America's Warzone: Modeling Armed Robberies in Chicago

Reuben K. McCreanor[*]
reuben.mccreanor@duke.edu

Anna K. Yanchenko[*]
anna.yanchenko@duke.edu

Lei Qian[*]
lei.qian@duke.edu

Megan S. Robertson[*]
megan.robertson@duke.edu

1 Introduction

The City of Chicago is frequently listed as one of the most dangerous and crime-ridden cities in the US. President Donald Trump frequently discusses the high-rate of crime in Chicago. According to the Chicago Tribune, there were 4,367 shooting victims in Chicago in 2016. In the same year there were also 785 homicides.[10] However, other reports conclude that Chicago should not be called the "crime capital" of America, as Chicago's violence rate is lower than cities like St. Louis and Detroit. [9] The goal of this project was to examine crime in Chicago, specifically armed robberies, from 2012-2016.

2 Data

The crime data used for this project came from the City of Chicago's website. [5]The data contained every reported crime in Chicago from 2001 to the present (Table 1). In addition to the type of crime reported (battery, assault, etc.), there was information on the location and time of the crime. The data set was reduced to only consider armed robberies, which we define to be robberies committed with a gun or another type of firearm.

Case Number	Date	Block	Description	Beat	District	Ward	Community Area	Location
HN180091	02/16/2007 03:20:00 PM	012XX W 103RD ST	OTHER WEAPONS VIOLATION	2232	22	21	73	(41.706819022, -87.654048084)
HN184333	02/15/2007 07:00:00 PM	093XX S WOODLAWN AVE	TELEPHONE THREAT	413	4	8	47	(41.725252492, -87.594860893)
HN182527	02/17/2007 09:00:00 PM	042XX S COTTAGE GROVE AVE	OTHER VIOLATION	213	2	4	38	(41.81741558, -87.606719823)
HN183814	02/18/2007 10:06:37 PM	063XX N SHERIDAN RD	TO PROPERTY	2433	24	49	77	(41.996866019, -87.655592844)
HN182579	02/18/2007 12:20:00 AM	033XX W HURON ST	DOMESTIC BATTERY SIMPLE	1121	11	27	23	(41.893682761, -87.710701702)
HN182986	02/18/2007 10:20:00 AM	042XX N CENTRAL AVE	OVER $500	1624	16	38	15	(41.957385814, -87.767141739)
HN183716	02/18/2007 08:40:00 PM	037XX S MICHIGAN AVE	DOMESTIC BATTERY SIMPLE	211	2	3	35	(41.827167256, -87.623160687)
HN184010	02/19/2007 02:40:00 AM	010XX N LAWNDALE AVE	DOMESTIC BATTERY SIMPLE	1112	11	27	23	(41.89998654, -87.71890157)
HN181071	02/17/2007 02:21:00 AM	051XX S CALUMET AVE	POSS FIREARM/AMMO:NO FOID CARD	232	2	3	40	(41.801609105, -87.617736187)
HN182079	02/17/2007 04:00:00 PM	034XX W FLOURNOY ST	DOMESTIC BATTERY SIMPLE	1133	11	24	27	(41.872709648, -87.711929688)
HN184350	02/19/2007 12:00:00 AM	013XX N WOLCOTT AVE	FROM BUILDING	1424	14	1	24	(41.905889056, -87.674299261)
HN184306	02/15/2007 12:01:00 AM	001XX N PARKSIDE AVE	ILLEGAL USE CASH CARD	1512	15	29	25	(41.88301997, -87.766605275)
HN183370	02/09/2007 10:00:00 AM	024XX W DEVON AVE	SIMPLE	2413	24	50	2	(41.99771689, -87.690448237)
HN183500	02/18/2007 05:10:00 PM	070XX S THROOP ST	DOMESTIC BATTERY SIMPLE	734	7	17	67	(41.7658997, -87.65656296)
HN184055	02/19/2007 05:30:00 AM	005XX W ROSCOE ST	TO RESIDENCE	2331	19	44	6	(41.943338611, -87.64332075)
HN184352	02/19/2007 10:45:00 AM	003XX N MICHIGAN AVE	FROM BUILDING	122	1	42	32	(41.887845852, -87.624560336)
HN181018	02/17/2007 01:31:00 AM	056XX S WABASH AVE	TO LAND	233	2	20	40	(41.792323044, -87.624025834)
HN182857	02/18/2007 07:39:41 AM	014XX W LUNT AVE	FORCIBLE ENTRY	2431	24	49	1	(42.009107852, -87.666843608)
HN177373	02/15/2007 08:44:00 AM	054XX S CORNELL AVE	TO VEHICLE	2132	2	5	41	(41.796263314, -87.585435453)
HN183280	01/14/2007 12:00:00 PM	021XX W BIRCHWOOD AVE	DOMESTIC BATTERY SIMPLE	2424	24	49	1	(42.017948603, -87.683418074)

Table 1: The City of Chicago website provides a data set containing information on crimes committed in the city from 2001 to present day.

---

[*]Department of Statistical Science, Duke University

3 Time Series Analysis

The City of Chicago is divided into regions known as sides (Figure 1), where each side is comprised of several neighborhoods. There is a lot of variation in the population (Figure 2) and the number of armed robberies per capita (Figure 3) for these sides. Additionally, some sides are more residential, while others are more commercial.
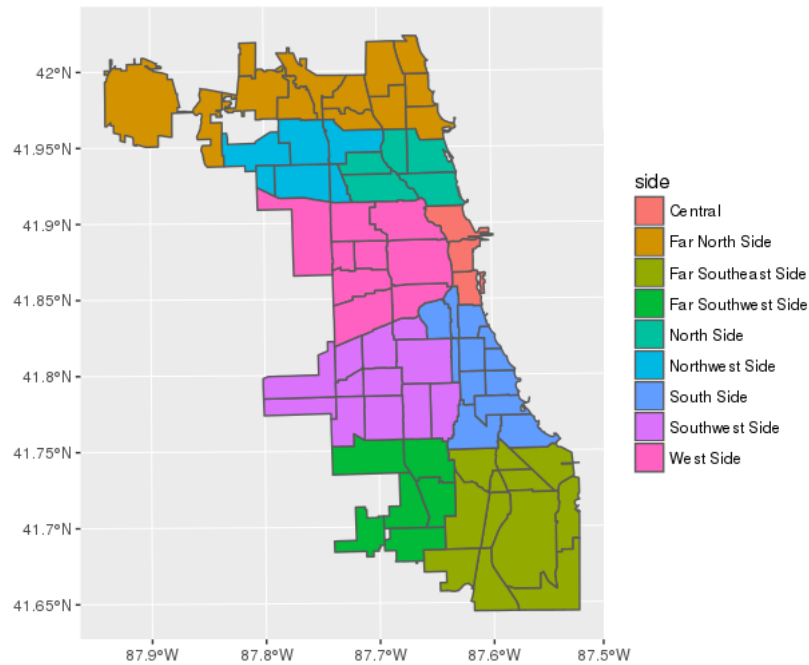

Figure 1: The "sides" of Chicago. The borders correspond to the boundaries of the community areas colored by the side.

