QUANTITATIVE RESEARCH, AND ANALYSIS OF BURGLARY AND THEFT IN PASCO COUNTY, FLORIDA

Asuelimen E. Ikhuoria
GIS Department, Pasco County Government

Abstract

This study examined the pattern of distribution of burglary and theft in Pasco County from 2010 and 2012. The methods used to analyze this pattern include spatial autocorrelation, high/low clustering, hot spot analysis, cluster and outlier analysis and CrimeStat's nearest neighbor hierarchical clustering (Nnh). The results showed hot spots around the northwestern region of the county. The nearest neighbor hierarchical clustering Nnh showed clusters in the first, second and third orders. These clusters were located in the western region of the county. This area also referred to the west market area.

1 Introduction

Crime incidents are based on the interaction between victims and offenders. These incidents are measured using a variety of socio-economic and crime opportunity variables, such as population density, economic opportunity and arrest rate and crime records (Wang et al., 2013).

Crime is a major part of our social lives. It's there when we wake up and when we go to bed. Crime over the past couple of years has gradually risen either due to the recent economic down turn that has plagued the country for a decade or due the inability of the police department to constantly keep up with its rise.

The spatial distribution of crime is very important in order to manage the limited resources at the police disposal. It is also important to understand the causes and therefore better enable the police department use their resources in areas that need it the most. The use of statistical analysis in GIS is important in analyzing crime and assists the police in a more effective way.

This project takes a brief look at the spatial distribution of crimes (burglary and theft to be specific) in Pasco County, Florida. These crimes were chosen because they are the most common crimes committed on a daily basis. This project will be look at the data collected between 2010 and 2012. This project will look at hotspots of these crimes and with the results generated provide the necessary recommendation to where and how the police should deploy.

Crime mapping is an important technical function that is part of the modern police enforcement. Crime analyst have routinely mapped crime incidents in order to detect general patterns of crime and to identify and apprehend specific offenders who commit these crimes (Levine, 2006).

The advancement in technologies have enabled agencies to collect enormous amounts of data. The rise in crime and an increase in the number of calls for service have led to a greater need to sort, organize, analyze, and disseminate data in the law enforcement. As a result, criminal justice agencies are turning to GIS software and the latest crime-mapping techniques to deliver data in a more efficient and instructive manner. In addition, using GIS to map crime and criminal behavior eliminates the rampant duplication of efforts among agencies (Stoe et al, 2003).

1.1 Objectives

The objectives to be undertaken in this project include

- Determine and identify the hot spots and cold spots of both the burglary and theft and which census tracts within the county contain these hotspots.

- Compare how these locations vary between 2010 and 2012.

- Observe the spatial autocorrelation between the number of burglary and thefts that occur within a specific census tract

1.2 Study Area

Pasco County is a county located in the mid-western part of Florida. It is part of the Tampa-St. Petersburg-Clearwater metropolitan area. Pasco County was created on June 2nd 1887 from Hernando County and is named after Samuel Pasco, a senator elected to the United States senate (“Pasco County, Florida,” 2015).

The county seat is located in Dade City and the largest city is New Port Richey located in the western section of the county. Cities in the county include New Port Richey, Port Richey, Dade City, San Antonio, and Zephyrhills. The population of the county, according to the 2010 census data, is 464697, a 34.8% increase from the 2000 census figures (“Pasco County, Florida,” 2015).

The total area of the county is 868 sq. miles with 747 sq. miles comprising of land while 122 sq. miles is water. The county includes parks and trails, which includes the Gulf of Mexico. It is also known fit its several nudist beaches and resorts (“Pasco County, Florida,” 2015).

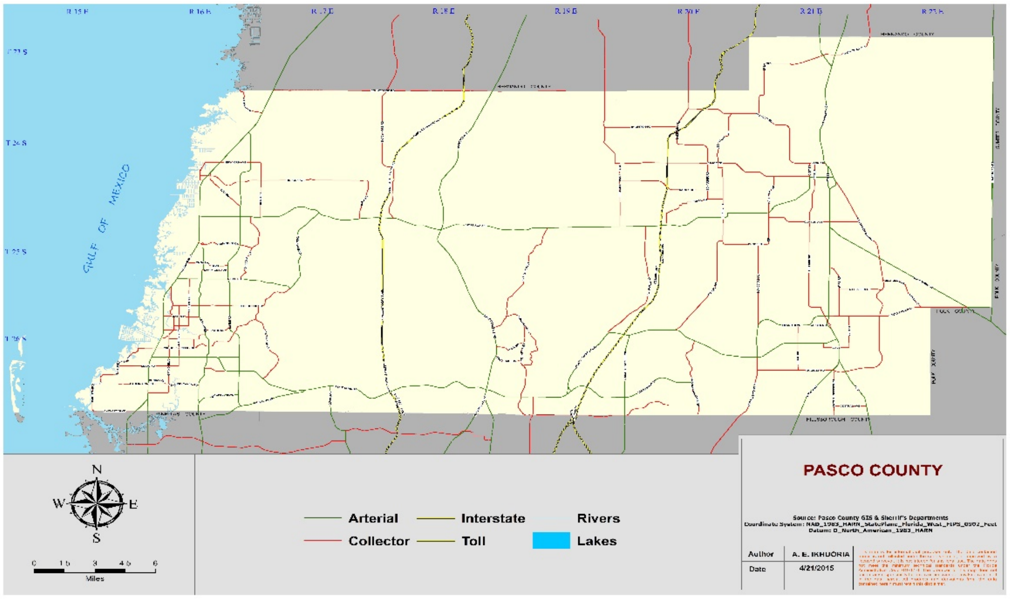

Fig. 1.1: Pasco County

2 Literature Review

The increased availability of Geographic Information Systems (GIS) data and software have contributed immensely to the growth and development of crime mapping among law enforcements both in the private and public sectors (Craglia et al, 2000). Crime mapping according to Cragalia et al (2000) is essentially divided into three main categories: Crime analysis, resource planning and community policing and dispatching. Each of these categories tends to operate in different spheres within a particular place and time (Craglia et al., 2000).

This project focuses on the various reviews of crime analysis, the role the internet plays in displaying the results generated from these analysis for the general public to see and also the police and other law enforcements to observe wherever and whenever they feel like. This section of the project is divided into two major categories: the crime analysis methods and the second section deals with the role and importance of the internet to crime mapping.

One major component to solving crime involves analyzing where crime takes place (Chainey et al, 2008). This is based on the recognition that crime takes place in time and space. Crime does not occur randomly, it does occur more in particular place than in others (Chainey et al., 2008). Chainey et al (2008) also states that hotspot mapping has become a very popular technique used by law enforcers and crime analysts all around the world to visually see and identify various crimes tends to be highest. It also aids decision-making and determines where limited resources can best be useful.

Brundson et al (2007) observed the various methods of visualizing space and time crime patterns. Brundson et al (2007) observed that previous research that focused on crime pace-time patterns focused more on the merits and the limitations of maps and interactive dynamic visualization techniques. They observed the merits of three visualization techniques (map animation, the comap and the isosurface) and also evaluated on their ability to assist in the exploration of space-time patterns of crime disturbance data. They concluded that these techniques had some advantages to crime analysis. Some of the advantages include being able to show patterns from the whole time period under study; it allowed a large degree of interaction; and intuitive.

