

# **The application of aberration detection methods to cold weather injury and a spatiotemporal analysis of influenza in Ontario**

by

Timothy Chisamore

A practicum report submitted to the Department of Public Health Sciences in conformity with the requirements for the degree of Master of Science

Queen's University  
Kingston, Ontario, Canada  
August, 2016

© Timothy Chisamore, 2016

## **Abstract**

**Objectives:** We wanted to both develop a novel warning system for cold weather related injuries in Ontario and make preliminary comparisons between this new method and the existing method. Further, we aimed at conducting a spatiotemporal analysis of seasonal influenza in Ontario to better understand the diffusion of this costly annual illness.

**Methods:** To develop the new cold related warning system; we used Spearman and Pearson correlations, cross-correlation functions and variants of Shewhart control charts, which are forms of aberration detection methods. To conduct the spatiotemporal analysis, we used cyclic regression, colloquially known as Serfling regression, as well the ubiquitous Fourier periodogram. Further, newer wavelet methods were applied to better understand the diffusion of seasonal influenza.

**Results:** We found weak and moderate negative correlations between daily meteorological variables and daily counts of cold related injuries in the public health units of interest. Further, we found results suggestive of an acute relationship between exposure to cold temperatures and cold injury, rather than a lag relationship. Our new method looks promising when compared to the old method, we feel prospective comparisons will demonstrate an increase in specificity and sensitivity.

A general northwest to southeast spread of seasonal influenza in Ontario was observed by the separate methods of analysis used. Additionally, it became evident that implementing Serfling curves in areas of Ontario could be beneficial in determining the official start of the flu season and whether excess cases of influenza are being observed.

**Conclusion:** The current cold weather warning system employed in Ontario can be improved upon, the addition of aberration detection methods may help. Further, the diffusion of seasonal

influenza in Ontario occurs from the northwest to the southeast, thus, we recommend vaccination campaigns follow a similar pattern.

## **Acknowledgments**

To begin, I want to thank my academic supervisor Dr. Keyue Ding; he gave me the opportunity to work in the area of most interest to me. His input, support and guidance were integral in the completion of practicum. It was a pleasure working with you and I am forever thankful for the opportunity you gave me.

Secondly, Dr. Paul Belanger was an amazing practicum supervisor, he provided me with applied public health data and meaningful research problems, but also supported me in researching these. He had a huge impact on my career as a student and as a biostatistician, and I am very appreciative of that. Moreover, I appreciate the support and teamwork from the staff at Kingston, Frontenac and Addington & Lennox Public Health.

I want to recognize the professors, TAs and staff in the Department of Public Health Sciences for their impact and hard work. Without them, I never would have been able to enroll, nor complete my degree.

Working alongside my fellow biostatistics students Hongbo and Lauren also had a huge impact on me, thank you for the support. Additionally, I want to thank the students enrolled in epidemiology and public health, it was a pleasure. Good luck to you all in your respective careers and futures.

Further, I want to acknowledge the funding of Dr. Keyue Ding, Kingston, Frontenac and Lennox & Addington Public Health and Queen's University.

Lastly, thank you to my family and friends for their support while pursuing my goals, without you; none of this would be possible.

# Table of Contents

Abstract .....	ii
Acknowledgments.....	iv
Table of Contents.....	v
List of Tables.....	vi
List of Figures.....	vii
Chapter 1: Introduction.....	1
1.1: Background.....	1
1.2: Objectives.....	1
Chapter 2: The Application of Aberration Detection Methods to Cold Weather Injury in Ontario.....	2
2.1: Introduction.....	2
2.2: Methods.....	2
2.3: Results.....	4
2.4: Discussion.....	8
2.5: Summary.....	9
Chapter 3: Spatiotemporal Analysis of Seasonal Influenza in Ontario.....	11
3.1: Introduction.....	11
3.2: Methods.....	11
3.3: Results.....	14
3.4: Discussion.....	26
3.5: Summary.....	27
Chapter 4: Future Work.....	29

References.....30

# List of Tables

Table 1: Correlations Between Meteorological Variables and Daily Counts of Cold Weather Injuries.....5

Table 2: Dates that Alarms were Issued Retrospectively by the Aberration Detection Methods Along with Daily Minimum Temperatures and Daily Counts of Cold Weather Injury.....7

# List of Figures

Figure 1: Cross-correlation plot for Toronto.....	6
Figure 2: The C1 surveillance plot for Toronto.....	9
Figure 3: The C1 surveillance plot for Northern Ontario.....	10
Figure 4: Serfling Curve for P9N in the 2009 & 2010 Flu Year.....	15
Figure 5: Serfling Curve for K6V in the 2010 & 2011 Flu Year.....	16
Figure 6: Diagnostic Plots for the 2010 & 2011 Flu Year Serfling Curve in K6V.....	17
Figure 7: Choropleth Map of the Median Start Week of Influenza by FSA.....	18
Figure 8: Zoomed in Figure 7 (GTA).....	18
Figure 9: A Choropleth Map of the Local Moran's I Statistic Values for the Median Start Week of Influenza by FSA.....	19
Figure 10: A Choropleth Map of the P-values Associated with the Local Moran's I Statistics Values for the Median Start Week of Influenza by FSA.....	19
Figure 11: A Choropleth Map of the Local Moran's I Statistic Values for the Excess Influenza Proportions by FSA for the 2007 & 2008 Flu Year.....	20
Figure 12: A Choropleth Map of the Local Moran's I Statistic Values for the Excess Influenza Proportions by FSA for the 2008 & 2009 Flu Year.....	20
Figure 13: A Choropleth Map of the Local Moran's I Statistic Values for the Excess Influenza Proportions by FSA for the 2009 & 2010 Flu Year.....	21
Figure 14: A Choropleth Map of the Local Moran's I Statistic Values for the Excess Influenza Proportions by FSA for the 2010 & 2011 Flu Year.....	21
Figure 15: A Choropleth Map of the Local Moran's I Statistic Values for the Excess Influenza Proportions by FSA for the 2011 & 2012 Flu Year.....	22

Figure 16: A Choropleth Map of the Local Moran’s I Statistic Values for the Excess Influenza Proportions by FSA for the 2012 & 2013 Flu Year.....	22
Figure 17: A Choropleth Map of the Local Moran’s I Statistic Values for the Excess Influenza Proportions by FSA for the 2013 & 2014 Flu Year.....	23
Figure 18: The Periodogram for K6Vs Weekly Proportion of Influenza Time Series.....	23
Figure 19: The Contour Plot of the Power Normalized by the Variance for the K6V Weekly Proportion of Influenza Time Series Morlet Wave Transform.....	24
Figure 20: The Contour Plot of the Wave Coherence Between M5A and K6V for the Weekly Proportions of Influenza Time Series Morlet Wave Transforms.....	25
Figure 21: A Choropleth Map of the Median Phase Differences Using M5A as a Comparator...27	
Figure 22: Zoomed in Figure 22 (GTA).....	28

# Chapter 1

## Introduction

### 1.1 Background

Although cold weather injuries are a real concern in Ontario, they do not receive as much research as we expected. It's unclear the extent of morbidity each year due to cold weather injuries in Ontario, but it is obvious that it is non-zero. Further, since these injuries are completely avoidable, a reduction of morbidity is easily attainable by proactive warning systems. Such systems would benefit the general population since they would reduce the burden on healthcare due to injuries caused both directly and indirectly by cold weather exposure.

Seasonal influenza results in considerable morbidity and mortality worldwide with an estimated 1 million deaths annually [14]. No estimates on the morbidity and mortality in Ontario were found, but it is likely that they are very significant. Further, this illness results in time off of school and work; thus, it has a large impact on society beyond just the morbidity and mortality. Vaccines exist for seasonal influenza and so understanding the spatiotemporal evolution of this malady across Ontario is essential and allows for preventive measures to be put into place.

### 1.2 Objectives

The main objective of our work was twofold; first, we wanted to develop a better warning system for cold related injury in Ontario. The current system relies only on meteorological variables and it is unclear whether these can reliably predict cold related injuries. Second, we desired a better understanding of the spatiotemporal diffusion of seasonal influenza in Ontario, doing so should lead to more informed vaccination campaigns and preparation.

## Chapter 2

# The Application of Aberration Detection Methods to Cold Weather Injury in Ontario

### 2.1 Introduction

The application of aberration detection methods to public health data has typically been constrained to monitoring counts of infectious diseases [9]. However, these methods are increasingly being applied to counts of chronic disease and injury. Kingston, Frontenac and Lennox & Addington (KFL&A) Public Health has access to daily counts of cold weather related injuries, among other injuries and diseases, for 127 hospitals in Ontario via the Acute Care Enhanced Surveillance (ACES) network. Traditionally, warnings for frostbite and other cold related injury have been based on meteorological data, yet this does not make use of any injury specific surveillance data collected. We aimed at employing methods from the Early Aberration Reporting System (EARS) in conjunction with meteorological data to improve reporting and warning systems for cold weather related injury in Ontario.

Cold weather injury is a collective diagnosis that is composed of both frostbite and hypothermia, as well as other injuries sustained either directly or indirectly due to an individual being exposed to cold temperatures. Upon an individual arriving in the emergency department, a chief complaint will be attached to their visit. We had access to these data via ACES, but there are some probable limitations. It isn't clear whether health seeking behavior for cold related injuries is homogenous across the areas studied, we would hypothesize that Toronto would see markedly less cases of reported cold related injury due to underreporting by vulnerable populations. Further, there is no feasible way to determine the percentage of cold weather injuries captured by the ACES data since no overall background rate is known.

### 2.2 Methods

Cross-correlation functions quantify the correlation between two time series at integer lag values. As usual, the value is bounded between -1 and 1; such functions allow us to determine whether a lag relationship exists amongst two time series. These functions look similar to the common autocorrelation function used in time series analysis; however, they are generalized to incorporate a second time series. We define cross-correlation functions in the following way:

$$\rho_{XY}(\tau) := \frac{\mathbb{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]}{\sigma_X \sigma_Y}$$

where  $X$  and  $Y$  are the given time series,  $\mathbb{E}$  is the conventional expectation operator and  $\tau$  is the given lag value. Further,  $\mu_X$  and  $\mu_Y$  are the means of time series  $X$  and  $Y$ , respectively, while  $\sigma_X$  and  $\sigma_Y$  are the standard deviations.

Statistical Process Control (SPC) methods, originally developed to monitor and control industrial processes, have been used in the surveillance of public health data [15]. More specifically, variants of Shewhart control charts have been adopted as reliable aberration detection methods for public health surveillance where limited historical data is available [5]. The  $C_1$ ,  $C_2$  and  $C_3$  methods, developed by the Centers for Disease Control and Prevention (CDC), are examples that show increasing sensitivity [8].

These methods use 7-day rolling averages and standard deviations to standardize the daily count, which in turn allows the use of quantiles to look for statistically significant counts, given a significance level  $\alpha$ .

$$C_1(t) := \frac{Y(t) - \bar{Y}_1(t)}{S_1(t)} \sim N(0, 1) \text{ under the null hypothesis of no outbreak [5]}$$

$$\text{where } Y(t) \text{ is the daily count, } \bar{Y}_1(t) := \frac{1}{7} \sum_{i=t-7}^{t-1} Y(i) \text{ and } S_1^2(t) := \frac{1}{6} \sum_{i=t-7}^{t-1} [Y(i) - \bar{Y}_1(i)]^2$$

An alert is set off if  $C_1(t) > z_{1-\alpha}$ , where  $\alpha$  is the desired significance level, though,  $\alpha = 0.001$  is recommended for the  $C_1$  algorithm [5]. Using this threshold, we can construct an upper confidence limit for daily counts:

$$U_1 := \bar{Y}_1(t) + z_{1-\alpha} S_1(t)$$

Similarly, the  $C_2$  method incorporates a 2-day lag into its baseline calculation as a way to increase the sensitivity by reducing the immediate effects of large increases on the baseline.

$$C_2(t) := \frac{Y(t) - \bar{Y}_2(t)}{S_2(t)} \sim N(0, 1) \text{ under the null hypothesis of no outbreak [5]}$$

$$\text{where } \bar{Y}_2(t) := \frac{1}{7} \sum_{i=t-9}^{t-3} Y(i) \text{ and } S_2^2(t) := \frac{1}{6} \sum_{i=t-9}^{t-3} [Y(i) - \bar{Y}_2(i)]^2$$

An alert is set off if  $C_2(t) > z_{1-\alpha}$ , where  $\alpha$  is the desired significance level, though,  $\alpha = 0.001$  is recommended for the  $C_2$  algorithm [5]. Using this threshold, we can construct an upper confidence limit for the daily count:

$$U_2 := \bar{Y}_2(t) + z_{1-\alpha} S_2(t)$$

The  $C_3$  method relies on the  $C_2$  method and looks similar to a CuSum method in that it uses a cumulative sum of the previous day's values [5].

$$C_3(t) := \sum_{i=t-2}^t \max \{0, C_2(i) - 1\}$$

Unlike the prior methods, the distribution of this statistic is not standard normal. This is easily seen since the statistic cannot be negative, by definition. It isn't clear what distribution the  $C_3$

statistic follows, but conceivably, it will be similar to that of the CuSum method. Using simulation, we found that a cut-off value of 2.6 yielded an approximate 5% significance level, although we do not explicitly use this method in our analysis.

