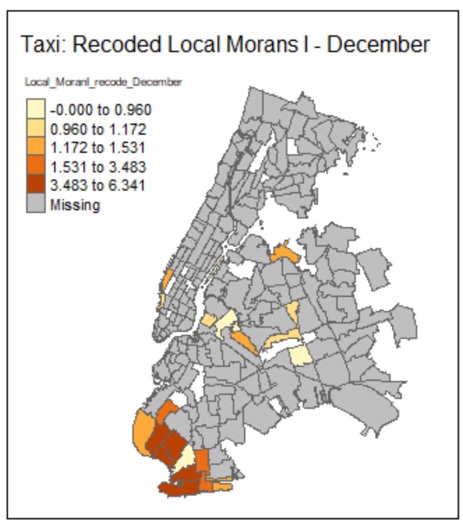

TNC: Recoded Local Morans I - November
Local_MoranI_recode_November
0.000 to 0.738
0.738 to 0.859
0.859 to 1.014
1.014 to 1.293
1.293 to 2.339
Missing

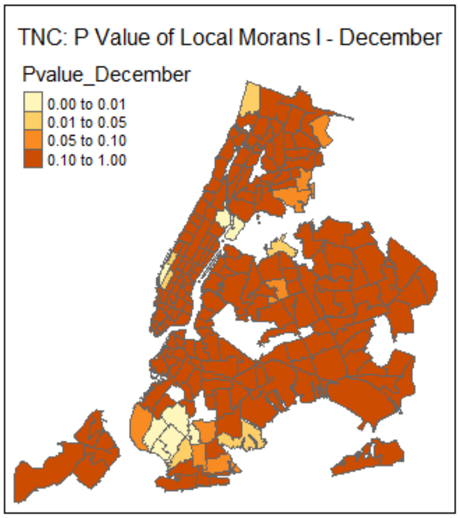

Taxi: Recoded Local Morans I - November
Local_MoranI_recode_November
-0.000 to 0.813
0.813 to 0.882
0.882 to 1.055
1.055 to 1.171
1.171 to 2.313
Missing


TNC: P Value of Local Morans I - November
Pvalue_November
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


Taxi: P Value Local Morans I - November
Pvalue_November
0.00 to 0.01
0.01 to 0.05
0.05 to 0.10
0.10 to 1.00


TNC: Recoded Local Morans I - December
Local_MoranI_recode_December
-0.000 to 0.690
0.690 to 0.912
0.912 to 1.074
1.074 to 1.491
1.491 to 4.880
Missing


Taxi: Recoded Local Morans I - December
Local_MoranI_recode_December
-0.000 to 0.960
0.960 to 1.172
1.172 to 1.531
1.531 to 3.483
3.483 to 6.341
Missing



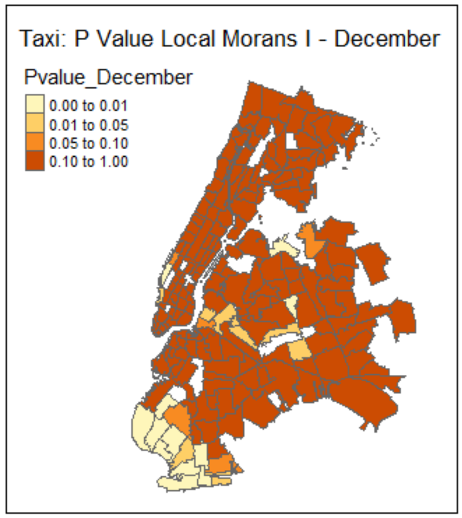


Figure B. 8 Recoded Local Moran's I and P-Value on ARIMA/GARCH/MLR Residuals – Month