Kuo et al (2013) noted that crime events, although they are known threats to the general public, can easily be focusing on high impact areas. One way they suggested in his article was to document the various ways to improve or change police or law enforcement dispatch time.

Wang et al (2013) noted that crimes were geographically clustered and as a result, the use of hotspot analysis, to identify and visualize crime, had become wide spread. They also noted that accurately identified crime hotspots can greatly benefit the public by creating accurate threat visualizations, more efficiently allocating police resources and crime prediction. But they pointed out that some crime mapping methods such as point mapping, choropleth mapping and kernel density estimation aggregates the density of a target crime and as a result lead to the loss of information.

Crime Analysis is defined as a set of systematic, analytical processes directed at providing timely and pertinent information relative to crime patterns and trend correlations to assist the operational and administrative personnel in planning the deployment of resources for the prevention and suppression of criminal activities, aiding the investigative process, and increasing apprehensions and the clearance of cases. It supports a number of department functions including patrol deployment, special operations, and tactical units, investigations, planning and research, crime prevention, and administrative services (Johnson, 2000).

Johnson (2000) observed the role of GIS in crime analysis. He noted that maps offer crime analysts graphic representations of crime related issues. An understanding of where and why crimes occur can improve attempts to fight crime. Mapping crime can help police protect citizens more effectively. Johnson (2000) also noted that maps that display the locations where crimes or concentrations of crime have occurred can be used to help direct patrols to places that are most needed. Policy makers in law enforcement use more complex maps to observe trends in criminal activity and has proved invaluable in solving criminal cases.

GIS plays an important role in crime mapping and analysis. Response capabilities often rely on a variety of data from multiple agencies and sources. The ability to access and process information quickly while displaying it in a spatial and visual medium allows agencies to allocate resources quickly and more effectively. In the 'mission-critical' nature of law enforcement, information about the location of a crime, incident, suspect, or victim is often crucial to determine the manner and size of the response(Johnson, 2000). GIS software helps coordinate vast amounts of location-based data from multiple sources. It enables the user to layer the data and view the data most critical to the particular issue or mission. It is used world over by police departments, both large and small, to provide mapping solutions for crime analysis, criminal tracking, traffic safety, community policing, Intranet/Internet mapping, and numerous other tasks(Johnson, 2000).

GIS helps crime officers determine potential crime sites by examining complex seemingly unrelated criteria and displaying them all in a graphical, layered, spatial interface or map. It also helps them map inmate populations, fixtures, and equipment to provide for the safety of inmates by separating gang members, identifying high-risk or potentially violent inmates, and identifying hazardous locations in an area. It reduces the potential for internal violence by providing better command and control. GIS functions, when combined with capabilities of location identification devices such as GPS facilitate tracking the movement of high-risk inmates or at-risk personnel throughout an area. It is more cost-effective for the crime analyst to come up with the information than for patrol officers to do it themselves(Johnson, 2000).

3 Data

3.1 Data Sources

The data required for this project will be obtained from two main sources. The crime data will be obtained from the sheriff's department while the rest of the data will be obtained from the GIS department (http://www.pascocountyfl.net/) in Pasco County. The data obtained from the sheriff's department and the Pasco County GIS department is very accurate and it is constantly being updated by both departments.

The data needed or acquired from sheriff's department and the GIS department include:

Data	Source	Type
2010 burglary and thefts	Sheriff's Department	Points
2012 burglary and thefts	Sheriff's Department	Points
Roads	County GIS Department	Lines
Census Tracts	County GIS Department	Lines
Police Station	County GIS Department	Points

Hospitals	County GIS Department	Points
County Boundary	County GIS Department	Polygon
Rivers	County GIS Department	Line
Lakes	County GIS Department	Polygon

Table 3.1: Data Sources and Types

3.2 Data processing

The crime data will be geocoded using the ESRI geocoder and re-project the geocoded data to conform to the same coordinate system the rest of the data has. A spatial join analysis will be made to determine the number of burglary and thefts within each of the census tracts. The next step is to conduct a spatial and statistical analysis using ESRI ArcGIS 10.2 and the CrimeStat software, on the theft and burglary for both years:

4 Methods

4.1 Spatial Join

This joins the attribute of two layers based on the location of the features in the layers. It appends the attributes of one to another. This creates a permanent link between both layers. The spatial join was used to link the 2010 census tract and the 2010 and 2012 crime data (burglary and theft) together. The outcome of the spatial join was new shapefile polygons showing the total number of burglaries and thefts in 2010 and 2012 respectively within each census tracts ("ArcGIS Help 10.2 - Spatial Join (Analysis)," n.d.).

4.2 Spatial Statistical Analysis

The spatial statistics analyzes the spatial distributions, patterns, processes and relationships of features. It was developed specifically for the use of geographic data. The spatial statistics is used to identify spatial clusters, outliers, access overall patterns of dispersion and explore spatial relationship ("ArcGIS Help 10.2 - An overview of the Spatial Statistics toolbox," n.d.).

To determine the spatial clusters, outliers and relationship between the crime and population distribution in Pasco County, four main processes were used, they are the hot spot analysis (Getis-Ord Gi*), Cluster and Outlier Analysis (Anselin Moran's I), Spatial Autocorrelation (Global Moran's I) and High/Low Clustering (Getis-Ord General G).

4.2.1 Spatial Autocorrelation

This measures the autocorrelation spatially based on where the features are located and what attribute values the analyst determines using the Global Moran's I statistics. The spatial autocorrelation decides if the pattern or the distribution of crime spatially clustered, dispersed or random ("ArcGIS Help 10.2 - Spatial Autocorrelation (Global Moran's I) (Spatial Statistics)," n.d.). The Global Moran's I was analyzed using ArcGIS for the number of Burglaries and Thefts for both 2010 and 2012. The shapefile that was analyzed was the previously generated spatially joined data. This joined shapefile contains the total population for each tract and a "COUNT" field where the total number of incidents per tract is shown.

The parameters for calculating Global Moran's I include:

Input Feature Class………………………….. 2010 and 2012 Crime Data (four analyses were calculated for this)
Input Field…………………………………… Count
Conceptualizing of Spatial Relationships........ Fixed Distance Band
Distance Method…………………………….. Euclidean Distance
Distance Band…………………………......... None
Weight Matrix File......................................…… None

The analysis returned five values: the Moran's I index (observed), the expected index, variance, z-score, and p-value. The results and interpretation are further discussed in the results section of the project.

4.2.2 **High/Low Clustering (Getis-Ord General G)**

The High/Low Clustering measures the degree of clustering for either high values or low values using the Getis-Ord General G statistics. The High/Low Clustering (Getis-Ord General G) tool is most appropriate when you have a fairly even distribution of values and are looking for unexpected spatial spikes of high values. Unfortunately, when both the high and low values cluster, they tend to cancel each other out ("ArcGIS Help 10.2 - High/Low Clustering (Getis-Ord General G) (Spatial Statistics)," n.d.).

The General G was analyzed using ArcGIS for the number of Burglaries and Thefts for both 2010 and 2012. The shapefile that was analyzed was the previously generated spatially joined data. This joined shapefile contains the total population for each tract and a "COUNT" field where the total number of incidents per tract is shown.