Additionally, we aggregated ACES cold weather injury data daily between the years of 2010 and 2016 in select public health unit regions, those being: Toronto, Northwestern, Thunder Bay and Porcupine. We also had access to daily meteorological information for these regions, variables of interest were: daily minimum temperature, daily maximum temperature, daily mean temperature, and diurnal range, which is defined as the difference between the daily maximum and daily minimum temperature.

We began our analysis by looking at correlations (both Spearman and Pearson) between daily counts of cold related injury in these select public health regions and the meteorological variables of interest. We chose to use both Spearman and Pearson correlations since it isn't clear whether a linear relationship should be expected; it is entirely possible that a non-linear monotonic relationship exists. Moreover, we looked for a lag relation between these by using cross-correlation functions (CCF) and plotting them for select lag values, which were based on physiologic data [6]. To finish our analysis, we retrospectively applied the family of  $C$  algorithms to our data in order to determine whether they were effective at capturing elevated counts and whether we could combine the traditional warning system, which bases its alerts off of meteorological data, with an aberration detection warning system. Since non-winter months correspond to long periods without cold related injury, we decided to restrict our data to fall between December 1<sup>st</sup> and March 31<sup>st</sup>; this prevents the baseline in the  $C$  methods from being affected by the overabundance of zero counts. Additionally, since natural cold related injury can only occur during this interval, the restriction of our data is theoretically consistent.

### 2.3 Results

Moderate negative correlations were found between daily counts of cold related injury and daily minimum, maximum and mean temperatures in each region of interest. Minimal positive correlations were found between daily counts of cold related injury and daily diurnal range in all regions.

From *Table 1*, there is some evidence towards a negative linear or negative non-linear monotonic relationship between daily minimum, maximum and mean temperatures and daily counts of cold related injury. Since correlation is a transitive property and daily minimum, maximum and mean temperature are all highly correlated, we would expect all of them to be correlated with daily cold related injury counts or none of them. Further, as these correlations are only moderate, this suggests a better warning system is needed. Traditionally, daily minimum temperatures have been used by Environment Canada to establish a warning system and so we will use this measure in our analysis.

It is well known that there is a lag between heat exposure and heat related mortality and morbidity [6], for these reasons, we were interested in knowing whether a similar lag existed between cold exposure and cold related mortality and morbidity. To determine this, we made use of cross-correlation functions and their plots using R. Initially, we needed to use first order

differencing to make both the count time series and the daily minimum temperature time series weakly stationary, this was further verified using autocorrelation functions that were plotted. When constructing cross-correlation plots between the time series, a maximum lag value of 14 days was used.

PHU	Daily Count and Daily Min. Temp.	
	Spearman	Pearson
<b>Toronto</b>	-0.3556186	-0.4227202
<b>Northwestern</b>	-0.2254353	-0.235652
<b>Thunder Bay</b>	-0.3119739	-0.2972599
<b>Porcupine</b>	-0.2652857	-0.2879026
PHU	Daily Count and Daily Diurnal Range	
	Spearman	Pearson
<b>Toronto</b>	0.0680736	0.1261248
<b>Northwestern</b>	0.02668961	0.003873671
<b>Thunder Bay</b>	0.02790925	0.008320056
<b>Porcupine</b>	0.1242861	0.1109848
PHU	Daily Count and Daily Mean Temp.	
	Spearman	Pearson
<b>Toronto</b>	-0.3635691	-0.4102995
<b>Northwestern</b>	-0.2333312	-0.2480959
<b>Thunder Bay</b>	-0.3449329	-0.3360687
<b>Porcupine</b>	-0.2677948	-0.2981839
PHU	Daily Count and Daily Max. Temp.	
	Spearman	Pearson
<b>Toronto</b>	-0.3508343	-0.3730992
<b>Northwestern</b>	-0.2311081	-0.2484023
<b>Thunder Bay</b>	-0.3360834	-0.3409816
<b>Porcupine</b>	-0.2414059	-0.2753933

*Table 1: Both Spearman and Pearson correlations between daily counts of cold related injury and meteorological variables of interest in each PHU.*

From Figure 1, there is evidence of statistically significant cross-correlation between the time series at lag values of 0, 2 and 4 days. Since the magnitude of these values is quite small ( $< 0.10$ ) and they occur within a short window of the initial exposure, it's not clear how clinically significant these lags are. Similar CCF plots were found for Thunder Bay, Porcupine and Northwestern, with significant positive and negative cross-correlations seen at positive lag values. However, the lag values in each were different, this may suggest some sort of underlying pattern or be due to random noise. Further investigation regarding the lag association should be

undertaken, ideally using similar methods that were able to demonstrate a lag association between heat exposure and morbidity [6].

### CCF Between Daily Minimum Temperature (C) and Daily ED Visits for Cold Related Injury

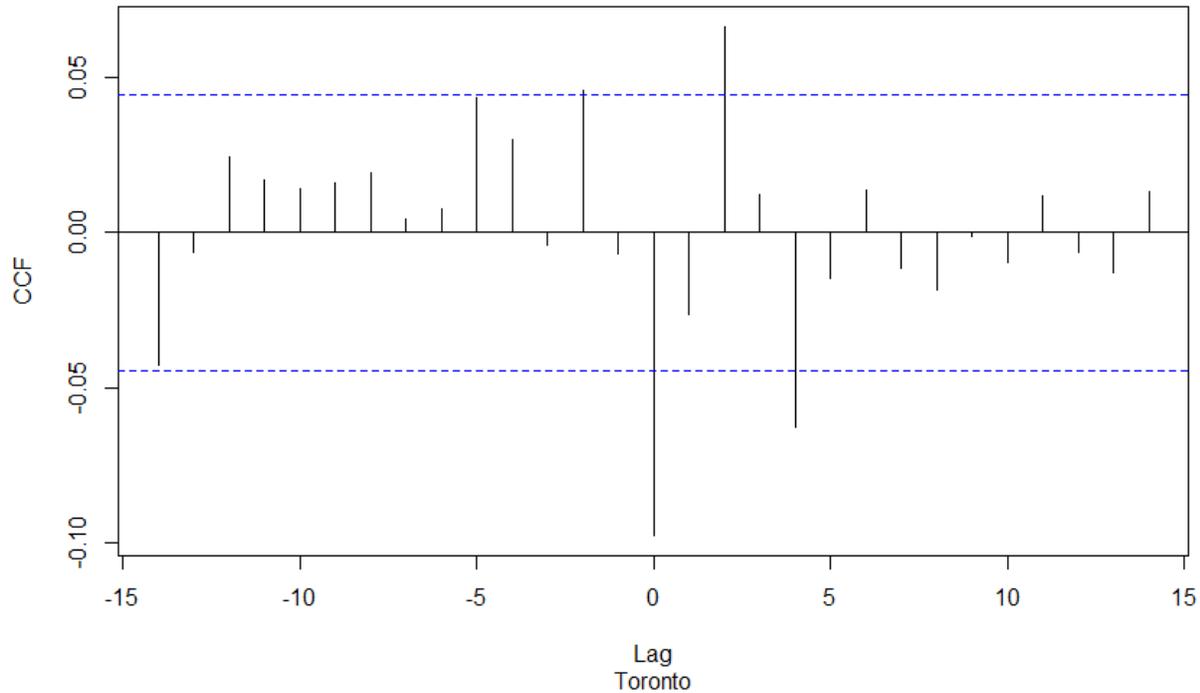


Figure 1: Cross-correlation plot for daily minimum temperature and daily ED visits for cold related injury in Toronto.

Our final component of the analysis was to retrospectively apply aberration detection algorithms to the data restricted between December 1<sup>st</sup> and March 31<sup>st</sup> each year. This restriction is necessary since natural cold related injuries are physiologically impossible during the non-winter months, and the abundance of zeros will heavily affect the baseline values.

The winters prior to 2014 did not lead to more than two cases of cold related injury on any day in Toronto. Due to the extreme temperatures experienced in Northern Ontario, we wanted to compare our results from Toronto with those from Northern Ontario, to do this; we aggregated the counts for the three northern public health regions previously mentioned. Again, the daily counts of cold related injury in Northern Ontario between the years 2011 and 2014 were low, with only two daily counts exceeding three. Moreover, it is known that the winters between 2011 and 2014 were unseasonably warm in Ontario [1] and so the lack of cold related injuries is not surprising, to facilitate a better comparison, we further restricted our data to the years 2014 and greater.

Applying the  $C_1$  method ( $\alpha=0.001$ ) retrospectively to both the Toronto and Northern Ontario data between December 1<sup>st</sup> and March 31<sup>st</sup> for the years 2014 and greater, we saw 20 and 21

alerts, respectively. *Table 2* highlights the corresponding date and daily minimum temperature when these alerts occurred. It is clear that Northern Ontario experiences colder minimum temperatures more frequently than Toronto does. However, larger counts occurred in Toronto, this is likely due to the larger total and vulnerable populations. Though, it is also possible that those in Northern Ontario can withstand colder temperatures before experiencing cold related injury or that they have a better understanding of the consequences of prolonged cold exposure. Regardless, it is evident that indicators based solely on meteorological variables do not work as well as we would like.

**Toronto C1 Aberration Detection Data (Winter Period 2014-2016)  $\alpha=0.001$**

Date	Number of Cases	Minimum Temp. (C)
03-01-14	5	-22.3
20-01-14	2	-16.5
15-02-14	3	-11.9
26-02-14	3	-14.2
13-03-14	3	-16.5
03-12-14	2	-0.6
18-12-14	2	-4.2
05-01-15	3	-12.5
08-01-15	5	-14.8
19-01-15	3	-7.7
30-01-15	2	-13.7
03-02-15	5	-14.0
15-02-15	11	-25.1
28-03-15	3	-10
19-12-15	2	-2.3
27-12-15	1	-0.6
31-12-15	2	0.6
13-02-16	4	-24.7
14-02-16	12	-22.4
02-03-16	1	-28.8

**Northern Ontario C1 Aberration Detection Data (Winter Period 2014-2016)  $\alpha=0.001$**

Date	Number of Cases	Minimum Temp. (C)
17-01-14	3	-23.0
24-02-14	2	-21.3
01-03-14	4	-30.1
02-03-14	7	-32.3
23-03-14	1	-27.0
31-03-14	1	-13.0
01-12-14	2	-25.3
13-12-14	1	-3.1
21-12-14	2	-9.4
31-12-14	2	-26.8

02-01-15	4	-30.1
30-01-15	3	-28.1
14-02-15	3	-31.0
23-03-15	1	-19.3
15-12-15	1	-9.4
01-01-16	1	-10.2
11-01-16	2	-27.5
08-02-16	1	-16.3
11-02-16	4	-28.5
02-01-16	1	-2.9
03-03-16	4	-30.9

*Table 2:* The dates, counts and minimum temperatures corresponding to the alerts in both Toronto and Northern Ontario using the C1 method for the study period.

## 2.4 Discussion

The application of non-linear lag distributed models may be necessary to further analyze whether a lag relationship exists between cold exposure and cold weather injury. This approach was developed and implemented by Gasparrinia to demonstrate the lag relationship between heat exposure and heat related injury [6]. It is clear that using cross-correlation functions alone cannot discount the fact that we may be seeing significant cross-correlations due to random noise alone.

We propose a novel approach to cold weather alerts and warnings, more specifically, we think a synthesis of aberration detection methods and meteorological variables will lead to a better warning system.

Using aberration detection methods alone can lead to missed alerts after a large shift in the baseline (Figure 2) or may lead to false alarms due to a very minimal baseline. We can use meteorological variables in the following way: suppose we see an aberrant count on any arbitrary day, if the meteorological variables are forecast to be similar the following days, we would expect further aberrant counts. This would lead to warnings that are based both on aberration detection methods and meteorological variables. Further, if we see elevated counts on a day with meteorological conditions not indicative of a warning, and the following days are projected to be the same, we could issue a warning.

However, we recognize that over-warning is not as worrisome as under-warning since the former will not lead to an increase in cold related injuries while the latter will. For these reasons, we should penalize under-warning more severely, however, we did demonstrate that the conventional method did fail to elicit warnings on days where aberrant counts occurred. Thus, a more rigorous comparison in terms of the type 1 and type 2 errors is required.