ARIMA models were fitted to predict the counts of monthly armed robberies in each side of the city between 2003 and 2016. In order to determine the type of model, the ACF and PACF plots were examined for the data for each of the sides. For example, if there was structure in the PACF plot beyond one lag, moving average terms were added. The model residuals were also examined to ensure that there was no remaining structure in the residuals. The PACF and ACF plots for the data from the South Side are displayed below. The ACF plot showed evidence of seasonality at lag 12 (i.e. yearly trends). After lag 1, there was no large values for the PACF, so no moving average terms were included in the model.
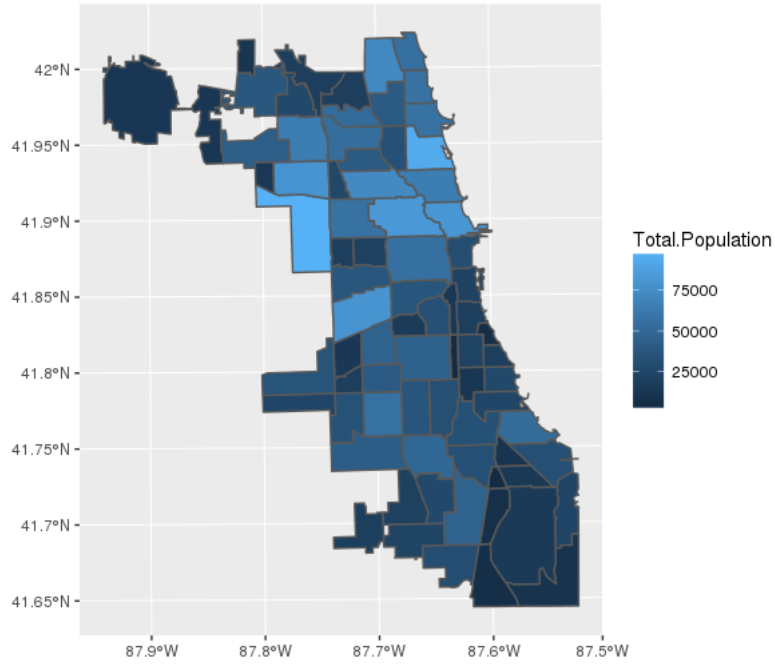
Figure 2: The population distribution of the "sides" of Chicago. The borders correspond to the boundaries of the community areas.
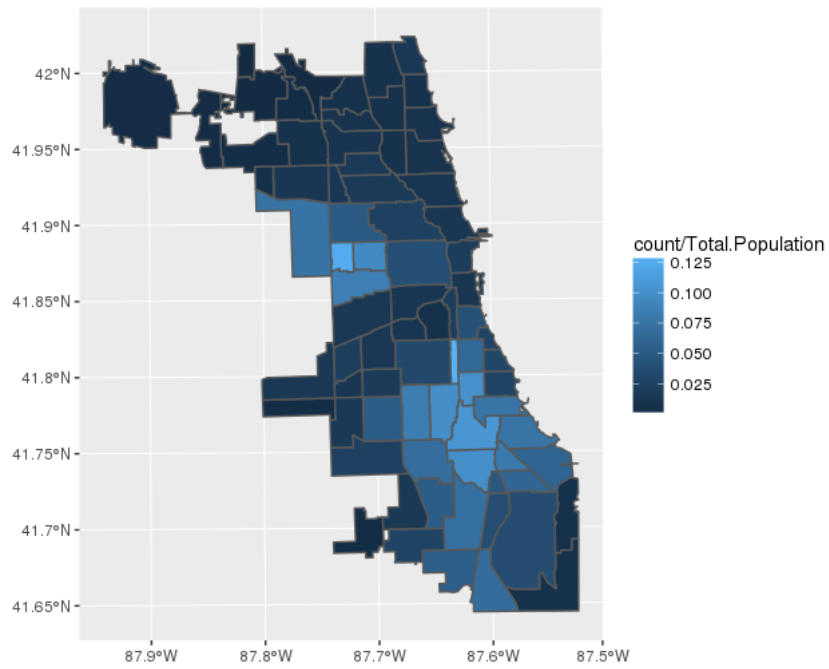

Figure 3: The number of armed robberies per capita for the "sides" of Chicago between 2003 and 2016. The borders correspond to the boundaries of the community areas.
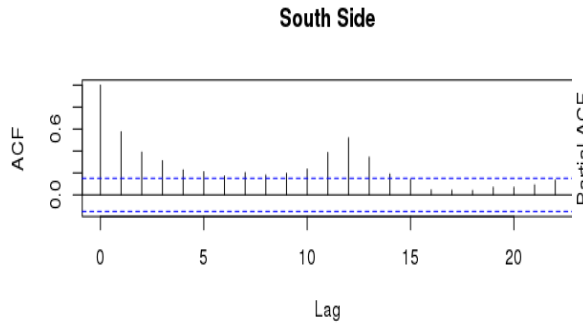
South Side   **South Side**

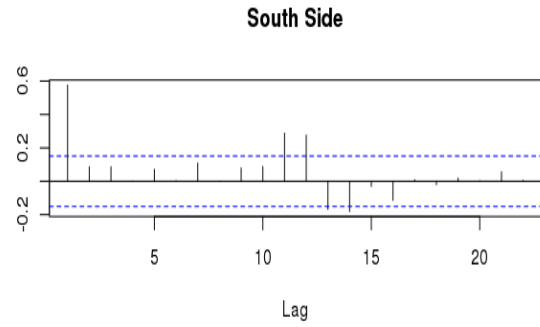Figure 4: ACF plot for the number of monthly armed robberies in the South Side of Chicago between 2003 and 2016.

Figure 5: PACF plot for the number of monthly armed robberies in the South Side of Chicago between 2003 and 2016.

Based on the ACF Figure 6 and PACF Figure 3 plots, an AR(4) model was fit with a seasonal component with period twelve. The residuals plot for this model did not display any remaining structure in the data and the coefficients are in Table 3.

Figure 6: ACF plot of the residuals of an AR(4) model with a period twelve seasonal component fit to the monthly count of armed robberies for the South Side.
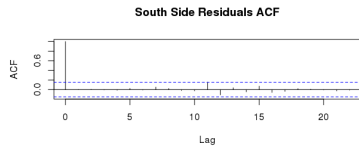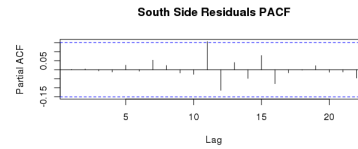
Figure 7: PACF plot of the residuals of an AR(4) model with a period twelve seasonal component fit to the monthly count of armed robberies for the South Side.

	ar1	ar2	ar3	ar4	sar1
Coefficient	0.3850	0.1328	0.0825	0.0167	0.4784
Standard Error	0.0793	0.0832	0.0824	0.0771	0.0708

Table 2: Summary of model fit for the AR(4) with period 12 seasonal component fit to the monthly count of armed robberies for the South Side.

The coefficient estimates for all of the different sides were very similar. While some sides displayed evidence of higher order autoregressive structure or the addition of moving average terms, as compared to the South Side, all sides had a clear period 12 seasonal component, indicating strong yearly trends for all sides of the city. The coefficient estimates were positive for the autoregressive terms, indicating that there was a positive correlation between the amount of monthly armed robberies over time. Plots of the various model fits can be found in subsection 5.1. It is interesting that although the sides of Chicago are quite diverse in terms of population and demographics, as well as the number of monthly armed robberies, the temporal trends for all of the sides are very similar. Although the count of the monthly armed robberies differs by side of the city, the overall temporal trend is the same across Chicago and has a strong yearly, autoregressive trend.

4 Spatial Models

4.1 Introduction

The City of Chicago is comprised of 77 distinct community areas. We predict for counts of armed robbery in each area using a Bayesian Spatial Latent Gaussian Process Poisson Regression Model (BSLGPPR) and a 100 explanatory variables provided by Chicago gathered from a range of years within that of the armed robbery data. We do make the assumption that these variables have not changed much over the next/previous few years and remain applicable. The explanatory variables come from 8 different datasets, average electricity usage, average gas usage, counts of graffiti removal, public health statistics, socioeconomic indicators, vacant properties, and demographics data.

Since we have numerous explanatory variables that range from demographic information to counts of graffiti art, we want to narrow down the number of variables to improve the model's prediction accuracy and interpretability. A popular frequentist method is the penalized LASSO regression; however, it does not take into account spatial information. Our model, the BSLGPPR, will simulate a LASSO regression while also modeling spatial random effects.

As the response variable of our data are counts of armed robbery in community areas, we model our observations $\{y_i\}_1^N$, $N = 77$, with a Poisson distribution.

$$y_i \sim \text{Poisson}(\lambda_i) \tag{1}$$

We know that λ_i must be positive, which is why we let it equal to the exponential of $\mathbf{X}\beta$. β are the coefficients for the explanatory variables and the intercept, which comprises the design matrix $\mathbf{X}$.

$$\log(\boldsymbol{\lambda}) = \mathbf{X}\beta \tag{2}$$

Equation 2 is for the LASSO regression, but we will add another variable ω_i to model random spatial effect in the BSLGPPR model.

$$\log(\boldsymbol{\lambda}) = \mathbf{X}\beta + \boldsymbol{\omega} \tag{3}$$

4.2 The LASSO Model

Our first model comprises of a simple penalized LASSO regression. Given the positive value of Moran's I, 0.5118399, this model is highly unlikely to outperform one that takes into account the spatial nature of crime. However, we implement this model as a baseline in order to illustrate why the BSLGPPR model performs better.