The parameters for calculating Global Moran's I include:

Input Feature Class………………………….. 2010 and 2012 Crime Data (four analyses were calculated for this)
Input Field…………………………………… Count
Conceptualizing of Spatial Relationships........ Fixed Distance Band
Distance Method…………………………….. Euclidean Distance
Distance Band…………………………......... None
Weight Matrix File......................................…… None

The analysis, like the spatial autocorrelation, returned five values: the Moran's I index (observed), the expected index, variance, z-score, and p-value. The results and interpretation are further discussed in the results section of the project.

4.2.3 **Hot Spot Analysis (Getis-Ord Gi*)**

The hot Spot Analysis identifies the areas that have significantly high clusters and areas with significantly low clusters. The result creates a new feature class with a z- score, p-value and confidence level bin for each feature in the input feature class. The z-score and p-values are indicators of statistical significance which enables the analyst to decide whether to accept or reject the null hypothesis. They also indicate whether the observed spatial clustering of high or low values is more pronounced than one would expect in a random distribution of those same values. A high z-score and small p-value for a feature indicates a spatial clustering of high values. A low negative z-score and small p-value indicates a spatial clustering of low values. The higher (or lower) the z-score, the more intense the clustering. A z-score near zero indicates no apparent spatial clustering ("ArcGIS Help 10.2 - Hot Spot Analysis (Getis-Ord Gi*) (Spatial Statistics)," n.d.).

The hot spot for Pasco County was calculated using the ArcGIS Hot Spot Analysis tool. Using the joined shapefile generated earlier, the hot spot and cold spots for burglaries and thefts between 2010 and 2012 were determined. The field containing the total number (count) of incidents per tract was used for this analysis

The parameters for calculating the Hot Spot Analysis is also similar to that used for calculating both the spatial autocorrelation and the High/Low Clustering. They include:

Input Feature Class………………………….. 2010 and 2012 Crime Data (four analyses were calculated for this)
Input Field…………………………………… Count
Conceptualizing of Spatial Relationships........ Fixed Distance Band
Distance Method…………………………….. Euclidean Distance
Distance Band…………………………......... None
Weight Matrix File.....................................…… None
Output Feature Class………………………… each analysis was given an output name in order to differentiate between the crime type and the year

The results from this analysis are further discussed the results section of this project. On major difference between the hot spot analysis and the high/low clustering is that while high/low clustering produces a single value for the z-score, p-value and variance, the hot spot analysis creates a whole new shapefile with each feature having its own z-score and p-value.

4.2.4 **Cluster and Outlier Analysis (Anselin Local Moran's I)**

The Cluster and Outlier Analysis tool identifies spatial clusters of features with high or low values. The tool also identifies spatial outliers. To do this, the tool calculates a local Moran's I value, a z-score, a p-value, and a code representing the cluster type for each statistically significant feature. The z-scores and p-values represent the statistical significance of the computed index values. A positive value for I specifies if a feature has neighboring features with similarly high or low attribute values; this feature is part of a cluster. A negative value for I indicates that a feature has neighboring features with dissimilar values; this feature is an outlier. In either instance, the p-value for the feature must be small enough for the cluster or outlier to be considered statistically significant. The cluster/outlier type field distinguishes between a statistically significant cluster of high values (HH), cluster of low values (LL), outlier in which a high value is surrounded primarily by low values (HL), and outlier in which a low value is surrounded primarily by high values (LH) ("ArcGIS Help 10.2 - Cluster and Outlier Analysis (Anselin Local Moran's I) (Spatial Statistics)," n.d.).

Using this tool, a Cluster and Outlier Analysis was conducted on the crime data (spatially joined) for both 2010 and 2012. The parameters for calculating the hot spot analysis was also used to calculate this also. They include:

Input Feature Class………………………….. 2010 and 2012 Crime Data (four analyses were calculated for this)
Input Field…………………………………… Count
Conceptualizing of Spatial Relationships........ Fixed Distance Band
Distance Method…………………………….. Euclidean Distance
Distance Band…………………………......... None
Weight Matrix File.....................................…… None
Output Feature Class………………………… each analysis was given an output name in order to differentiate between the crime type and the year

The findings and interpretation are further discussed in the results section.

4.3 Nearest Neighbor Hierarchical Analysis (Nnh)

This method is derived from the CrimeStat software. CrimeStat is a stand-alone Windows spatial statistics program for the analysis of crime incident locations that can interface with most desktop GIS programs. The purpose is to provide supplemental statistical tools to aid law enforcement agencies and criminal justice researchers in their crime mapping efforts (Levine, 2006)

The nearest neighbor hierarchical clustering (Nnh) identifies groups of incidents that are spatially close. It is a hierarchical clustering routine that clusters points together on the basis of several criteria. The clustering is repeated until either all points are grouped into a single cluster (Levine, 2013).

The nearest neighbor hierarchical clustering uses a method that defines a threshold distance and compares the threshold to the distances for all pairs of points. Only points that are closer to one or more other points than the threshold distance are selected for clustering. In addition, the user can specify a minimum number of points to be included in a cluster. Only points that fit both criteria - closer than the threshold and belonging to a group having the minimum number of points, are clustered at the first level (first-order clusters). The routine conducts subsequent clustering to produce a hierarchy of clusters. The first-order clusters are themselves clustered into second-order clusters. Again, only clusters that are spatially closer than a threshold distance (calculated anew for the second level) are included. The second-order clusters, in turn, are clustered into third-order clusters, and this re-clustering process is continued until either all clusters converge into a single cluster (Levine, 2013).

The hotspot analysis was calculated using the CrimeStat's Nearest Neighbor Hierarchal Cluster method. This was achieved by using the originally geocoded point data for the 2010 and 2012 crime data. While the input data changed to reflect each crime data for each year, every other parameters used for the analysis remained the same. These parameters are shown in the results section of the project

5 Results and Discussion

5.1 ArcGIS Statistical Analysis

5.1.1 Spatial Autocorrelation (Moran's I)

BURGLARY	OBSERVED INDEX	EXPECTED INDEX	VARIANCE	Z-SCORE
2010	0.041548	-0.007519	0.000378	2.522765
2012	0.180703	-0.007519	0.000385	9.591119