This combination of methods should increase both the specificity and sensitivity of the warning system since we are eliminating potential days where the original warning system produced false alarms as well as including potential days where the original system missed true alarms.

### C1 Surveillance of Cold ED Visits for Toronto Between 2014 and 2016

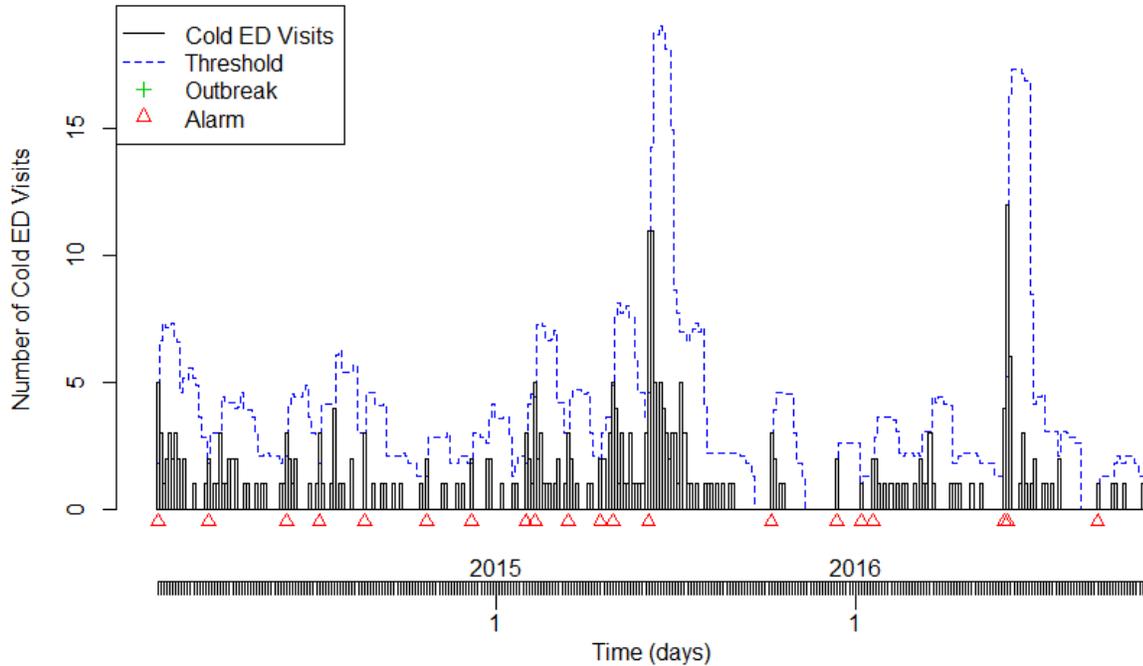


Figure 2: The C1 surveillance plot ( $\alpha=0.001$ ) for Toronto restricted to winter data (December 1<sup>st</sup> to March 31<sup>st</sup>) for the years 2014 and up [15].

#### 2.4 Summary

We found that the current cold weather warning system employed in Ontario had some flaws, it failed to elicit warnings on days where significant counts of injury occurred. We believe combining the current warning criterion with aberration detection methods such as variants of Shewhart control charts or CuSum should lead to a system with increased sensitivity and specificity. This could reduce the burden of preventable morbidity on the healthcare system in Ontario and lead to safer winters for residents of the province.

### C1 Surveillance of Cold ED Visits for Northern Ontario Between 2014 and 2016

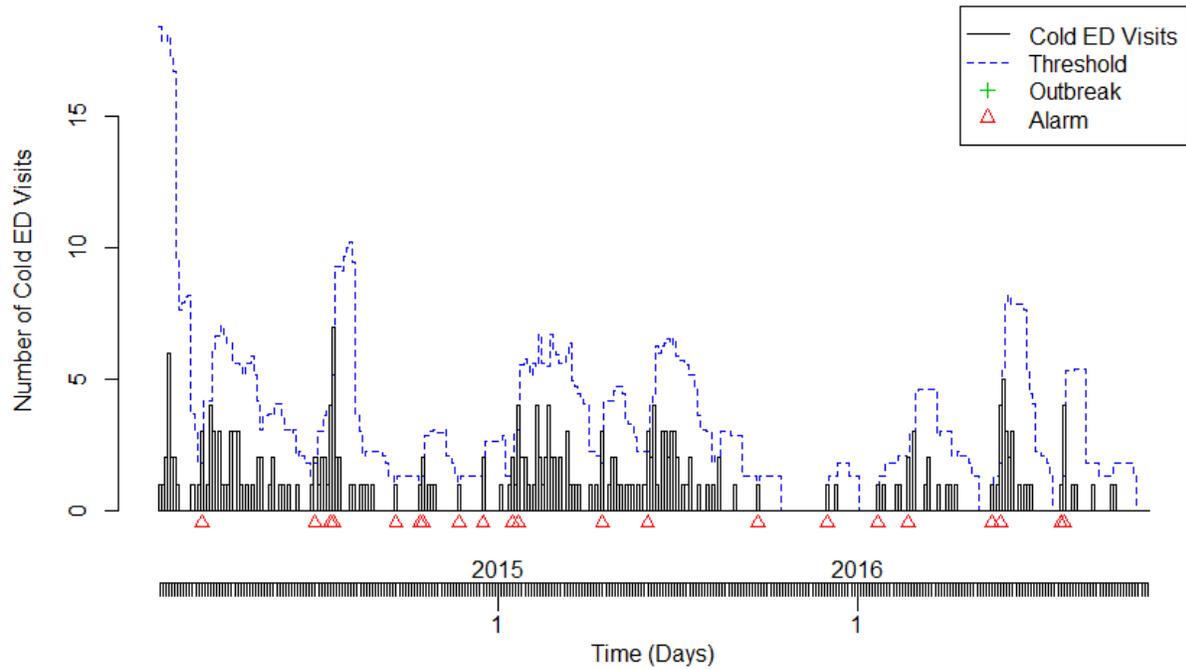


Figure 3: The C1 surveillance plot ( $\alpha=0.001$ ) for Northern Ontario restricted to winter data (December 1<sup>st</sup> to March 31<sup>st</sup>) for the years 2014 and up [15].

## Chapter 3

# Spatiotemporal Analysis of Seasonal Influenza in Ontario

### 3.1 Introduction

Spatiotemporal patterns of influenza are still not fully understood; a study conducted in the USA found a westerly to easterly spread [23], however, it isn't clear whether Ontario is subject to this spread. Knowledge of the spatiotemporal evolution of influenza is important to public health officials, as this information would allow a more refined stock of antivirals, better vaccination campaigns, and appropriate staffing of hospitals, for example. This would result in a more informed planning process and, thus, likely reduce the amount of money spent preparing for the influenza season.

Currently, methods based on time series and cyclic regression are used to study the spatiotemporal spread of influenza [20]. However, the dataset used in this study provides time series data for over 500 geographic units in Ontario and the aforementioned methods are largely aspatial and require plots to both select suitable parameters and to perform diagnostic checks. Moreover, it isn't clear how comparisons would be made between these models since they may differ significantly. For these reasons, a novel approach capable of making comparisons between the large number of geographical units is necessary.

### 3.2 Methods

Through Kingston, Frontenac and Lennox & Addington (KFL&A) Public Health, we had access to the total daily number of emergency department (ED) visits and the daily number of ED visits for influenza-like illness (ILI) in all 516 Forward Sortation Areas (FSA) of Ontario. FSAs are geographical units corresponding to a triplet of characters. The first character indicates the area of Ontario, the following character indicates whether the area is rural or urban and the final character encodes the size of the area. These data spanned January 1<sup>st</sup>, 2002 to March 31<sup>st</sup>, 2014. Influenza-like illness is a diagnosis assigned to symptoms that can be caused by various respiratory infections, for this reason, it's likely many of the cases caught are due to respiratory infections other than influenza. Further, it is unclear whether the reporting rate across public health units is homogenous, it is likely that areas with larger densities see much higher reporting rates due to increased transmission of respiratory infections and a greater presence of public health and community health programs.

After cleaning the data, we aggregated them by week in each FSA allowing us to calculate the proportion of weekly ED visits due to ILI per FSA. In other words, we had time series data for all 516 FSA composed of the weekly proportion of ED visits due to ILI between January 1<sup>st</sup> 2002 and March 31<sup>st</sup> 2014. However, some of these time series were missing data, while others were oversaturated with zero counts, which was likely due to their small sizes. For these reasons, we removed the following FSAs: K1P, K1X, K2R, K2V, K6T, L1L, L4V, L5S, L5T, L6G, L7K, L9J, L9Z, M5H, N3E, N6L, N6N, N8V, P0Y and P7L. The instability of the rates in these FSAs

as well as the nature of the missing data should be further investigated and methods to deal with these issues should be explored, however, we will not pursue this in our research.

The next step of our analysis was to construct Serfling curves for each FSA using five years of data in our baseline. This type of curve relies on cyclic regression and was originally developed by Robert Serfling in 1963 [19], but has been updated and modified several times over the years [18]. Traditionally, Serfling's method requires the removal of trend from the baseline data and to do this, we decomposed the time series data into their trend, seasonal and irregular components and kept only the latter two. To do this, we made use of the `stl` function built into R, which applies non-parametric regression. It is important to note that this method of removing trend is different from both Serfling's original method and the common current methods. For each FSA, we constructed models of the form:

$$y_t := \beta_0 + \beta_1 \cos\left(\frac{2\pi t}{52.18}\right) + \beta_2 \sin\left(\frac{2\pi t}{52.18}\right)$$

where  $t$  refers to the week,  $y_t$  is the sum of the seasonal and irregular components of the time series at week  $t$ ,  $\beta_i$  are regression coefficients and  $\frac{2\pi}{52.18}$  is included to account for the assumed periodicity of influenza since there are 52.18 weeks in a year, on average.

We then plotted the predicted curves based on these regressions, as well as the upper confidence interval at two standard deviations. Weekly data observed in the present flu year could then be overlaid on these plots to determine whether the observed data exceeds the expected data in each FSA. Summing the area above this threshold and observing the first week the threshold was crossed then provided us with both the total excess flu morbidity and the starting week of the flu season in each FSA. This information allowed us to examine whether seasonal flu was spatially structured or random in the study period. We then used Moran's  $I$  to test for global spatial autocorrelation in both the start week and excess influenza, and if present, we used local indicators of spatial autocorrelation (LISA) to examine local spatial autocorrelation. Conceptually, spatial autocorrelation refers to whether units in space are similar, dissimilar or random. Positive values of  $I$  indicate spatial units are similar while negative values indicate they are dissimilar, further,  $I = 0$  indicates spatial randomness.

Global Moran's  $I$  is given by:

$$I := \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

where  $w_{ij}$  is the  $ij^{th}$  entry of the matrix of spatial weights,  $N$  is the total number of spatial units,  $X$  is the variable of interest and  $\bar{X}$  is the arithmetic mean of the variable. Additionally, the matrix of spatial weights is constructed such that  $w_{ij} = 1$  if  $i \neq j$  share a border and 0 otherwise. It is important to note that  $I \notin [-1, 1]$  like the traditional measures of correlation. Finally, under  $H_0$  of no spatial autocorrelation, we can standardize  $I$  and thus determine the associated p-value [16].

Similarly, LISA can be defined as:

$$I_i := \frac{(X_i - \bar{X})}{m_2} \sum_j w_{ij} (X_j - \bar{X})$$

$$\text{where } m_2 := \frac{1}{N} \sum_i (X_i - \bar{X})^2$$

and  $X$ ,  $w_{ij}$  and  $N$  are as defined before. Further,  $I = \frac{1}{N} \sum_i I_i$ , which we would expect [2]. Note, multiple comparison correction is implemented by dividing  $\alpha$  for a given spatial unit by the number of neighbors that unit has plus one [3]. For example, if an FSA has six neighbors, then the corrected significance level would be  $\frac{\alpha}{6+1}$ . Moreover, global spatial autocorrelation is a single measure of the overall spatial autocorrelation present in the entire spatial entity composed of the spatial units while local spatial autocorrelation is a value assigned to each spatial unit measuring its spatial autocorrelation relative to the surrounding spatial units.

To complete our analysis, we wanted to consider periodograms of our time series to get an initial look at the period of the influenza season by FSA. We also performed wavelet analyses of these time series to understand the spatiotemporal patterns in the data. Wavelet methods are useful when time series are non-stationary and show changing variance, additionally; they allow us to make comparisons between multiple time series [22]. Selecting an FSA as a comparator, we can look at the differences in phase angles of the wavelets between FSAs to determine whether the flu seasons in these FSAs are in sync or out of sync. We chose to use M5A as the comparator since it lies in the greater Toronto area (GTA), which contains a majority of Ontario's population.