In this model, without the random spatial effect, we do make the assumption that the observations are independent of one another.

From Equation 2, it is simple to realize that

$$\boldsymbol{\lambda} = e^{\beta'\mathbf{X}}$$

. Given a set of parameters β, and explanatory variables $\mathbf{X}$, we observe the counts of armed robbery $\mathbf{Y}$ with probability

$$p(y_1, ..., y_N | \mathbf{x}_1, ..., \mathbf{x}_N, \beta) = \prod_{i=1}^{N} \frac{e^{y_i \beta' \mathbf{x}_i} e^{-e^{\beta' \mathbf{x}_i}}}{y_i!} \tag{4}$$

Equation 4 can be obtained by plugging Equation 2 into the Poisson probability distribution.

We then want to use the maximum likelihood method to find a set of β that will maximize the likelihood, Equation 4 which is the same as maximizing the log-likelihood:

$$l(\beta | \mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{N} \left(y_i(\beta' \mathbf{x}_i) - e^{\beta' x_i} \right) \tag{5}$$

To implement the penalized LASSO regression, we instead optimize the penalized log-likelihood:

$$\min_{\beta} -\frac{1}{N} l(\beta | \mathbf{X}, \mathbf{Y}) + \lambda \frac{1}{2} \left((1-\alpha) \sum_{i=1}^{N} \beta_i^2 + \alpha \sum_{i=1}^{N} |\beta_i| \right) \tag{6}$$

and set $\alpha = 1$.

This method will allow us to obtain more accurate predictions than regular OLS and perform variable selection to prevent overfitting and for interpretability purposes.

4.3 The Bayesian Spatial Latent Gaussian Process Poisson Regression Model

Our second model uses a double exponential or Laplace prior to emulate the LASSO penalized regression model as the distribution sharply peaks at zero; concentrating the probability mass at zero. While this prior will not cause our coefficients to go to zero as in the case of a LASSO–instead behaving more like Ridge regression–we get around this by constructing 95 percent credible intervals around the coefficients and finding the ones that contain zero.

In our Bayesian model, we use the same data as in subsection 4.2 and supplement in spatial data in terms of spatial polygons for the 77 areas.

For our model, we take Equation 1 and Equation 3 and further specify by setting:

$$\beta_j \overset{iid}{\sim} \text{Laplace}(0, \eta)$$
$$\boldsymbol{\omega} \sim \text{MVN}(\mathbf{0}, \tau(D - \phi W))$$
$$\tau \sim \text{Gamma}(2, 2)$$
$$\phi \sim \text{Unif}(0, 0.99)$$
$$\eta \sim \text{Unif}(0.001, 10)$$

where $\{D : d_{\{jj\}} = $ total number of neighboring community areas for community area j$|j \in 1...77\}$

where $\{W : w_{\{jk\}} = $ whether community area j shares boundaries with community area k$|j, k \in 1...77\}$

Like before, we say that the armed robbery data can be modeled using a Poisson distribution with λ_i. We know that λ_i must be positive, which is why we let it equal to the exponential of $\mathbf{X}\beta + \omega_i$. β are the coefficients for each explanatory variable $\mathbf{X}$ and the intercept. We model the random spatial effect by putting a multivariate gaussian prior on ω_i and setting the mean to zero,we expect the average effect to be 0, and a correlation matrix $\tau(D - \phi W))$, we believe community areas to be affected by neighboring community areas and the total number of neighboring community areas, which is then scaled by τ and ϕ. We set τ to have a Gamma(2,2) prior because we believe τ is positive and also heavily concentrated between 0 and 4. ϕ has a Unif(0, 0.99) prior to allow for equal probability for any value within the specified range; as we are not biased towards any weight for W. The βs are iid Laplace to emulate a penalized LASSO regression. We set η to have a Unif(0.001, 10) distribution because we believe that is how concentrated the β's will be around the mean, 0; the smaller the η the more heavily concentrated at the mean.

4.4 Results

β coefficients for the two models can be seen in Table 13 and Table 14 along with 95% credible intervals for the BSLGPPR model. To summarize, the LASSO model selects 30 variables whereas BSLGPPR selects 15 variables as seen in Table 3 and Table 4. Predictions are also more accurate as seen in Figure 8; the residuals for the BSLGPPR model are minute compared to that of the LASSO model as the LASSO scale would be approximately 100 times wider than that of the BSLGPPR model. While it seems the predictions in comparison to that of other community areas are similar between the two models, it does not change that BSLGPPR is more accurate. Take a community area such as Austin, predicted to have the highest amount of armed robberies for both models, yet the predicted amount differs by approximately 1000 counts–which is very signficant in this context.

Variable	Lasso
Intercept	6.5489
Male: 65 and 66 years	-0.3789
Not Hispanic or Latino, Black or African American alone	0.2773
Assault (Homicide)	0.2175
Unemployment	-0.2116
Infant Mortality Rate	0.2030
Vacant Housing Units	0.1868
Female: 30 to 34 years	0.1732
Prenatal Care Beginning in First Trimester	-0.1557
Breast cancer in females	0.1407
N_Graffiti	0.1310
Hispanic or Latino	0.1115
Female: 22 to 24 years	0.1054
Low Birth Weight	0.0916
Female: 80 to 84 years	0.0863
Female: 18 and 19 years	0.0851
PERCENT AGED 16+ UNEMPLOYED	0.0847
Tuberculosis	0.0790
Female: 15 to 17 years	0.0785
Female: 5 to 9 years	0.0642
Crowded Housing	-0.0503
Firearm-related	0.0400
Female: 21 years	0.0381
Female: 85 years and over	0.0377
Stroke (Cerebrovascular Disease)	-0.0358
Prostate Cancer in Males	0.0227
Colorectal Cancer	0.0124
vacantLots	-0.0080
Not Hispanic or Latino, Asian alone	-0.0045
N_Aff_Housing	0.0006

Table 3: LASSO model's non-zero Beta Coefficients sorted from Most to Least Impact

2.5%	97.5%	Spatial	Variable
6.2136	6.2683	6.2427	Intercept
-0.8244	-0.3868	-0.5845	Unemployment
-0.7821	-0.1529	-0.5485	Male: 45 to 49 years
-0.9097	-0.1008	-0.4947	Female: 10 to 14 years
0.1699	0.5196	0.3433	PERCENT AGED 16+ UNEMPLOYED
-0.5641	-0.0527	-0.2933	Male: 65 and 66 years
0.1618	0.4959	0.2893	Cancer (All Sites)
0.0753	0.5511	0.2888	Male: 10 to 14 years
0.1335	0.3318	0.2309	Infant Mortality Rate
0.0626	0.3982	0.2145	Birth Rate
-0.3259	-0.0923	-0.2129	Prenatal Care Beginning in First Trimester
0.0968	0.3364	0.1923	N_Graffiti
0.0697	0.2227	0.1423	Breast cancer in females
-0.2459	-0.0209	-0.1398	Not Hispanic or Latino, Asian alone
0.0144	0.2186	0.1282	Tuberculosis
0.0107	0.2456	0.1021	Low Birth Weight