Table 5.1: *2010 and 2012 Spatial Autocorrelation for Burglary in Pasco County*

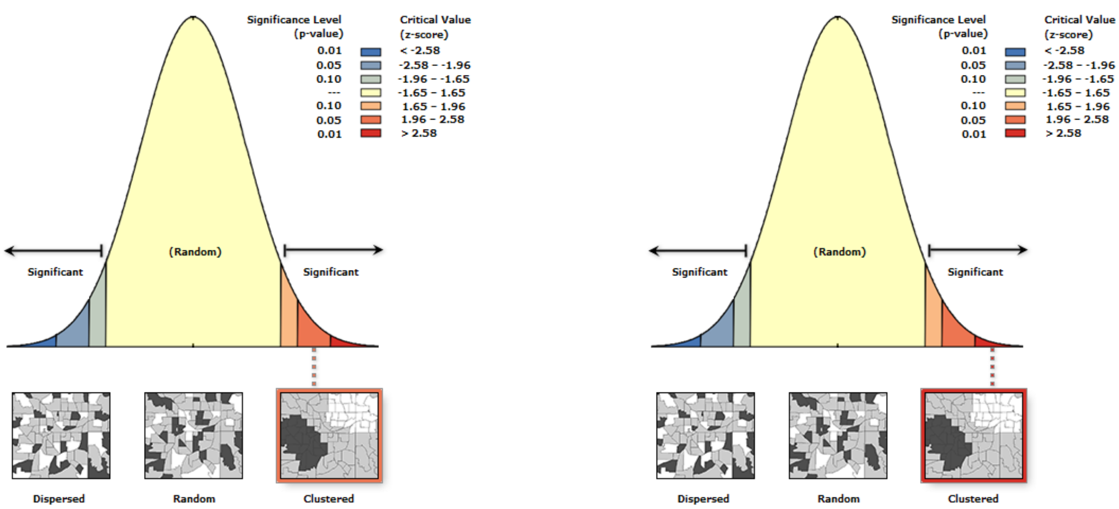

Fig 5.1: Moran's I Index for burglary in 2010 and 2012

The results from the Moran's I index analysis shows that burglary between 2010 and 2012 are highly clustered. The z-score for both years are high although 2012 is higher than 2010. The z-score for 2010 burglary shows that there is less than 5% chance that the clustered pattern could be random while the 2012 burglary shows that there is less than1% chance that the clustered pattern could be random. This indicates that burglary for both years were highly clustered and concentrated in a particular location. This is further illustrated in the normal curve or distribution shown in the figures above.

THEFT	OBSERVED INDEX	EXPECTED INDEX	VARIANCE	Z-SCORE
2010	0.17031	-0.007519	0.000367	9.277961
2012	0.124863	-0.007519	0.000365	6.929155

Table 5.2: *2010 and 2012 Spatial Autocorrelation for Theft in Pasco County*

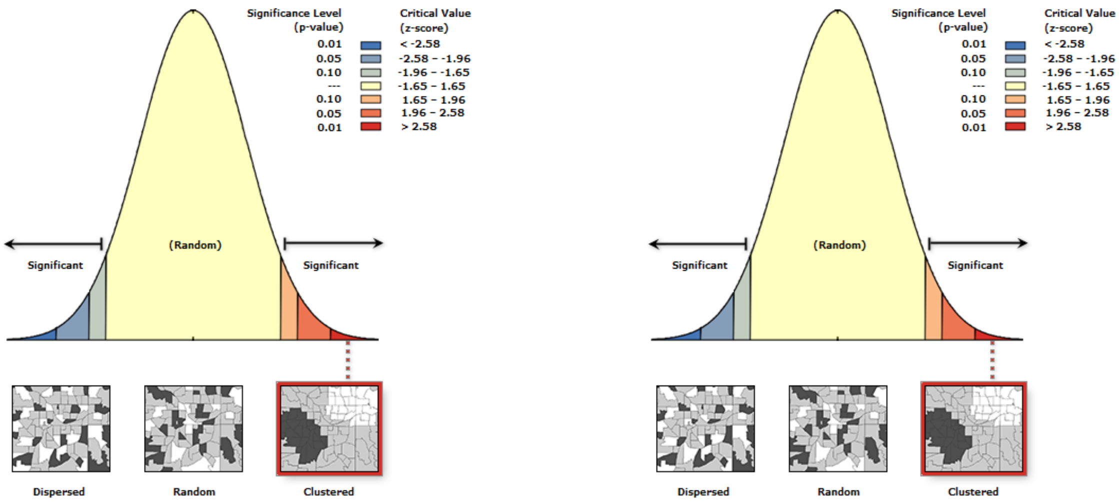

Fig 5.2: Moran's I Index for theft in 2010 and 2012

The results from analyzing the thefts from 2010 and 2012 shows that both years were spatially correlated and highly clustered. The z-score for both years are extremely with 2012 slightly lower than 2010. The z-score for both years showed that there was less than 1% chance that the clustered pattern could random leading to the rejection of the null hypothesis. This indicates that theft just like burglary was clustered for both years. This result is further illustrated in the normal curve or distribution in the figures above.

5.1.2 **High/Low Clustering (Getis-Ord General G)**

BURGLARY	OBSERVED INDEX	EXPECTED INDEX	VARIANCE	Z-SCORE
2010	0.203383	0.198968	0.000127	0.391125
2012	0.233497	0.198968	0.000125	3.091367

Table 5.3: *2010 and 2012 High/Low Clustering for Burglary in Pasco County*

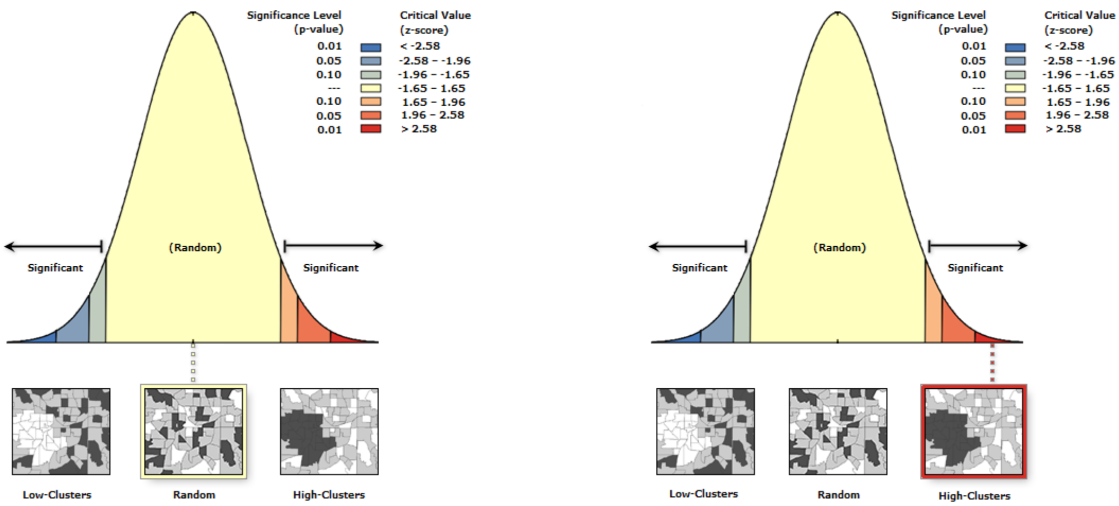

Fig 5.3: General G Index for burglary in 2010 and 2012

The table above shows results from the General G analysis. For 2010, the result indicates that given that the z-score is 0.39, the pattern seem to be random. This is a low z-score, therefore we have to accept the null hypothesis that the high cluster values does appear to be random. The 2012 burglary shows a different result. It shows a high z-score indicating that there is a 1% chance that the high-clustered pattern is not random. The results show that there is no concentration of high or low values of burglary in 2010 but the same cannot be said of 2012. The normal distribution curve further illustrates this result.