The periodogram is defined as:

$$I\left(\frac{j}{n}\right) := \frac{n}{2} (\hat{A}_j^2 + \hat{B}_j^2)$$

$$\text{where } \hat{A}_j := \frac{2}{n} \sum_{t=1}^n Y_t \cos\left(\frac{2\pi t j}{n}\right),$$

$$\hat{B}_j := \frac{2}{n} \sum_{t=1}^n Y_t \sin\left(\frac{2\pi t j}{n}\right)$$

and  $f := \frac{j}{n}$  for  $j = 1, 2, \dots, k-1$  is the given Fourier frequency and  $Y_t$  is the  $t^{\text{th}}$  observation of the time series of interest [13].

Additionally, the wavelet transformation we will use is the Morlet wave given by:

$$W_n(s) := \sum_{n'=0}^{N-1} x_{n'} \psi^* \left[ \frac{(n-n')\delta t}{s} \right] \text{ where } \psi^*(\eta) := \left( \frac{\delta t}{s} \right)^{\frac{1}{2}} \pi^{-\frac{1}{4}} e^{i\omega_0 \eta} e^{-\frac{\eta^2}{2}}, \omega_0 := 6 \text{ and}$$

$$\eta := \frac{(n-n')\delta t}{s}$$

here,  $n$  is the localized time index,  $s$  is the wavelet scale,  $N$  is the number of elements in the time series being transformed,  $x_{n'}$  is the  $(n')^{th}$  observation from the time series,  $\delta t$  is the time step of the series and  $*$  refers to the complex conjugate operator [4, 17, 22].

These wavelet transforms allow us to look at both the phases and power spectra across multiple periods (they are powers of 2 to facilitate fast Fourier transforms (FFT)). Notice, this will be a 3-dimensional plot projected into 2-dimensions, thus, it will be a contour plot. We can calculate the power spectrum and phase angle as:

$$|W_n(s)|^2 \text{ and } \tan^{-1} \left[ \frac{\Im\{W_n(s)\}}{\Re\{W_n(s)\}} \right]$$

respectively [22]. Here,  $\Im$  refers to the imaginary component and  $\Re$  refers to the real component. It is important to note that this method requires the time series to be padded with zeros to facilitate proper comparisons for larger periods; this results in a cone of influence on the periphery of the plot whereby results may not be as accurate. We thus need to be careful when analyzing such plots and results [22].

### 3.3 Results

Using Serfling regression, we were able to create curves for all FSAs considered. Figure 1 shows an example of a Serfling curve for P9N in the 2009 and 2010 flu year.

It is clear that the 2009 H1N1 epidemic was picked up by the Serfling curve, evident by the marked increase above the epidemic threshold. Notice, further, the initial spike around week 35, which was seen to be in excess. It is possible that early information like this could lead to more timely responses.

From Figure 2, we see that the Serfling curve suggests the flu season began on week 37 and by summing the excess area; we can quantify the magnitude of the flu season. This was found to be 0.1174795, thus, in the 2010 and 2011 flu year; there was an excess weekly ILI proportion of almost 0.12 in K6V. These means that we saw about a 10% increase in the weekly ILI proportion compared to what we expected.

It is of interest to consider the diagnostic plots for these predictive curves, however, since we are creating one for 496 FSAs each year, this would be hard to implement. From Figure 3, there is some fanning of the residuals. There is also a slight deviation from the expected line in the normal QQ plot on the right tail. We should further examine these and consider either weighted least squares regression or potentially a Poisson regression model using an offset.

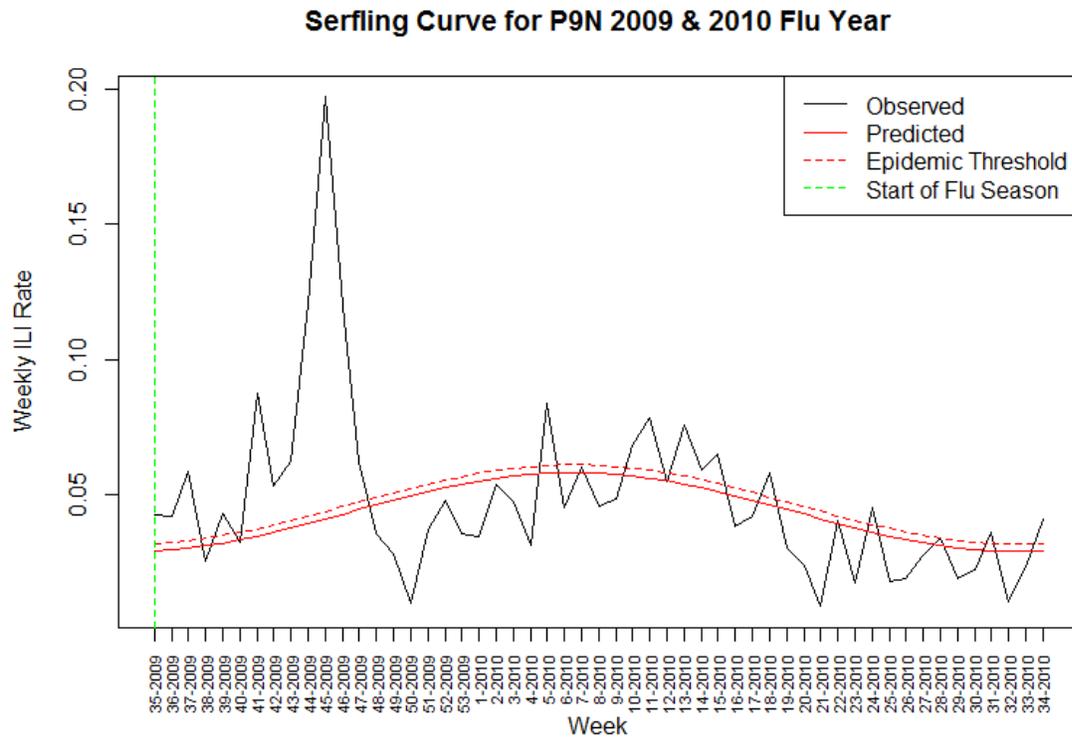


Figure 4: The Serfling curve for P9N in the 2009 and 2010 flu year, notice the sharp peak due to the H1N1 outbreak that flu year.

We were interested in knowing whether there was significant and annually consistent global spatial autocorrelation amongst flu season start weeks across our FSAs of interest. We derived a summary measure of the seven start weeks for each FSA corresponding to the seven Serfling curves constructed for the seven flu weeks included in the study. We used the median since it is unclear whether the distribution is symmetric and, hence, the mean may not be a meaningful measure of the central tendency. From Figures 4 and 5, notice the somewhat general trend of flu beginning later in the southern FSAs. The global Moran's  $I$  test returned  $I = 0.09$  which corresponds to a p-value of less than 0.001. Although the Moran's  $I$  statistic is quite small, it is still evidence of significant positive global spatial autocorrelation and so we carried out a LISA analysis and made use of choropleth plots to examine local spatial autocorrelation.

**Serfling Curve for K6V 2010 & 2011 Flu Year**

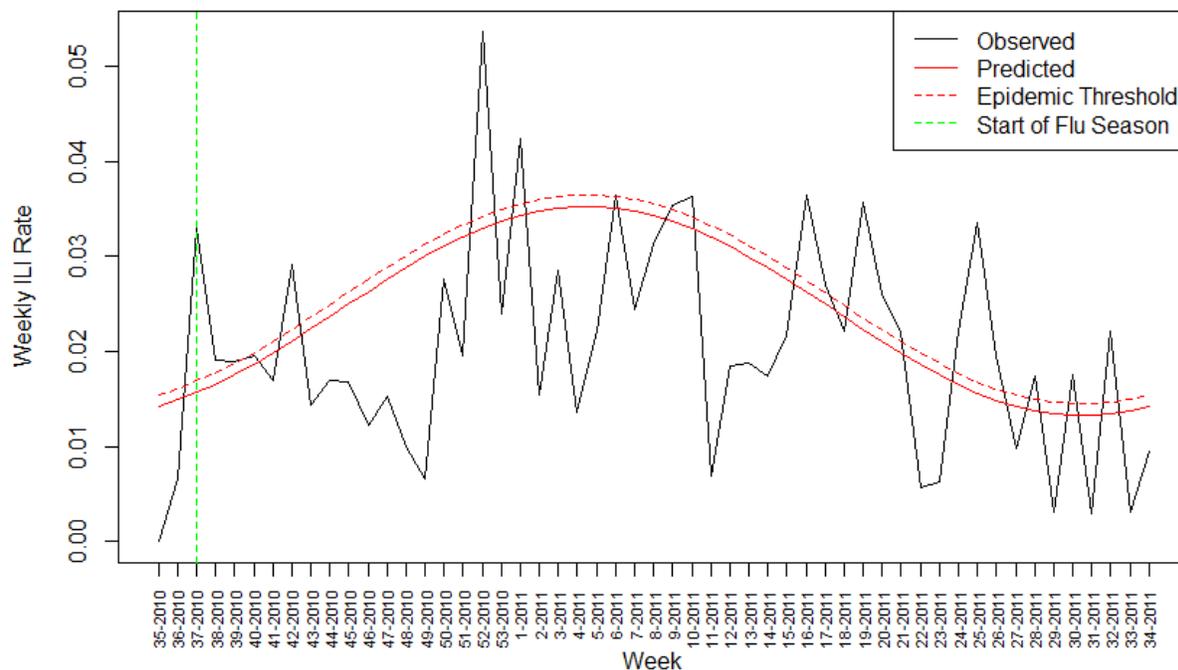


Figure 5: The Serfling curve for K6V in the 2010 and 2011 flu year.

Figures 6 and 7 illustrate the local Moran's I values and the associated p-values, respectively. There are areas with large positive local Moran's I values that correspond with significant corrected p-values, thus we see that some FSAs in Ontario are positively influenced by their neighbour FSAs in terms of the median starting week of the flu season.

We further wanted to investigate the excess rates of influenza yearly by FSA using again the same approach as was used above, but applying it to each flu year we have data for, rather than using the median value over the seven years of data as we did before.

Significant global spatial autocorrelation was found for each flu year observed, the p-values all being less than 0.001. Moreover, the global Moran's I statistics were 0.09, 0.21, 0.22, 0.18, 0.10, 0.18 and 0.22, respectively. We again appealed to LISA in order to determine the local spatial autocorrelation driving this process.

Figures 8 through 14 only partially uncover the spatiotemporal pattern of influenza in Ontario, they highlight that the spatial dependence of excess influenza is conditional on the flu year. This is troubling since our intent is to establish the spatiotemporal diffusion of influenza in Ontario.

Since these methods are somewhat descriptive, a more analytic technique is required to quantify the spatiotemporal patterns of influenza in Ontario. To do this, we applied both Fourier methods and wavelet analysis techniques to our data.

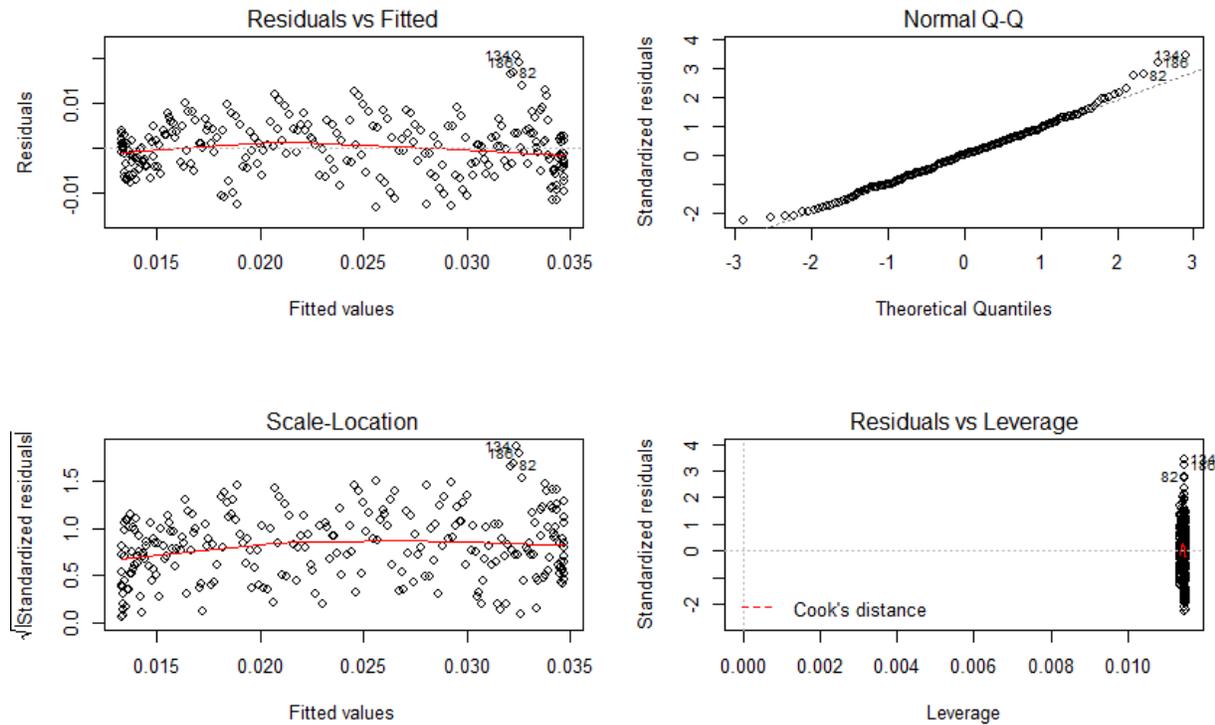


Figure 6: The regression diagnostic plots for the 2010 and 2011 flu year Serfling curve

Periodograms, as briefly described before, use given Fourier frequencies up to half the total length of the time series and roughly examine the magnitude of the periodic behavior at these frequencies. Higher values indicate stronger periodicity in the data for that given period.