Table 4: BSLGPPR model's Beta Coefficients that did not contain zero in their 95% credible interval, sorted from Most to Least Impact
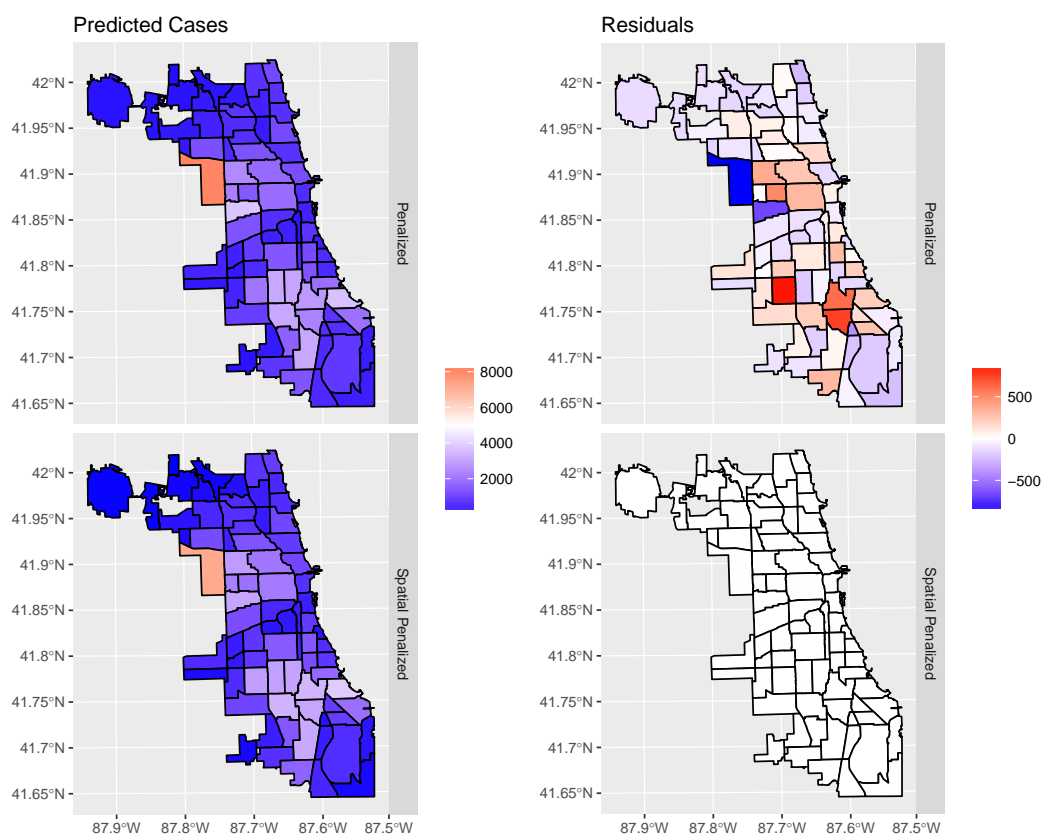
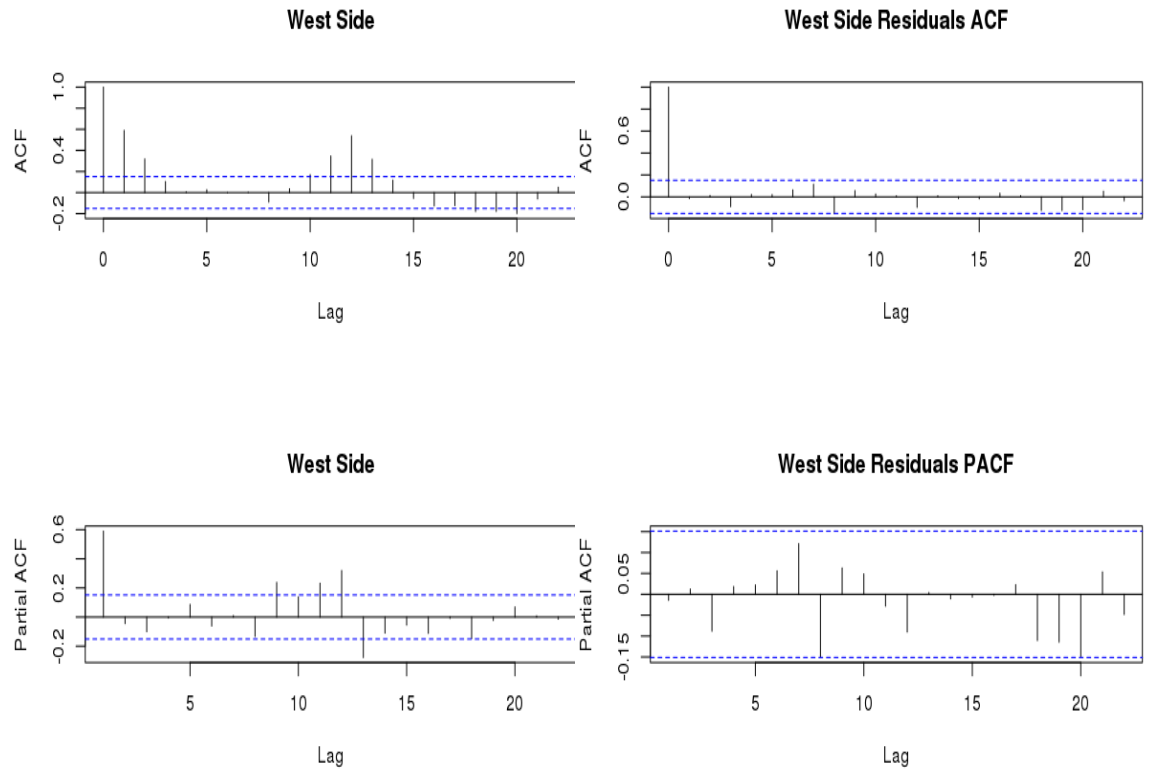Figure 8: Comparison of Two Model Predictions and Residuals

References

[1] City of Chicago, "Affordable Rental Housing Developments", *City of Chicago Open Data Project*, https://data.cityofchicago.org/Community-Economic-Development/Affordable-Rental-Housing-Developments/s6ha-ppgi, May 2017.

[2] City of Chicago, "Average Gas Usage per Square Foot by Community Area 2010", *City of Chicago Open Data Project*, https://data.cityofchicago.org/Service-Requests/311-Service-Requests-Graffiti-Removal/hec5-y4x5, May 2017.

[3] City of Chicago, "Census Data - Selected socioeconomic indicators in Chicago, 2008 2012", *City of Chicago Open Data Project*, https://data.cityofchicago.org/Health-Human-Services/Census-Data-Selected-socioeconomic-indicators-in-C/kn9c-c2s2, May 2017.

[4] City of Chicago, "Public Health Statistics- Selected public health indicators by Chicago community area", *City of Chicago Open Data Project*, https://data.cityofchicago.org/Health-Human-Services/Public-Health-Statistics-Selected-public-health-in/iqnk-2tcu, May 2017.

[5] City of Chicago, "Public Safety: Crimes 2001 to Present", *City of Chicago Open Data Project*, https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2/data, May 2017.

[6] City of Chicago, "Vacant Properties by Community Area", *City of Chicago Open Data Project*, https://data.cityofchicago.org/Community-Economic-Development/Vacant-Properties-by-Community-Area/cjr7-qyw5, May 2017.

[7] City of Chicago, "311 Service Requests - Graffiti Removal", *City of Chicago Open Data Project*, https://data.cityofchicago.org/Environment-Sustainable-Development/Average-Gas-Usage-per-Square-Foot-by-Community-Are/jp7r-k6h2, May 2017.

[8] CMAP Data Hub, "2010 Census Data Summarized to Chicago Community Area", *Chicago Metropolitan Agency for Planning*, https://datahub.cmap.illinois.gov/dataset/2010-census-data-summarized-to-chicago-community-areas/resource/b30b47bf-bb0d-46b6-853b-47270fb7f626?inner_span=True, May 2017.

[9] Papachristos, Andrew V., "48 Years of Crime in Chicago: An Analysis of of Serious Crime Trends from 1965-2013", http://isps.yale.edu/sites/default/files/publication/2013/12/48yearsofcrime_final_ispsworkingpaper023.pdf, December 2013.

[10] Pearson, Rick, "Trump Again Assails Chicago gun violence in speech to Congress", *Chicago Tribune*, http://www.chicagotribune.com/news/local/politics/ct-donald-trump-congress-speech-chicago-met-20170228-story.html, March 2017.

5 Appendix

5.1 Time Series Modeling Plots

5.1.1 West Side - ARIMA(2,0,0) x (1,0,0)[12]

	ar1	ar2	sar1
Coefficient	0.5742	0.0329	0.5735
Standard Error	0.0791	0.0796	0.0643

Table 5: Summary of model fit for the AR(2) with period 12 seasonal component fit to the monthly count of armed robberies for the West Side.