THEFT	OBSERVED INDEX	EXPECTED INDEX	VARIANCE	Z-SCORE
2010	0.268828	0.198968	0.00027	4.252543
2012	0.241776	0.198968	0.000259	2.657416

Table 5.4: *2010 and 2012 High/Low Clustering for Theft in Pasco County*

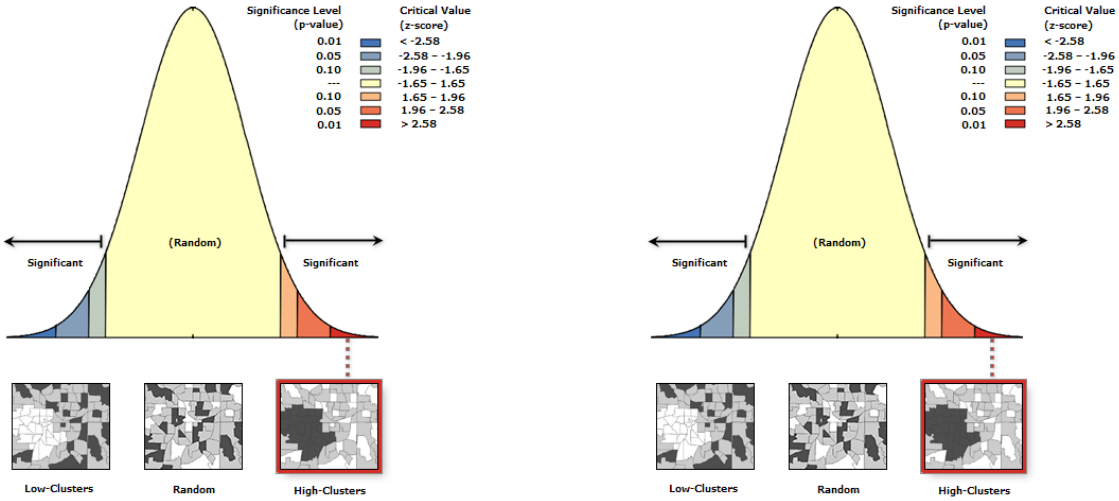

Fig 5.4: General G Index for theft in 2010 and 2012

The result from the General G analysis shows a high-clustered pattern of theft in 2010 and 2012. With a z-score of 4.25 and 2.66, it was observed that there was less than 1% chance that the high-clustered pattern could be the result of random chance. In this case the null hypothesis was rejected for both years due to a high z-score value. This result is also illustrated using the normal distribution curve shown in the figures above.

5.1.3 **Hotspot (General G)**

Burglary

The results of the local G statisitcs of burglary in 2010 and 2012 are shown in figure 5.5. From the maps below it can be observed that the hot spot in both years were concentrated in the northwestern region of the county. In 2010 the hot spots were relatively fewer compared to that in 2012. From the analysis, the 19 out of the 133 census tracts in the county were recoded as hotspots, 7 were recoded as cold spots while the rest were not significant. In 2012 however, 37 out of the 133 tracts were recoded as hot spots, 24 were recoded as cold spots and the rest were not significant. This showed a major difference between both years. Despite a reduction in the overall rate and number of burglary within the county, it could be observed it had spread out to other census tracts in 2012. One reason for the spread of burglary could be due to the economic hardship and loss of jobs for many middle and low income earners. Another likely reason could be that as the police cracked down more on the hot spots in 2010, these burglars spread out to other places of opportunity.

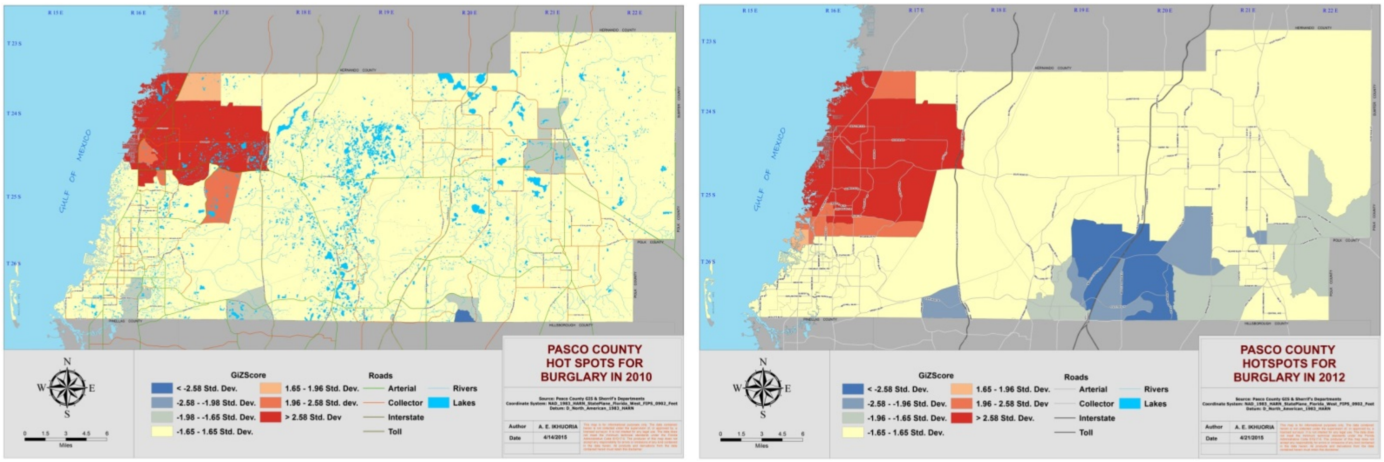

Fig. 5.5: Hot Spots of Burglary in 2010 and 2012

Theft

The results of the hot spot analysis of theft in 2010 and 2012 are shown in figure 5.6. From the maps below it can be observed that all the hot spots were concentrated in the northwestern region of the county in both years, while the cold spots were concentrated in the southeastern region of the county. From the map it can also be observe the almost all the hotspots and cold spots remained the same for both years. The only difference was the census tracts at the lower section of the hot spots. These tracts are 314.05, 314.08, 314.06, 314.01, and 314.07. There were also some slight change in the cold spots between 2010 and 2012. These tracts changed from cold spot to becoming not significant. They are census tract 321.08. In 2010 there were 41 census tracts with hot spots and 24 cold spot census tracts. In 2012 there were 35 hot spot census tracts and 21 cold spot census tracts. Although there was also a slight drop in the in the rate of theft between 2010 and 2012, most of the census tracts retained they hot spots and cold spot over this two year period.