The periodogram analysis of the time series corresponding to each FSA yielded a period of 52 weeks for all but 49 FSAs, inspection of these 49 revealed quite sparse counts due to either small population size or incomplete data. This resulted in a periods ranging from three weeks to the entire study length, rather than the expected 52 weeks. Additionally, a few FSAs had periods surrounding 52 weeks, but this is likely due to leakage between Fourier frequencies, (i.e., effects from 51, 52 and 53 leaking into the surrounding frequencies). This can lead to large periodograms associated with periods very near 52 weeks.

Custom Breaks (Median)

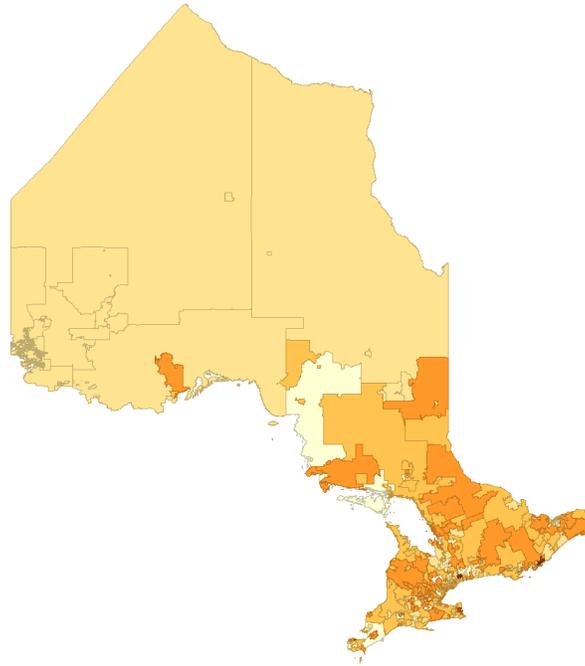
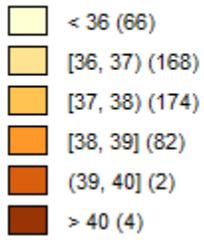
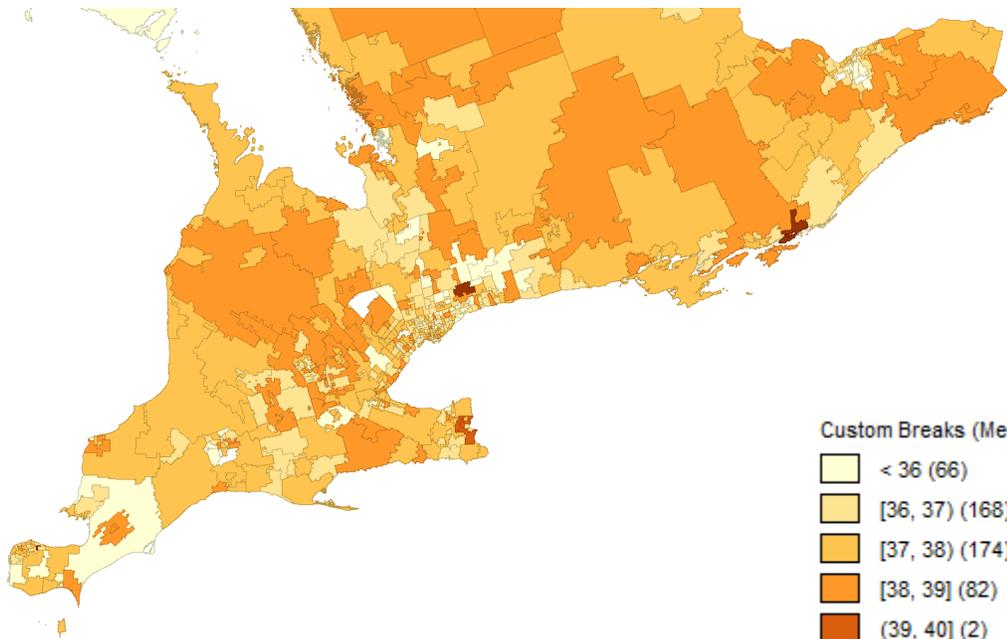


Figure 7: The choropleth map of the median start of the flu week in each FSA based off our Serfling curves.



Custom Breaks (Median)

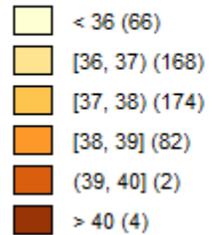


Figure 8: A zoomed in version of Figure 4 highlighting the Greater Toronto Area.

Custom Breaks (LMoran):

- < -1 (5)
- [-1, -0.5) (30)
- [-0.5, 0) (169)
- [0, 0.5] (237)
- (0.5, 1] (35)
- > 1 (20)

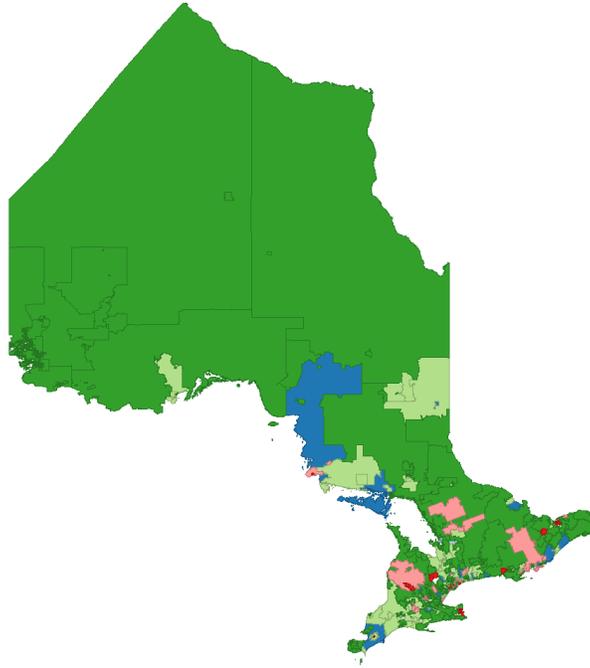
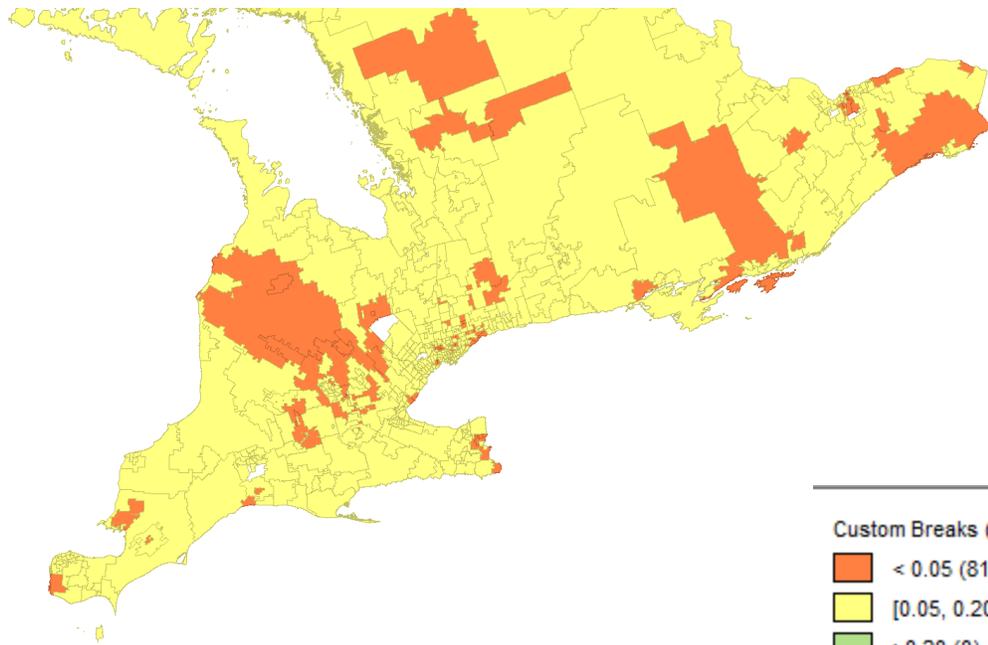


Figure 9: A choropleth map of the local Moran's I statistic value associated with the median start week of influenza in each FSA.



Custom Breaks (Pval):

- < 0.05 (81)
- [0.05, 0.20] (415)
- >0.20 (0)

Figure 10: A choropleth map of the p-values associated with the local Moran's I statistic values seen in Figure 6 for the Greater Toronto Area.

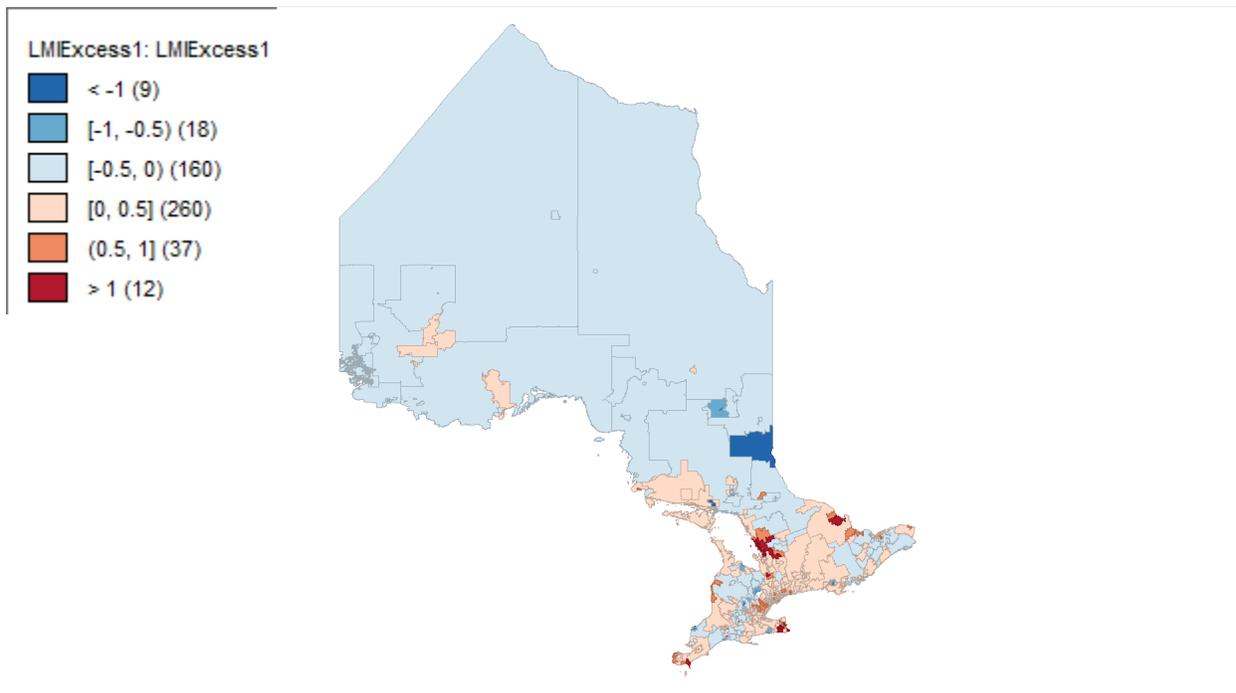


Figure 11: A choropleth map of the local Moran's I statistics associated with excess influenza rates by FSA for the 2007 and 2008 flu year.