5.1.2 Southwest Side - ARIMA(12,0,0) x (1,0,0)[12]
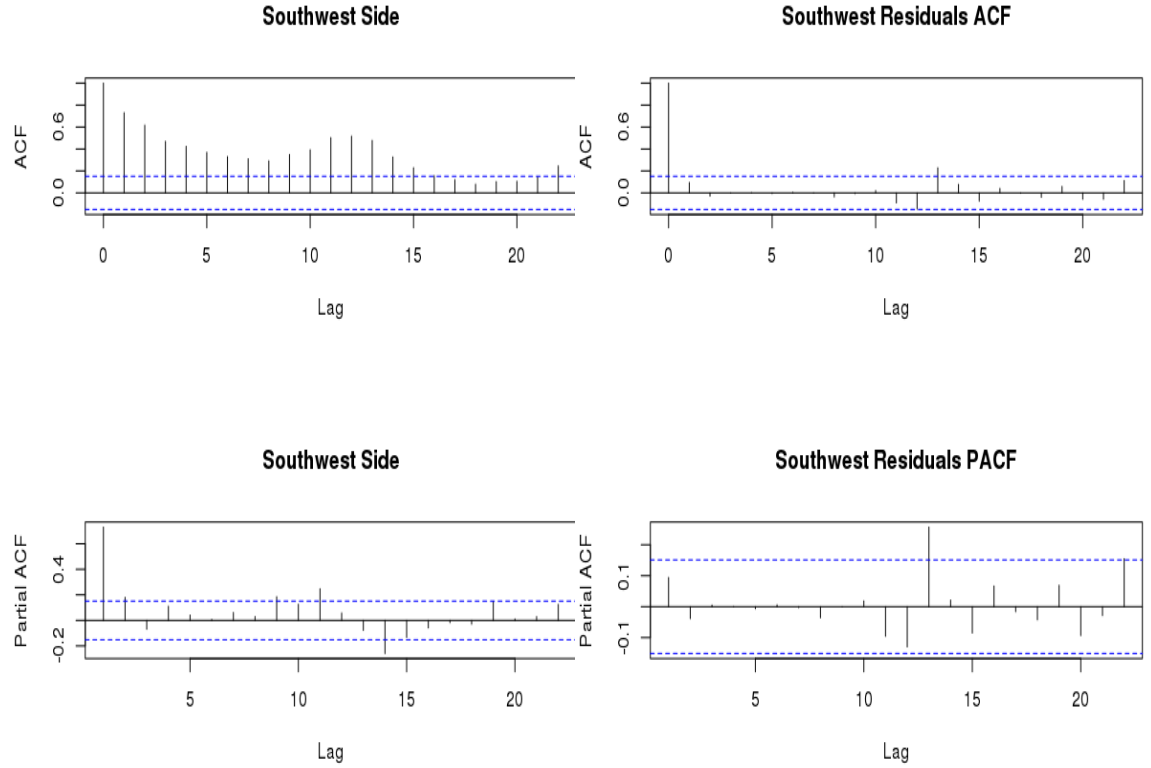
	ar1	ar2	ar3	ar4	ar5	ar6	ar7
Coefficient	0.3718	0.2471	-0.0041	0.1178	0.0534	0.0521	-0.0722
Standard Error	0.0861	0.0791	0.0837	0.0811	0.0829	0.0844	0.0820
		ar8	ar9	ar10	ar11	ar12	sar1
Coefficient		-0.0419	0.0943	-0.0330	0.1937	-0.3336	0.6148
Standard Error		0.0819	0.0810	0.0827	0.0843	0.0978	0.1077

Table 6: Summary of model fit for the AR(12) with period 12 seasonal component fit to the monthly count of armed robberies for the Southwest Side.

5.1.3 Far Southeast Side - ARIMA(12,0,0) x (1,0,0)[12]
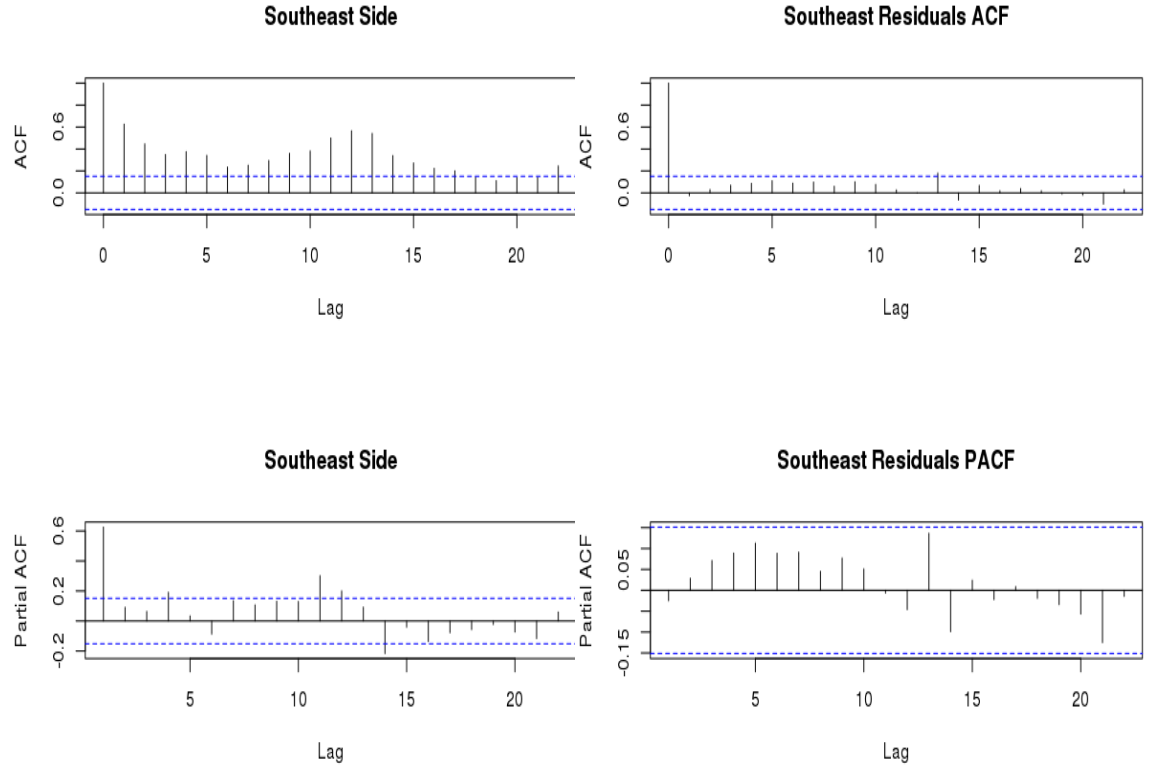
	ar1	ar2	ar3	ar4	ar5	ar6	ar7
Coefficient	0.3841	0.0168	-0.0954	0.1271	0.1040	-0.1575	-0.0059
Standard Error	0.0742	0.0832	0.0797	0.0800	0.0796	0.0791	0.0815
	ar8	ar9	ar10	ar11	ar12	sar1	
Coefficient	0.0193	0.0552	-0.0223	0.2122	0.2868	-0.0748	
Standard Error	0.0802	0.0837	0.0795	0.0827	0.1119	0.1373	

Table 7: Summary of model fit for the AR(12) with period 12 seasonal component fit to the monthly count of armed robberies for the Far Southeast Side.