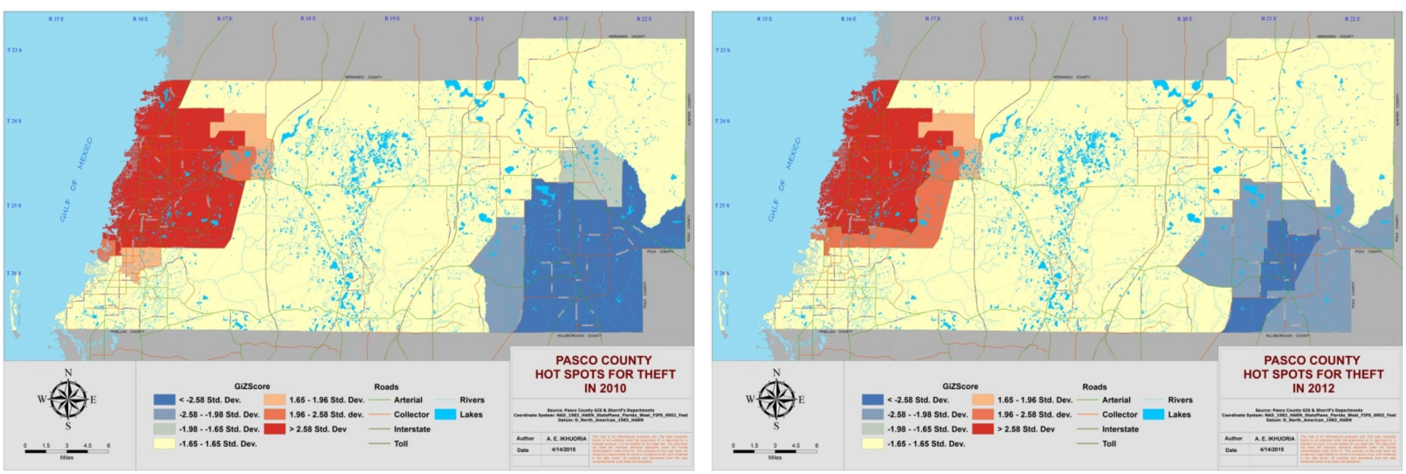

Fig. 5.6: Hot Spots of theft in 2010 and 2012

5.1.4 **Clusters and Outliers (Moran's I)**

Burglary

The results from the cluster and outlier analysis for 2010 and 2012 are shown in figure 5.7. From the maps below it can be observed that the results are just as similar to that obtained from the hot spot analysis. From the results it can be observed that the census tracts with high index value are located on the northwestern region of the county, while the rest such as the outliers are not concentration in a particular region. In 2010 it was observed that of the 133 census tract within the county only 8 had high burglary cluster values. There were 3 high-low outliers, and the rest were recorded as not significant. No low-high outlier or low clusters were recorded for that year. In 2012 however, there were more concentration of high cluster index in the northwestern region of the county. The result also showed that of the 133 tracts within the county, 22 census tracts with high cluster indexes, 5 high-low outliers, 4 low-high outliers and 7 low clusters. This was a big jump from the 2010 despite a reduction in burglary rate in the county.

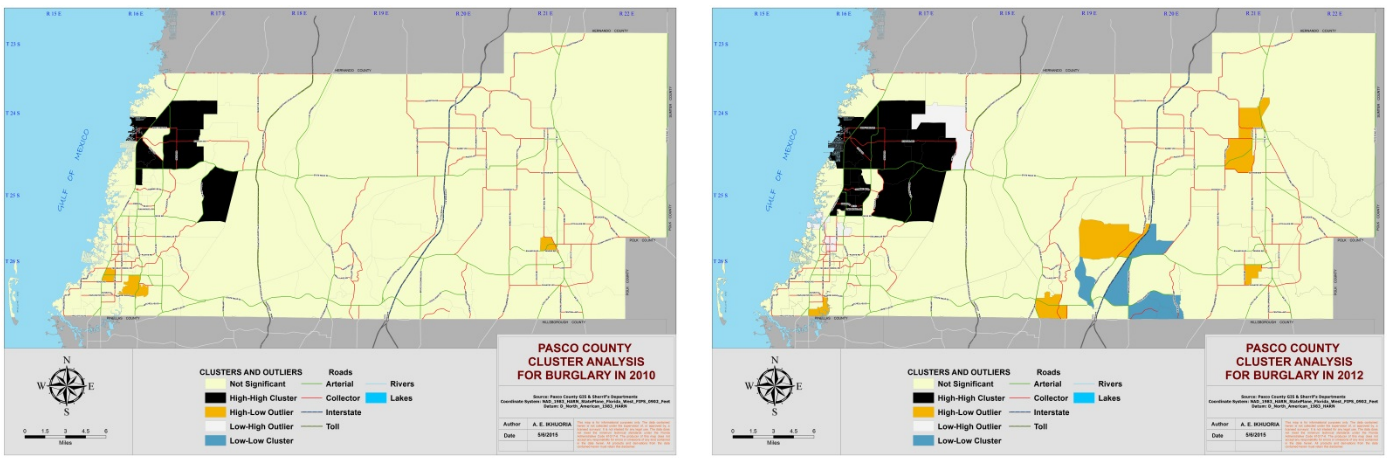

Fig. 5.7: Clusters and Outliers of Burglary in 2010 and 2012

Theft

The results from this analysis are show in the maps below. From the maps it can be observed that the high cluster index or values are concentrated in the northwestern region of the county for both years. In 2010, there were 17 tracts with high cluster values, I high-low outlier and low-high outlier respectively, and 9 tracts with low cluster values. The census tracts with low clusters were located at the southeastern region of the county. The 2012 results was also similar to that of the 2010 results. Of the 133 census tracts, 15 of them were recorded with high cluster values, 1 census tract has a high-low outlier, 2 tracts had low-high outliers and 6 tracts had low cluster values. This was the only analysis that mirrored the drop in the rate of theft within the county.

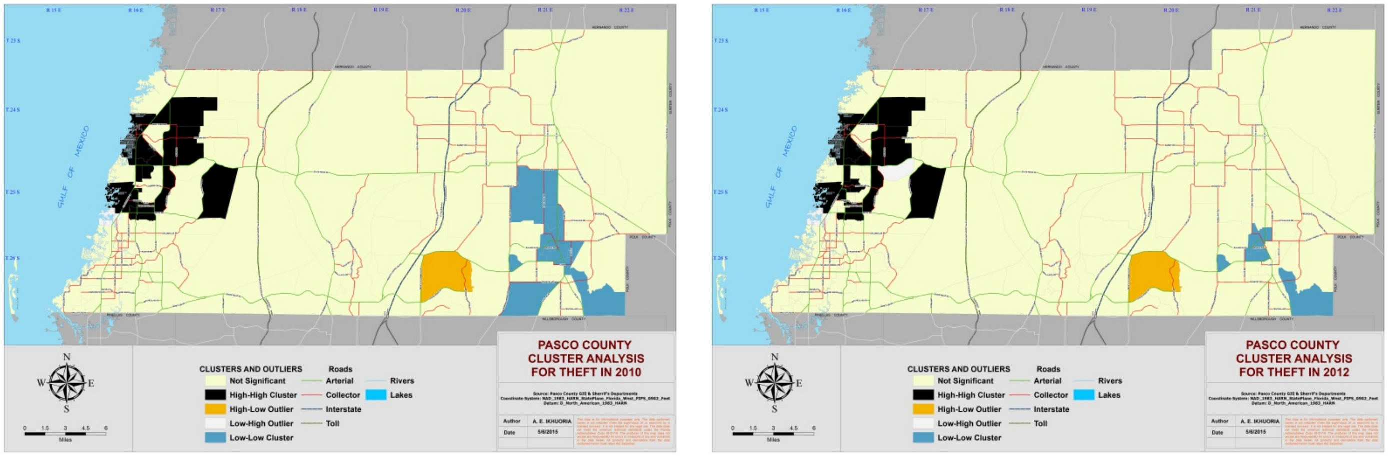

Fig. 5.8: Clusters and Outliers of Theft in 2010 and 2012

5.2 CrimeStat

5.2.1 Nearest Neighbor Hierarchical Clusters (Nnh) Burglary

The results of the analysis are shown in the summary and maps below. The results from the analysis showed that in 2010, 130 clusters were found. Of the 130 clusters 117 were first order ellipses, 11 were second order ellipses and 3 were third order ellipses. Almost all the clusters were located in the western region of the county between the Gulf of Mexico and Suncoast highway. The third order ellipse located at the north-western section of the county covered a larger area compared to the second which only covers a smaller section. Both third order ellipses were oriented towards the central section of the western region. One of the second order ellipses was located in the City of Zephyrhills. The cluster or ellipse with the highest number of burglary incidents was also located in the City of Zepyrhills with a total number of 61 burglary incidents.