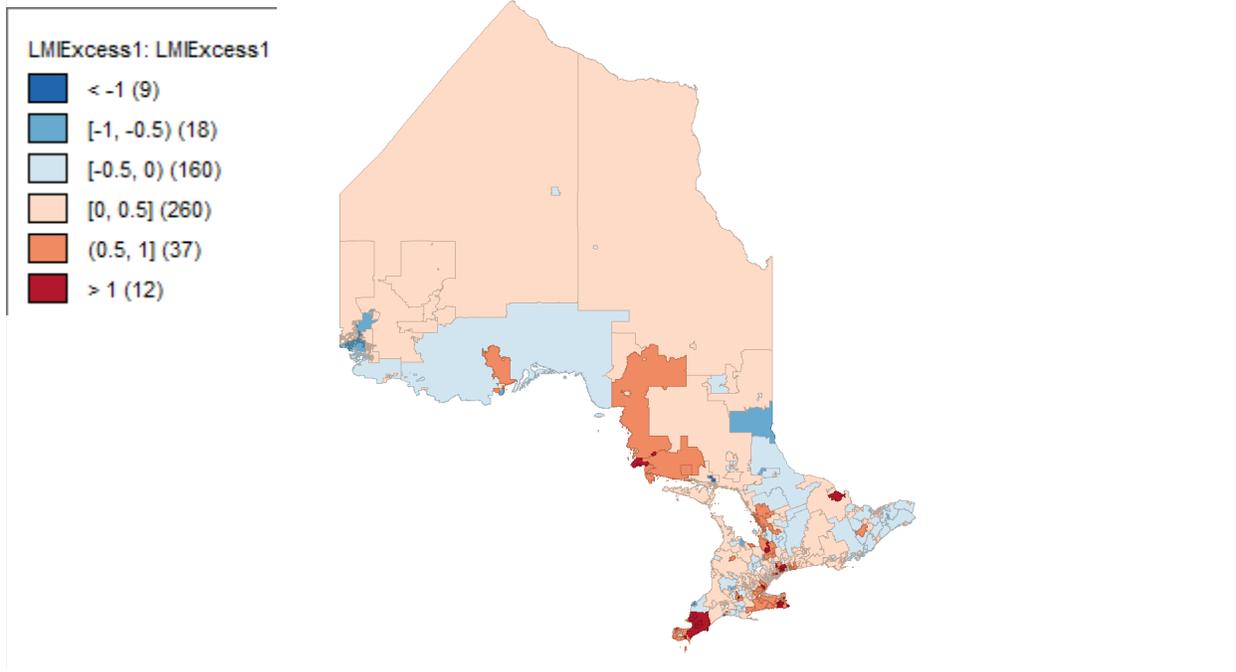


Figure 12: A choropleth map of the local Moran's I statistics associated with excess influenza rates by FSA for the 2008 and 2009 flu year.

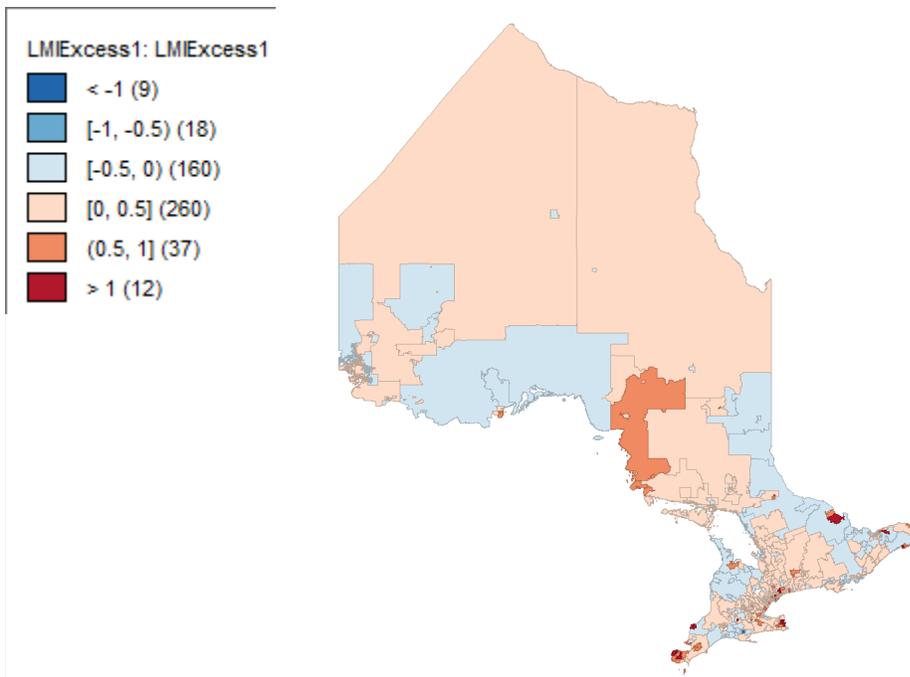


Figure 13: A choropleth map of the local Moran's I statistics associated with excess influenza rates by FSA for the 2009 and 2010 flu year.

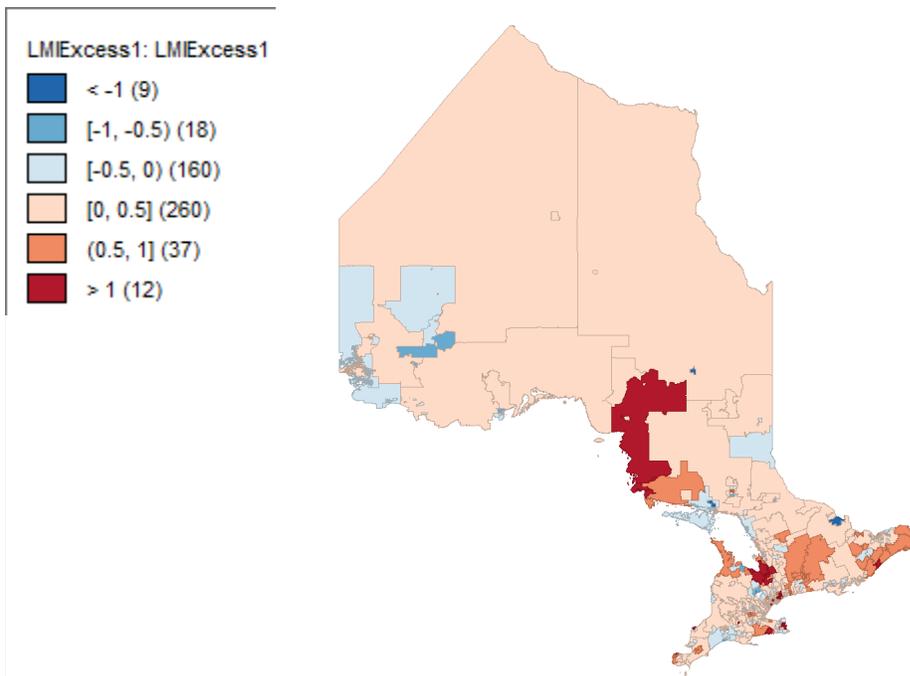


Figure 14: A choropleth map of the local Moran's I statistics associated with excess influenza rates by FSA for the 2010 and 2011 flu year.

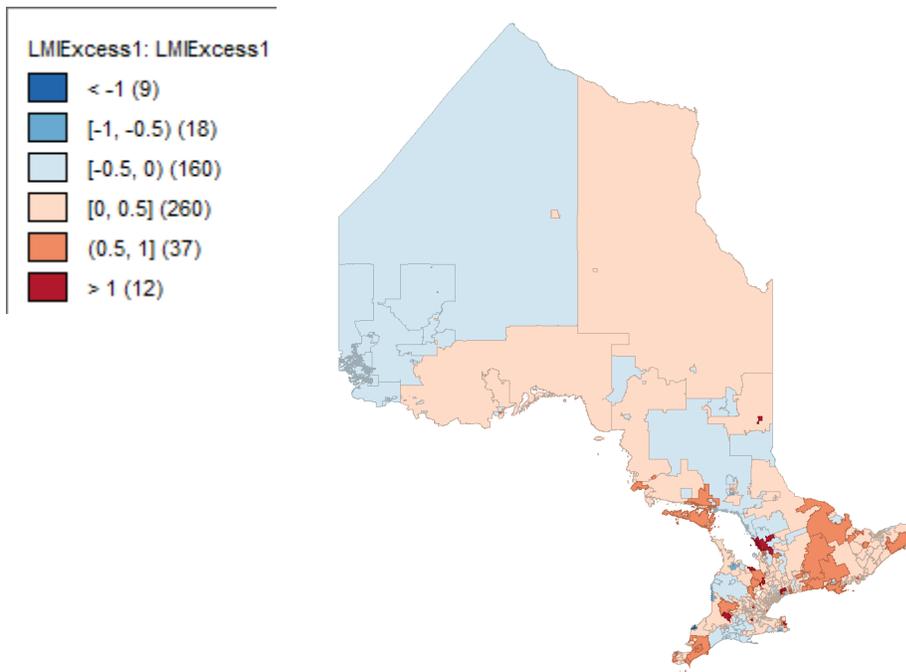


Figure 15: A choropleth map of the local Moran's I statistics associated with excess influenza rates by FSA for the 2011 and 2012 flu year.

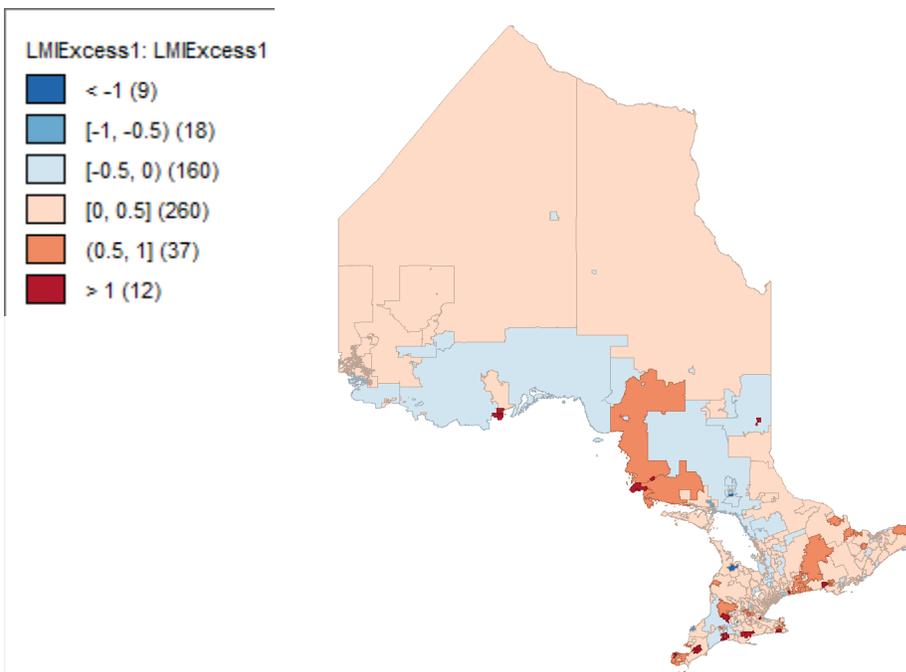


Figure 16: A choropleth map of the local Moran's I statistics associated with excess influenza rates by FSA for the 2012 and 2013 flu year.

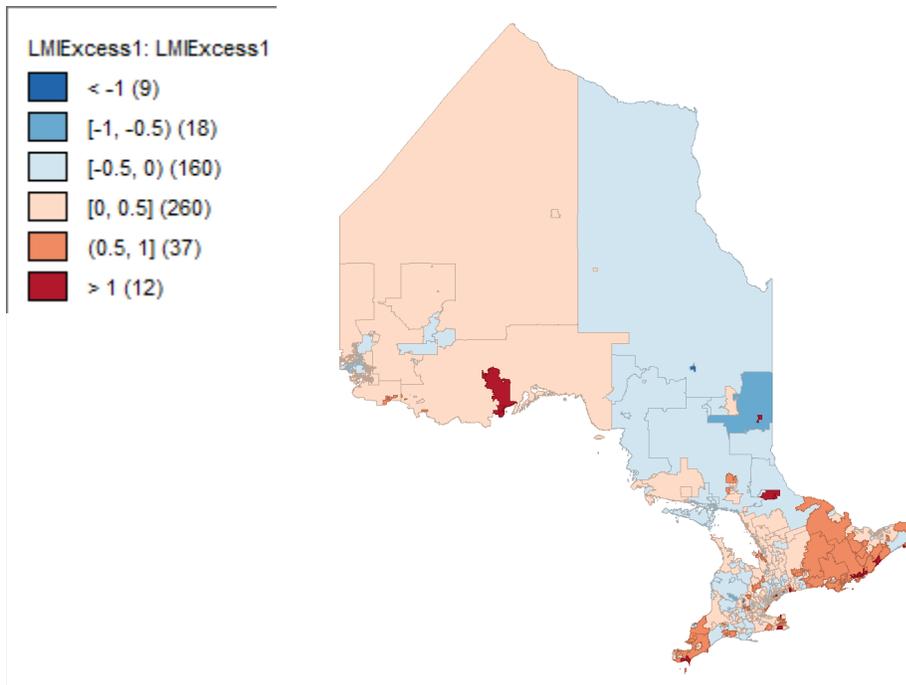


Figure 17: A choropleth map of the local Moran's I statistics associated with excess influenza rates by FSA for the 2013 and 2014 flu year.