12

5.1.4 Northwest Side - ARIMA(3,0,0) x (1,0,0)[12]
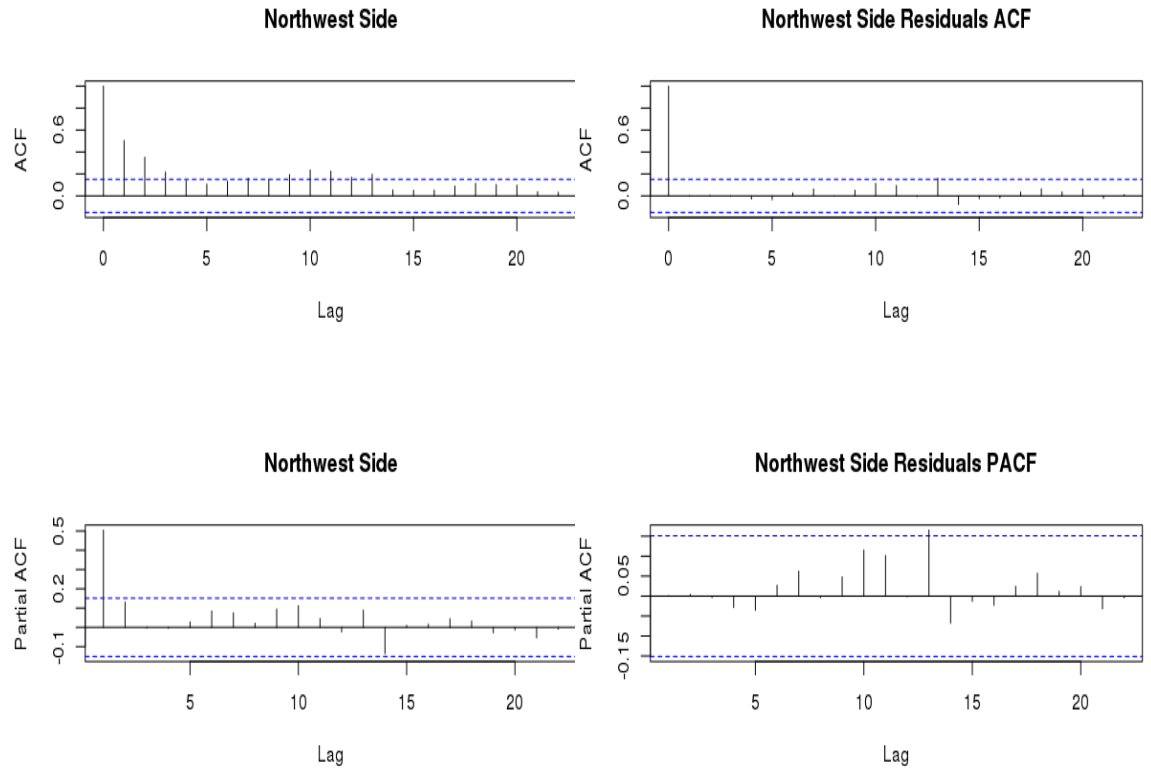
	ar1	ar2	ar3	sar1
Coefficient	0.4330	0.1290	0.0021	0.0169
Standard Error	0.0798	0.0842	0.0782	0.0848

Table 8: Summary of model fit for the AR(3) with period 12 seasonal component fit to the monthly count of armed robberies for the Northwest Side.

5.1.5 North Side - ARIMA(3,0,0) x (1,0,0)[12]
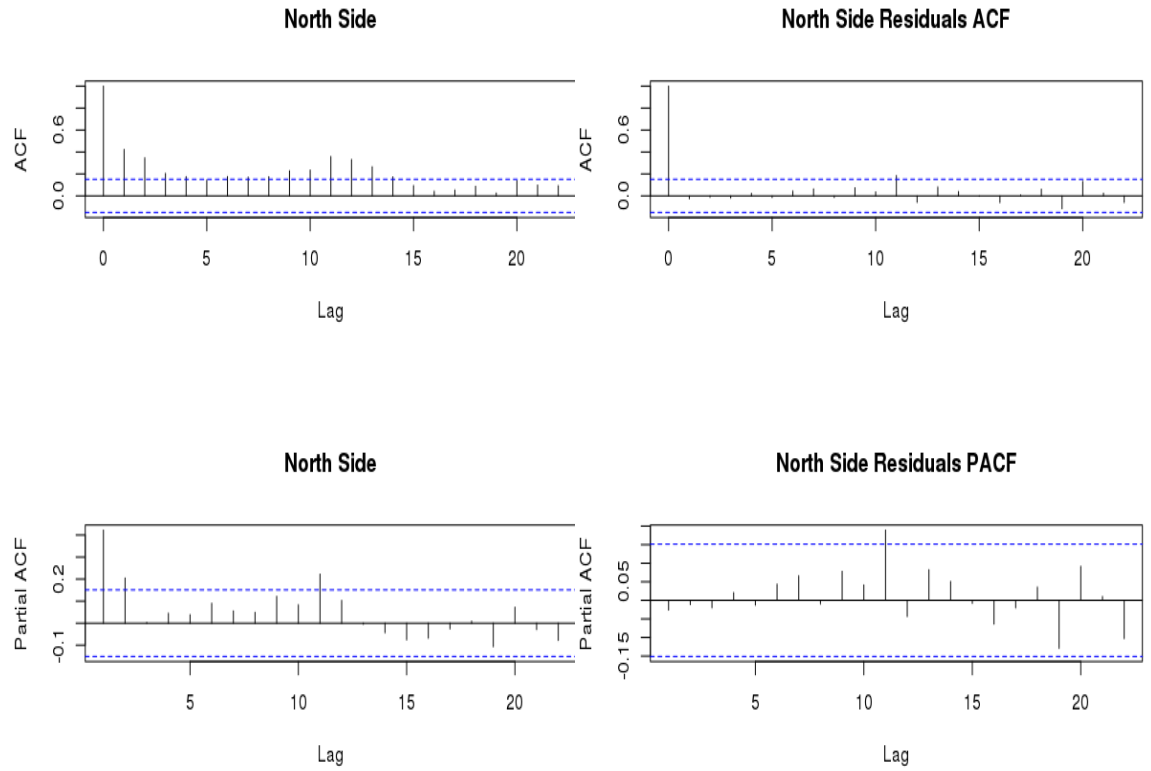
	ar1	ar2	ar3	sar1
Coefficient	0.2810	0.2050	0.0014	0.2441
Standard Error	0.0825	0.0816	0.0809	0.0871

Table 9: Summary of model fit for the AR(3) with period 12 seasonal component fit to the monthly count of armed robberies for the North Side.

5.1.6 Far Southwest Side - ARIMA(5,0,0) x (1,0,0)[12]
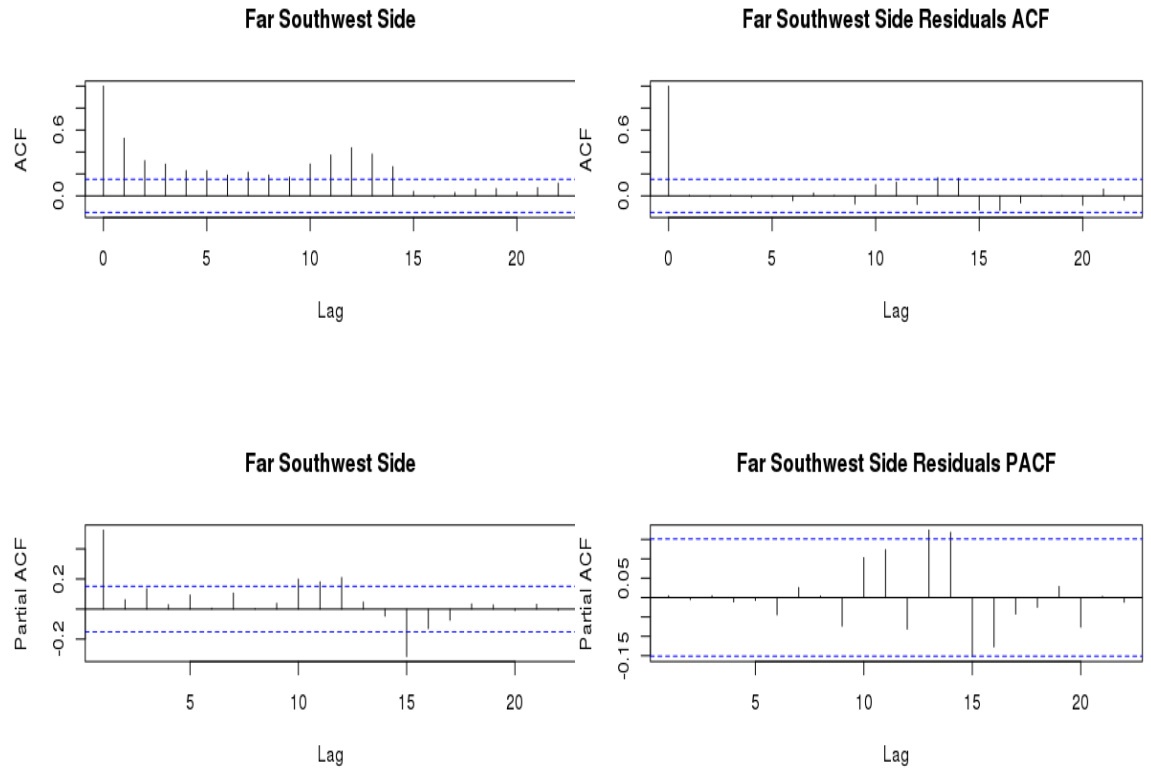
	ar1	ar2	ar3	ar4	ar5	sar1
Coefficient	0.3477	-0.0629	0.2363	0.0084	0.0928	0.3577
Standard Error	0.0825	0.0825	0.0823	0.0821	0.0774	0.0819

Table 10: Summary of model fit for the AR(5) with period 12 seasonal component fit to the monthly count of armed robberies for the Far Southwest Side.