The 2012 burglary results on the other hand had a total of 121 clusters or ellipses. Of the 121 clusters, 108 were first order ellipses, 11 were second order ellipses and 2 were third order ellipses. Over ninety-five percent of the clusters were located in the western region of the county between the Gulf of Mexico and Suncoast highway. Some of the first order ellipses were found in the southern and south-eastern region of the county. Two second order ellipses were located in the City of Zephyrhills. The third order ellipses are slightly bigger than those in the 2010 burglary analysis. The cluster or ellipse with the highest number of burglary incidents was located in the City of Zephyrhills with a total number of 28 burglary incidents.

Nearest Neighbor Hierarchical Clustering: 2010 Burglary

Sample size..................: 6135
Likelihood of grouping pair of points by chance...: 0.50000 (50.000%)
Z-value for confidence interval...................: 0.000
Measurement type.............: Direct
Output units.................: Miles, Square Miles, Points per Square Miles
Standard Deviations ....: 1.0
Clusters found...............: 130
Simulation runs..............: 0

Nearest Neighbor Hierarchical Clustering: 2012 Burglary

Sample size..................: 5203
Likelihood of grouping pair of points by chance...: 0.50000 (50.000%)
Z-value for confidence interval...................: 0.000
Measurement type.............: Direct
Output units.................: Miles, Square Miles, Points per Square Miles
Standard Deviations ....: 1.0
Clusters found...............: 121
Simulation runs..............: 0

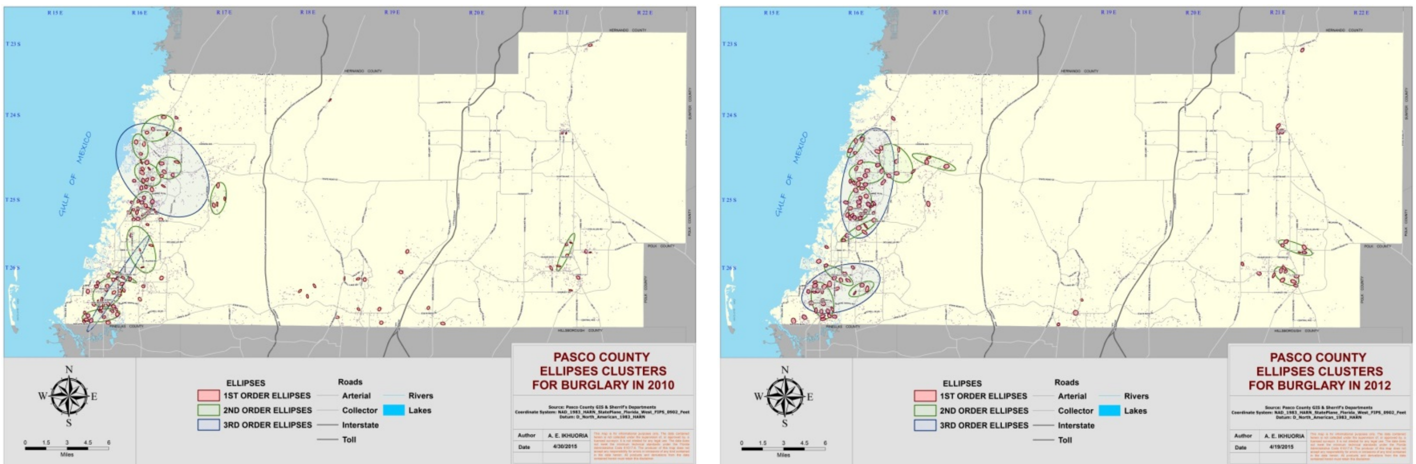

Fig. 5.9: 2010 and 2012 First, Second and Third Order Ellipses of Burglary in Pasco County

Theft

The results of the analysis are shown in the summary and maps below. The results from the analysis showed that in 2010, over ninety-five percent of the county were located in the western region of the county. There were a total of 123 clusters or ellipses and of these 110 were first order ellipses, 11 were second order and 2 were third orders. The results further showed that the cluster with the largest number of theft was located in the Hudson area with a total number of 262 incidents. This was by far the largest number of incidents ever recorded in the project. The coverage area for the second and third ellipses were much bigger than that generated for the burglary. The third order ellipses faced a southwest-northwest orientation.

The 2012 analysis on the other hand showed that a total of 104 clusters was found. This was a reduction of 19 clusters from the 2010 analysis. 95 clusters of the total 104 clusters were first order ellipses, 8 were second order and one was a third order. Over ninety-five percent of all the clusters were located in the western region of the county. The ellipses especially the third order was much narrower ad longer than those generated for the 2010 crime analysis. The third order ellipses has a north-south orientation. The cluster with the highest number of theft incidents was located around Hudson with a total number of 180 incidents.

Nearest Neighbor Hierarchical Clustering: 2010 Theft

Sample size..................: 6251
Likelihood of grouping pair of points by chance...: 0.50000 (50.000%)
Z-value for confidence interval...................: 0.000
Measurement type.............: Direct
Output units.................: Miles, Square Miles, Points per Square Miles
Standard Deviations ....: 1.0
Clusters found...............: 123
Simulation runs..............: 0

Nearest Neighbor Hierarchical Clustering: 2012 Theft

Sample size..................: 5743
Likelihood of grouping pair of points by chance...: 0.50000 (50.000%)
Z-value for confidence interval...................: 0.000
Measurement type.............: Direct
Output units.................: Miles, Square Miles, Points per Square Miles
Standard Deviations ....: 1.0
Clusters found...............: 104
Simulation runs..............: 0