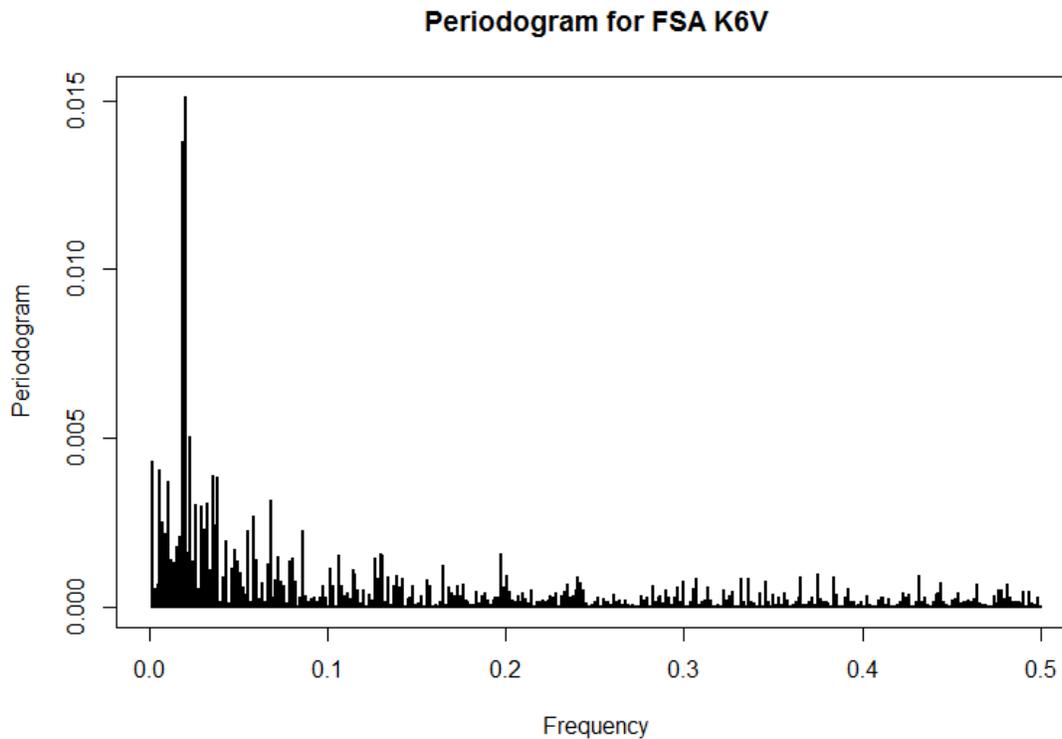


Figure 18: The periodogram for K6Vs time series, note the aforementioned leaking effect at the peak.

Figure 15 shows the periodogram for K6V, notice a sharp peak around a frequency of 0.02006173, which corresponds to a period of approximately 50 weeks.

Similarly, we can view the power spectrum contour plot for this FSA, which corresponds to looking at given periods versus time, in terms of week. Further, the contouring roughly equates to the strength of the given period at a given time, higher values indicating a greater association.

To read these plots, we look at the x-axis to see the given week of the time series and the y-axis to look at the potential period, in terms of week. Then, the contouring color indicates the strength of the period at that given week of the time series. Further, notice the white shading on the border, this is the cone influence, which was previously mentioned. For large periods being tested (such as 128 weeks), the length of the data may not be enough to facilitate proper testing and so it is padded with zeros to increase the length. This obfuscates the results and so we cannot be as certain in the regions that required padding. This cone indicates the boundary of the region and so we must be careful interpreting exterior of it.

Additionally, we can also analyze the coherence between two waves using similar plotting; in this case, the contouring indicates the level of coherence ranging between 0 and 1. Figure 17 is an example of this type of plot.

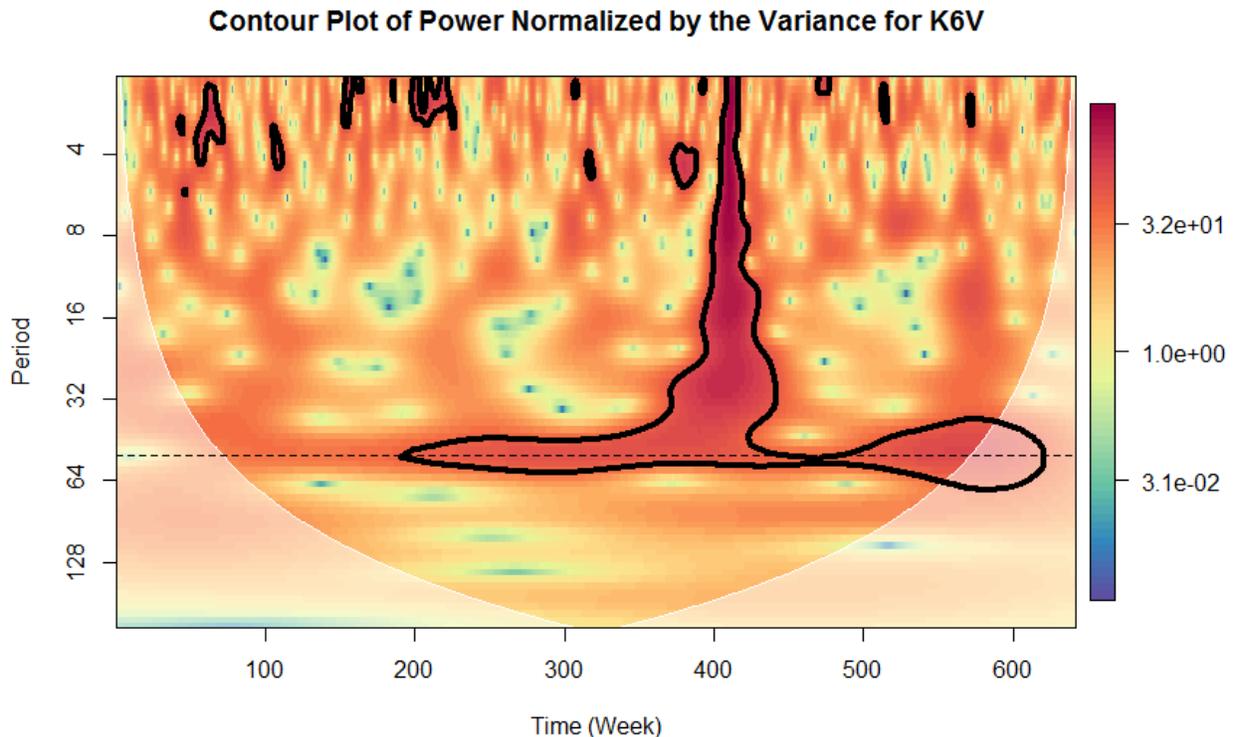


Figure 19: The contour plot of power normalized by the variance for K6V.

Figure 16 indicates similar results in terms of the periodicity of influenza in this FSA, here; the horizontal dashed line corresponds with a period of 52 weeks. Additionally, the thick black outlines indicate significant areas based on 300 Monte Carlo simulations with an AR(1) process. Therefore, we see strong evidence for a 52-week period for a large portion of the study period in this FSA.

More importantly, we are interested in the wavelet coherence between FSAs, we chose to use M5A as a comparator within Ontario since it is in the GTA. The choice of a comparator is arbitrary though, and we could have performed this analysis using any FSA as a comparator. Figure 17 highlights areas with high wavelet coherence and also provides visual indicators of the phase angle coherence between the waves. Notice the strong coherence around the dashed horizontal line representing the 52-week period. The arrows indicate the phase difference between waves, an arrow pointing to the right implies the waves are in phase, an arrow pointing up implies the phase difference is positive, an arrow pointing down implies the phase difference is negative and an arrow pointing to the left implies the waves are in anti-phase. Note we are considering these angles modulo  $2\pi$ .

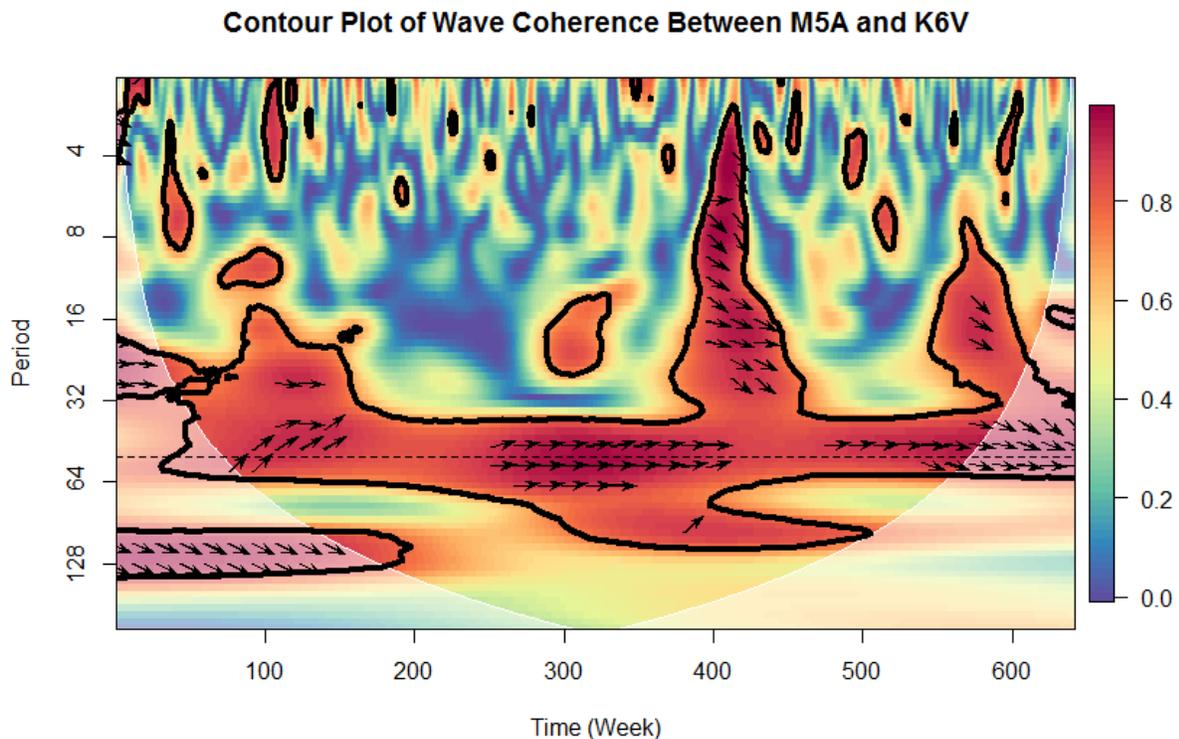


Figure 20: The contour plot of the wavelet coherence between M5A and K6V.

Since we had phase angles corresponding to the different wavelet scales for each time step in our time series, we needed to develop a summary statistic for the wavelet scale of 52 weeks, the median was used since it isn't clear whether the distribution will be symmetric. With a summary

measure for each FSA, we then created a choropleth plot to examine the spatiotemporal trend of influenza in Ontario.

Looking at the plot in Figure 18, we see a general trend of influenza spreading from the northwest to the southeast; this is because both the magnitude and sign of the median phase differences contain information about the temporality between influenza in each FSA compared to M5A. More importantly, negative values indicate that the FSA in question enters the influenza season prior to M5A, while a positive value indicates the opposite. Further, the magnitude corresponds to  $\frac{52}{2\pi} \approx 8.3$  weeks. This implies that a one-unit difference in phase angle correlates with an 8.3-week difference in the start of the respective influenza seasons. However, since the spatial units in the north are large, these may be driving the apparent trend we see. For these reasons, analyses taking into account population densities need to be done as well to verify if the same pattern is seen.

### 3.4 Discussion

Since our data describes influenza-like illness rather than influenza, it is possible that many of our counts are due to respiratory infections other than influenza. This may conclude in us declaring the beginning of influenza season or excess influenza morbidity when there is truly increased morbidity due to another pathogen. Clearly, this is a major impediment that must be considered upon interpretation of our results. Performing laboratory analysis on a sample of patients to determine an underlying percentage of true influenza infection could resolve this issue. Doing so would allow us to more reliably comment on whether influenza is driving the influenza-like illness being observed.

It is clear that the construction of Serfling curves in FSAs of Ontario can lead to useful implementation in both determining the beginning of the flu season and quantifying excess morbidity. More importantly, we have the existing retrospective data to operationalize this method and should look into further developments of the theory. Trends in these curves can allow us to distribute flu vaccine, antivirals and healthcare providers more equitably, in terms of space and time. However, these curves cannot tell us much about the simultaneous spatiotemporal behavior of influenza, as discussed earlier. Further, since these curves are easy to implement, there is huge potential for them be used in FSAs with the capability of monitoring them. Rather than using them retrospectively as we did, it would be wise to employ them prospectively to determine both the start week of the flu season and the weekly excess influenza morbidity.