5.1.7 Far North Side - ARIMA(3,0,0) x (1,0,0)[12]
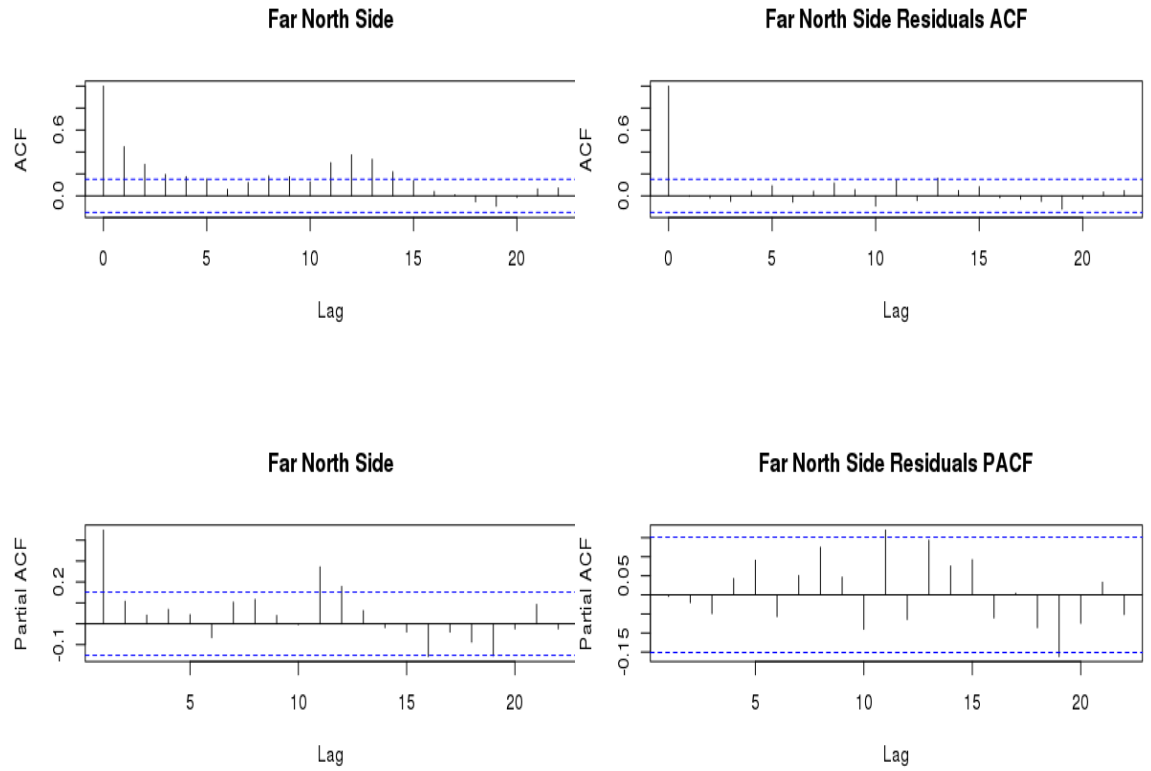
	ar1	ar2	ar3	sar1
Coefficient	0.3000	0.1374	0.0190	0.2987
Standard Error	0.0823	0.0805	0.0781	0.0848

Table 11: Summary of model fit for the AR(3) with period 12 seasonal component fit to the monthly count of armed robberies for the Far North Side.

5.1.8 Central - ARIMA(2,0,0) x (1,0,0)[12]

	ar1	ar2	sar1
Coefficient	0.3372	0.1944	0.2026
Standard Error	0.0793	0.0773	0.0921

Table 12: Summary of model fit for the AR(2) with period 12 seasonal component fit to the monthly count of armed robberies for the Central Side.
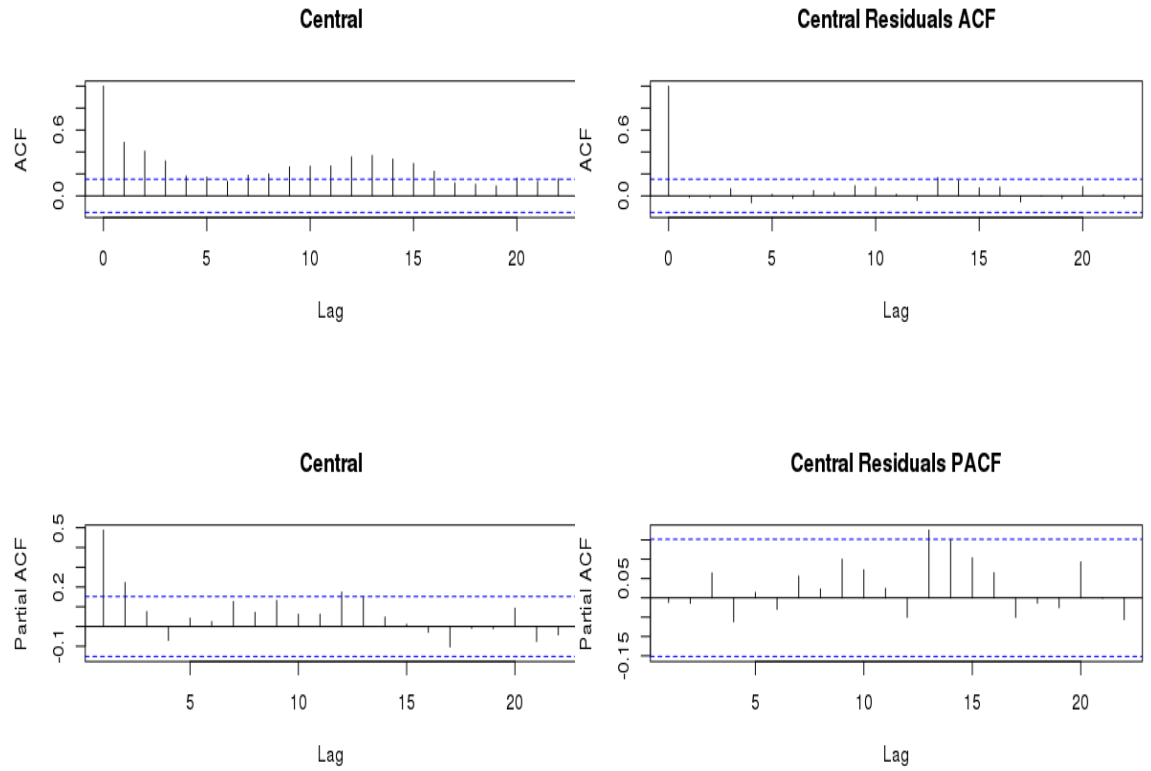
5.2 Spatial Analysis