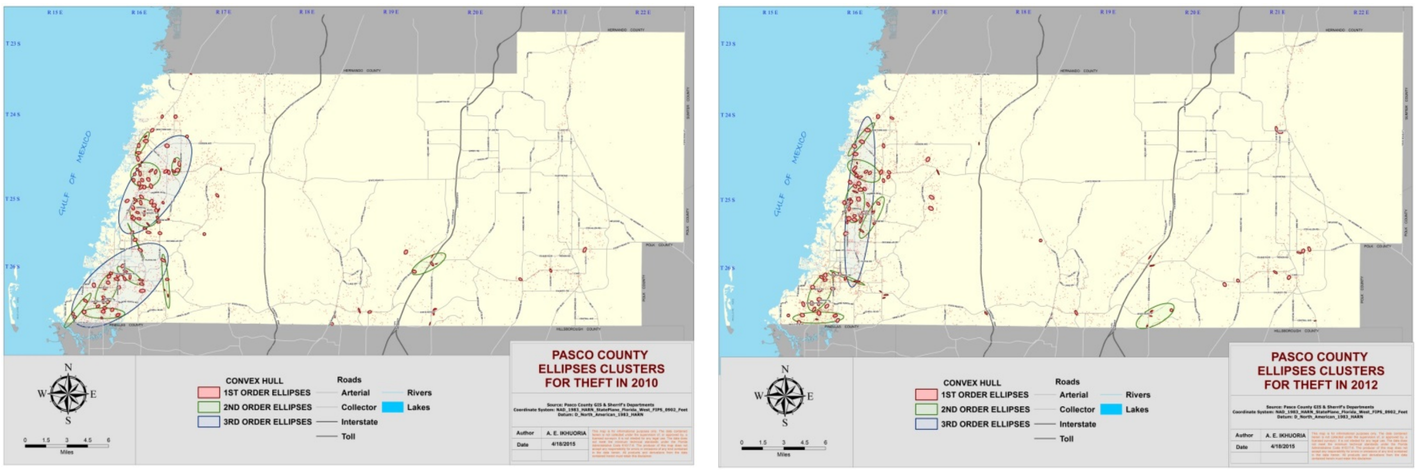

Fig. 5.9: 2010 and 2012 First, Second and Third Order Ellipses of Theft in Pasco County

6 Conclusion

This project observed the spatial distribution of burglary and theft in Pasco County from 2010 to 2012. The study used different methods to compare the patterns of distribution of burglary and theft. The methods used for this study include spatial autocorrelation, high/low clustering, hot spot analysis, cluster and outlier analysis and crimestat's nearest neighbor hierarchical clustering analysis.

The spatial autocorrelation measured clustered, dispersed or random distribution of both burglary and theft. The results showed that both burglary and theft were highly clustered in 2010 and 2012.

The high/low clustered measured the degree of clustering. All results showed a high degree of cluster of both crimes in both years except for burglary in 2010 which indicated that it was random.

The hot spot analysis identified areas within the county that were significantly clustered using the z-score. The results showed hot spots in the northwestern region of the county and cold spots in the southern and southeastern region of the county. This was the same for both crimes in 2010 and 2012.

The clusters and outliers identified the clusters with high values and low values. The results showed the census tracts with high clusters were located in the northwestern region of the county. A few low clusters and outliers were also observed in tracts around the county

The Nearest Neighbor Hierarchical Clustered (Nnh) identified crime (burglary and theft) that were spatially close. The crime incidents were grouped into clusters. The results showed three orders of ellipses. These clusters/ellipses were concentrated in the western region of the county.

The results shows the vast concentration of crime is located in western region of the county. This region is also the urban section of the county, it’s highly populated and this was the earliest place that was settled in. A majority of the middle and high income earners also live in this section of the county and most of the businesses are found there. This has created an atmosphere that would attract crime rate.

7 Recommendation

More detailed study of the impact of crime on population growth is needed. This would go a long way to observe the influx people in the county and how this increase affect the county both positively and negatively. A web application is needed to show the public where the hot spots and cold spots are located in the county. This would go a long way to increase the public trust in the police force. The police department needs to concentrate more police in the areas of hot spots

8 References

ArcGIS Help 10.2 - An overview of the Spatial Statistics toolbox. (n.d.). Retrieved May 17, 2015, from http://resources.arcgis.com/en/help/main/10.2/index.html#/An_overview_of_the_Spatial_Statistics_toolbox/005p00000002000000/

ArcGIS Help 10.2 - Cluster and Outlier Analysis (Anselin Local Moran’s I) (Spatial Statistics). (n.d.). Retrieved May 17, 2015, from http://resources.arcgis.com/en/help/main/10.2/index.html#//005p0000000z000000

ArcGIS Help 10.2 - High/Low Clustering (Getis-Ord General G) (Spatial Statistics). (n.d.). Retrieved May 17, 2015, from http://resources.arcgis.com/en/help/main/10.2/index.html#/High_Low_Clustering_Getis_Ord_General_G/005p00000009000000/

ArcGIS Help 10.2 - Hot Spot Analysis (Getis-Ord Gi*) (Spatial Statistics). (n.d.). Retrieved May 17, 2015, from http://resources.arcgis.com/en/help/main/10.2/index.html#/Hot_Spot_Analysis_Getis_Ord_Gi/005p00000010000000/

ArcGIS Help 10.2 - Spatial Autocorrelation (Global Moran's I) (Spatial Statistics). (n.d.). Retrieved May 17, 2015, from http://resources.arcgis.com/en/help/main/10.2/index.html#/Spatial_Autocorrelation_Global_Moran_s_I/005p0000000n000000/

ArcGIS Help 10.2 - Spatial Join (Analysis). (n.d.). Retrieved May 17, 2015, from http://resources.arcgis.com/en/help/main/10.2/index.html#//00080000000q000000

Brunsdon, C., Corcoran, J., Higgs, G., & Extracting Information from Spatial Datasets. (2007). Visualising space and time in crime patterns: A comparison of methods. *Computers, Environment and Urban Systems*, *31*(1), 52–75.

Chainey, S., Tompson, L., & Uhlig, S. (2008). The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime. *Security Journal*, *21*(1-2), 4–28. http://doi.org/http://dx.doi.org/10.1057/palgrave.sj.8350066

Craglia, M., Haining, R., & Wiles, P. (2000). A Comparative Evaluation of Approaches to Urban Crime Pattern Analysis. *Urban Studies*, *37*(4), 711–729. http://doi.org/10.1080/00420980050003982

Johnson, C. P. (2000). Crime mapping and analysis using GIS. In *Geomatics Conference on Geomatics in Electronic Governance, January*. Retrieved from http://fac.ksu.edu.sa/sites/default/files/crim_mapping.pdf

Kuo, P.-F., Lord, D., & Walden, T. D. (2013). Using geographical information systems to organize police patrol routes effectively by grouping hotspots of crash and crime data. *Journal of Transport Geography*, *30*, 138–148. http://doi.org/10.1016/j.jtrangeo.2013.04.006

Levine, N. (2006). Crime Mapping and the Crimestat Program. *Geographical Analysis*, *38*(1), 41–56. http://doi.org/10.1111/j.0016-7363.2005.00673.x

Levine, N. (2013). *Introduction to CrimeStat IV (CrimeStat IV: A Spatial Statistics Program for the Analysis of Crime Incident Locations, Version 4.0)*. United States of America.

Pasco County, Florida. (2015, March 5). In *Wikipedia, the free encyclopedia*. Retrieved from http://en.wikipedia.org/w/index.php?title=Pasco_County,_Florida&oldid=650037306

Stoe, D. A., United States, Office of Justice Programs, & Office for Victims of Crime. (2003). *Using geographic information systems to map crime victim services a guide for state Victims of Crime Act administrators and victim service providers*. Washington, DC: U.S. Dept. of Justice, Office of Justice Programs, Office for Victims of Crime. Retrieved from http://purl.access.gpo.gov/GPO/LPS27208

Wang, D., Ding, W., Lo, H., Morabito, M., Chen, P., Salazar, J., & Stepinski, T. (2013). Understanding the spatial distribution of crime based on its related variables using geospatial discriminative patterns. *Computers, Environment and Urban Systems*, *39*, 93–106. http://doi.org/10.1016/j.compenvurbsys.2013.01.008