Wavelet methods are a powerful way to look at the spatiotemporal patterns of influenza across Ontario. However, there are two obvious shortcomings of these methods: firstly, the lack of polyvariate comparisons with associated statistical significance testing. Second, applying these methods to problems with such a large number of geographical units overtly requires multiple comparison corrections.

Ultimately, our research suggests a northwest to southeast diffusion of influenza over the 12 years we had data for. It is unclear what role population density plays in the trend, however, we expect it to play a significant one since influenza's mode of transmission is via droplets. Further,

different types and subtypes of influenza may have unique diffusion patterns that our methods are not able to parse.

With this information, we might consider the timing of vaccination campaigns to begin in the northwest prior to the 35<sup>th</sup> week of the calendar year and continue southeast ahead of the diffusion.

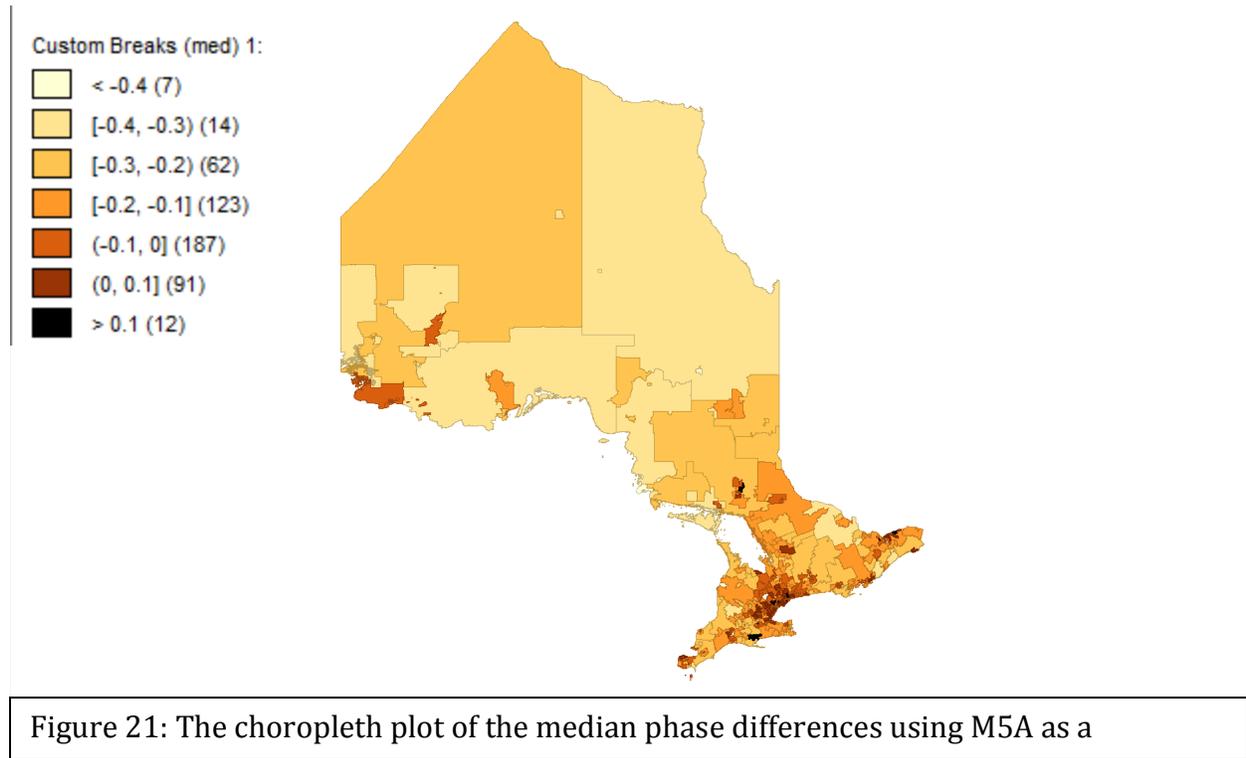


Figure 21: The choropleth plot of the median phase differences using M5A as a

### 3.5 Summary

Using two separate methods, we found a northwest to southeast diffusion of seasonal influenza in Ontario. This suggests that influenza vaccination should begin in the northwest and extend to the southeast, just as the pathogen does. Further, implementing Serfling curves in public health units capable of doing so and monitoring them could prove useful in the fight against seasonal influenza.

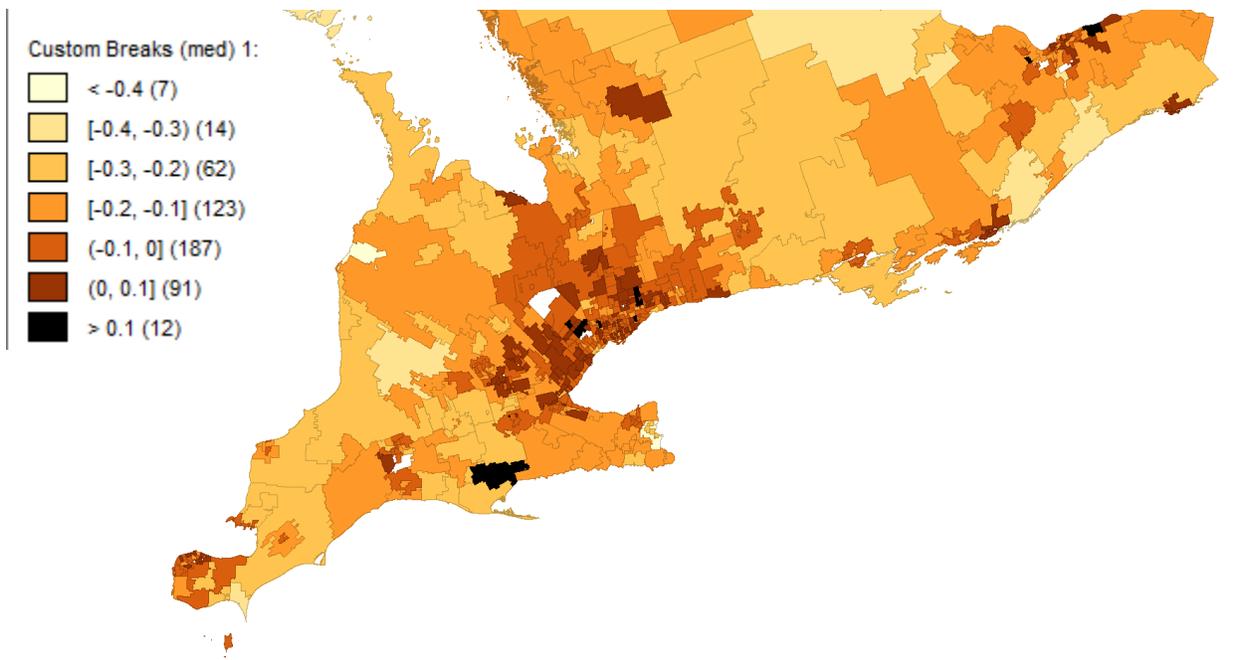


Figure 22: A zoomed in view of figure 18 highlighting the Greater Toronto Area.

## Chapter 4

### Future Work

Further development and refinement of our new cold weather warning system is necessary, as well are formal comparisons in terms of performance relative to the conventional method. No repository containing the dates of historical cold weather warnings exists to our knowledge and so making this comparison retrospectively would not be possible. Finally, we want to consider the use of CuSum methods rather than control chart methods, as they are capable of detecting much smaller shifts in the mean.

For the spatiotemporal analysis of seasonal influenza in Ontario, another available approach is that of Bayesian vector autoregression (BVAR) [11] or spatiotemporal autoregressive integrated moving average (STARIMA) [12]. These methods allow us to take into account both the spatial and temporal dimension of data and to construct models with the ability to make explicit forecasts. Unfortunately, these methods are still not widely available on computers due to their complexity, which is primarily a result of the fact that we are considering both spatial and temporal components.

## References

- [1] ARCHIVED - Environment And Climate Change Canada - Ontario Weather Review - March 2012. *Ec.gc.ca*. 2014. Web. 11 Aug. 2016.
- [2] Anselin L. Local Indicators of Spatial Association-LISA. *Geogr Anal*. 1995;27(2). doi:10.1111/j.1538-4632.1995.tb00338.x.
- [3] Bivand R, Piras G. Comparing Implementations of Estimation Methods for Spatial Econometrics. *JSS J Stat Softw*. 2015;63(18). <http://www.jstatsoft.org/>.
- [4] Cazelles B, Chavez M, Constantin De Magny G, Guégan J-F, Hales S. Time-dependent spectral analysis of epidemiological time-series with wavelets. *J R Soc Interface*. 2007. doi:10.1098/rsif.2007.0212.
- [5] Fricker RD, Hegler BL, Dunfee DA. Comparing syndromic surveillance detection methods: EARS' versus a CUSUM-based methodology. *Stat Med*. 2008. doi:10.1002/sim.3197.
- [6] Gasparrinia A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Stat Med*. 2010. doi:10.1002/sim.3940.
- [7] Hutwagner LC, Thompson WW, Seeman GM, Treadwell T. A Simulation Model for Assessing Aberration Detection Methods used in Public Health Surveillance for Systems with Limited Baselines. *Stat Med*. 2005;24:543-550. doi:10.1002/sim.2034.
- [8] Hutwagner L, Browne T, Seeman GM, Fleischauer AT. Comparing Aberration Detection Methods with Simulated Data. *Emerg Infect Dis*. 2005;11(2).
- [9] Hutwagner L, Thompson W, Seeman GM, Treadwell T. The Bioterrorism Preparedness and Response Early Aberration Reporting System (EARS). *J Urban Heal Bull New York Acad Med*. 2003;80(1).
- [10] Jackson ML, Baer A, Painter I, Duchin J. A simulation study comparing aberration detection algorithms for syndromic surveillance. *BMC Med Inform Decis Mak*. 2007. doi:10.1186/1472-6947-7-6.
- [11] Kamarianakis Y. Spatial-Time Series Modeling: A Review of the Proposed Methodologies. *Reg Econ Appl Laboratory*. 2003;(217217):61801-63671. [www.uiuc.edu/unit/real](http://www.uiuc.edu/unit/real).
- [12] Kamarianakis Y, Prastacos P. Space-time modeling of traffic flow. *Comput Geosci*. 2005. doi:10.1016/j.cageo.2004.05.012.
- [13] Kokoszka P, Mikosch T. The periodogram at the Fourier frequencies. *Stoch Process their Appl*. 2000;86:49-79. [www.elsevier.com/locate/spa](http://www.elsevier.com/locate/spa).
- [14] Kwong JC, Li P, Redelmeier DA (2009) Influenza Morbidity and Mortality in Elderly Patients Receiving Statins: A Cohort Study. *PLoS ONE* 4(11): e8087. doi:10.1371/journal.pone.0008087
- [15] Modeling and Monitoring of Epidemic Phenomena. R package version 1.12.1. <https://CRAN.R-project.org/package=surveillance>
- [16] Moran PAP. Notes on Continuous Stochastic Phenomena. *Biometrika Trust Notes Contin Stoch Phenom Author Biometrika*. 1950;37(12):17-23. <http://www.jstor.org/stable/2332142>.

- [17] Omata K, Takahashi Y. Spatiotemporal Analysis of Influenza Epidemics in Japan. *Proc Int Conf Soc Model Simul*. 2014. doi:10.1007/978-3-319-20591-5\_15.
- [18] Schanzer DL, Sevenhuysen C, Winchester B, Mersereau T. Estimating influenza deaths in Canada, 1992-2009. *PLoS One*. 2013. doi:10.1371/journal.pone.0080481.
- [19] Serfling RE. Methods for Current Statistical Analysis of Excess Pneumonia-influenza Deaths. 1963.
- [20] Thompson WW, Weintraub E, Dhankhar P, et al. Estimates of US influenza-associated deaths made using four different methods. *Influenza Other Respi Viruses*. 2009. doi:10.1111/j.1750-2659.2009.00073.x.
- [21] T.C. Gouhier, A. Grinstead and V. Simko (2016). biwavelet: Conduct univariate and bivariate wavelet analyses (Version 0.20.7). Available from <http://github.com/tgouhier/biwavelet>
- [22] Torrence C, Compo GP. A Practical Guide to Wavelet Analysis. *Bull Am Meteorol Soc*. 1997.
- [23] Wenger JB, Naumova EN, Gething PW. Seasonal Synchronization of Influenza in the United States Older Adult Population. *PLoS One*. 2010;5(4). doi:10.1371/.