Variable	Lasso	2.5%	97.5%	Spatial	IncludeZero
Intercept	6.5489	6.2136	6.2683	6.2427	FALSE
KWH Total SQFT	0.0000	-0.1712	0.4157	0.0727	TRUE
THERMS Total SQFT	0.0000	-0.2902	0.3763	0.0674	TRUE
N_Graffiti	0.1310	0.0968	0.3364	0.1923	FALSE
Birth Rate	0.0000	0.0626	0.3982	0.2145	FALSE
General Fertility Rate	0.0000	-0.3869	0.0174	-0.1396	TRUE
Low Birth Weight	0.0916	0.0107	0.2456	0.1021	FALSE
Prenatal Care Beginning in First Trimester	-0.1557	-0.3259	-0.0923	-0.2129	FALSE
Preterm Births	0.0000	-0.1675	0.0465	-0.0610	TRUE
Teen Birth Rate	0.0000	-0.1547	0.1216	-0.0162	TRUE
Assault (Homicide)	0.2175	-0.1223	0.3060	0.0898	TRUE
Breast cancer in females	0.1407	0.0697	0.2227	0.1423	FALSE
Cancer (All Sites)	0.0000	0.1618	0.4959	0.2893	FALSE
Colorectal Cancer	0.0124	-0.0272	0.1851	0.0822	TRUE
Diabetes-related	0.0000	-0.1356	0.0732	-0.0433	TRUE
Firearm-related	0.0400	-0.0884	0.1132	0.0265	TRUE
Infant Mortality Rate	0.2030	0.1335	0.3318	0.2309	FALSE
Lung Cancer	0.0000	-0.1050	0.1438	0.0014	TRUE
Prostate Cancer in Males	0.0227	-0.0799	0.1437	0.0198	TRUE
Stroke (Cerebrovascular Disease)	-0.0358	-0.1426	0.0592	-0.0379	TRUE
Tuberculosis	0.0790	0.0144	0.2186	0.1282	FALSE
Below Poverty Level	0.0000	-0.0601	0.2703	0.0843	TRUE
Crowded Housing	-0.0503	-0.3545	0.0036	-0.1928	TRUE
Dependency	0.0000	-0.0464	0.2258	0.0723	TRUE
No High School Diploma	0.0000	-0.1632	0.3010	0.0759	TRUE
Per Capita Income	0.0000	-0.2785	0.2422	-0.0064	TRUE
Unemployment	-0.2116	-0.8244	-0.3868	-0.5845	FALSE
N_Aff_Housing	0.0006	-0.0934	0.0205	-0.0323	TRUE
PERCENT OF HOUSING CROWDED	0.0000	-0.1809	0.0931	-0.0349	TRUE
PERCENT HOUSEHOLDS BELOW POVERTY	0.0000	-0.2042	0.2440	-0.0041	TRUE
PERCENT AGED 16+ UNEMPLOYED	0.0847	0.1699	0.5196	0.3433	FALSE
PERCENT AGED 25+ WITHOUT HIGH SCHOOL DIPLOMA	0.0000	-0.2138	0.4319	0.1477	TRUE
PERCENT AGED UNDER 18 OR OVER 64	0.0000	-0.2190	0.2157	-0.0505	TRUE
PER CAPITA INCOME	0.0000	-0.1572	0.1933	0.0253	TRUE
HARDSHIP INDEX	0.0000	-0.1290	0.6274	0.1752	TRUE
vacantLots	-0.0080	-0.0200	0.0946	0.0451	TRUE
Total Population	0.0000	-0.0719	0.3466	0.1129	TRUE
Not Hispanic or Latino, White alone	0.0000	-0.3291	0.1993	-0.0750	TRUE
Not Hispanic or Latino, Black or African American alone	0.2773	-0.1242	0.1845	0.0646	TRUE
Not Hispanic or Latino, American Indian and Alaska Native alone	0.0000	-0.1491	0.2238	0.0764	TRUE
Not Hispanic or Latino, Asian alone	-0.0045	-0.2459	-0.0209	-0.1398	FALSE
Not Hispanic or Latino, Native Hawaiian and Other Pacific Islander alone	0.0000	-0.0508	0.0893	0.0065	TRUE
Not Hispanic or Latino, Some Other Race alone	0.0000	-0.1073	0.3309	0.1343	TRUE
Not Hispanic or Latino, Two or More Races	0.0000	-0.3185	0.3276	0.0182	TRUE
Hispanic or Latino	0.1115	-0.2077	0.1797	0.0237	TRUE
Male: Under 5 years old	0.0000	-0.1909	0.1494	-0.0237	TRUE
Male: 5 to 9 years	0.0000	-0.2622	0.0710	-0.0919	TRUE
Male: 10 to 14 years	0.0000	0.0753	0.5511	0.2888	FALSE
Male: 15 to 17 years	0.0000	-0.3138	0.2186	0.0197	TRUE
Male: 18 and 19 years	0.0000	-0.0289	0.7760	0.3415	TRUE
Male: 20 years	0.0000	-0.0030	0.3619	0.1537	TRUE

Table 13: Beta Coefficients for LASSO and BSLGPPR Models.

Variable	Lasso	2.5%	97.5%	Spatial	IncludeZero
Male: 21 years	0.0000	-0.3656	0.1339	-0.0838	TRUE
Male: 22 to 24 years	0.0000	-0.1454	0.3723	0.0717	TRUE
Male: 25 to 29 years	0.0000	-0.1612	0.4841	0.1711	TRUE
Male: 30 to 34 years	0.0000	-0.4044	0.1591	-0.0959	TRUE
Male: 35 to 39 years	0.0000	-0.4585	0.3307	0.0278	TRUE
Male: 40 to 44 years	0.0000	-0.2925	0.2505	0.0403	TRUE
Male: 45 to 49 years	0.0000	-0.7821	-0.1529	-0.5485	FALSE
Male: 50 to 54 years	0.0000	-0.2311	0.2038	-0.0092	TRUE
Male: 55 to 59 years	0.0000	-0.1085	0.5342	0.1661	TRUE
Male: 60 and 61 years	0.0000	-0.1383	0.3300	0.0697	TRUE
Male: 62 to 64 years	0.0000	-0.4629	0.0427	-0.2165	TRUE
Male: 65 and 66 years	-0.3789	-0.5641	-0.0527	-0.2933	FALSE
Male: 67 to 69 years	0.0000	-0.2136	0.2511	0.0377	TRUE
Male: 70 to 74 years	0.0000	-0.4171	0.0526	-0.1750	TRUE
Male: 75 to 79 years	0.0000	-0.2765	0.1980	-0.0323	TRUE
Male: 80 to 84 years	0.0000	-0.1820	0.1216	-0.0021	TRUE
Male: 85 years and over	0.0000	-0.1266	0.2376	0.0535	TRUE
Female: Under 5 years old	0.0000	-0.0854	0.3244	0.0571	TRUE
Female: 5 to 9 years	0.0642	-0.2919	0.4038	0.0265	TRUE
Female: 10 to 14 years	0.0000	-0.9097	-0.1008	-0.4947	FALSE
Female: 15 to 17 years	0.0785	-0.1623	0.2829	0.0487	TRUE
Female: 18 and 19 years	0.0851	-0.2976	0.4667	0.0772	TRUE
Female: 20 years	0.0000	-0.3636	0.1191	-0.1461	TRUE
Female: 21 years	0.0381	-0.2372	0.1521	-0.0438	TRUE
Female: 22 to 24 years	0.1054	-0.2446	0.1472	-0.0319	TRUE
Female: 25 to 29 years	0.0000	-0.2606	0.2080	-0.0201	TRUE
Female: 30 to 34 years	0.1732	-0.5229	0.2026	-0.1613	TRUE
Female: 35 to 39 years	0.0000	-0.1611	0.4163	0.1265	TRUE
Female: 40 to 44 years	0.0000	-0.0885	0.4475	0.1260	TRUE
Female: 45 to 49 years	0.0000	-0.0847	0.3460	0.1258	TRUE
Female: 50 to 54 years	0.0000	-0.2909	0.1965	-0.0274	TRUE
Female: 55 to 59 years	0.0000	-0.1331	0.2360	0.0520	TRUE
Female: 60 and 61 years	0.0000	-0.1194	0.5062	0.2276	TRUE
Female: 62 to 64 years	0.0000	-0.3584	0.1626	-0.1239	TRUE
Female: 65 and 66 years	0.0000	-0.1784	0.4086	0.1105	TRUE
Female: 67 to 69 years	0.0000	-0.2072	0.3396	0.1342	TRUE
Female: 70 to 74 years	0.0000	-0.2815	0.2031	-0.0507	TRUE
Female: 75 to 79 years	0.0000	-0.2354	0.2365	-0.0560	TRUE
Female: 80 to 84 years	0.0863	-0.1139	0.4210	0.1324	TRUE
Female: 85 years and over	0.0377	-0.1938	0.1619	-0.0117	TRUE
Median Age	0.0000	-0.0419	0.2316	0.0797	TRUE
Total Households	0.0000	-0.2695	0.3955	0.0862	TRUE
Average Household Size	0.0000	-0.0545	0.2095	0.0747	TRUE
Total Housing Units	0.0000	-0.0947	0.9917	0.3907	TRUE
Occupied Housing Units	0.0000	-0.4459	0.1310	-0.1418	TRUE
Vacant Housing Units	0.1868	-0.3502	0.0549	-0.1469	TRUE
Occupied Housing Units__1	0.0000	-0.2624	0.2954	0.0121	TRUE
Owned with a mortgage or a loan	0.0000	-0.1566	0.6600	0.2384	TRUE
Owned free and clear	0.0000	-0.4927	0.1434	-0.0713	TRUE
Renter occupied	0.0000	-0.0892	0.5038	0.1460	TRUE

Table 14: Beta Coefficients for LASSO and BSLGPPR Models, continued.