Spatial Analysis Leveraging Geographic Information System and Kernel Density Estimation for Crime Prediction

Roger S. Mission*

Abstract: This study explored the potential of spatial analysis techniques to predict crime hotspots and coldspots within the province of Antique, Philippines. The unique blend of geographical features (coastal and mountainous regions) and socio-economic characteristics in Antique presented a compelling opportunity for the research. Crime data from 2018 to 2023, encompassing direct assaults, homicides, murders, and parricides, was acquired from the Police Regional Office 6. Using Quantum Geographic Information System (QGIS) software and Kernel Density Estimation (KDE) algorithms, the researchers analyzed these crime records. The findings highlighted a positive trend – a decrease in direct assaults over the study period. However, homicide data presented a more concerning picture, with specific locations consistently reporting higher rates. While murder cases showed a decline, parricide incidents remained relatively low, with occasional fluctuations. In conclusion, this data-driven approach, utilizing advanced crime analysis tools, offers valuable insights for policymakers in Antique. By investing in such technology and adopting data-driven strategies, policymakers can work towards creating a safer and more secure future for their communities.

Keywords: Kernel density estimation algorithm, Mapping, Crime analysis, Antique, QGIS

1. Introduction

In response to the escalating societal challenge of crime, criminology has become increasingly reliant on sophisticated research methodologies to gather and analyze crime data. The recent surge in criminal activities underscores the pressing need for law enforcement agencies to proactively manage and mitigate unlawful behavior. Recognizing the imperative for advanced methods to combat crime effectively, this study integrates Information Technology (IT) into the realm of criminological research. By combining clustering techniques and pattern recognition with IT solutions, the study aims to extract valuable insights from criminal activities, focusing on predictive hotspot mapping methods to identify and forecast potential crime locations. This integration of IT enhances the capacity for data collection, storage, and analysis, enabling a more comprehensive approach to understanding and managing crime.

* College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: rsmission@antiquespride.edu.ph

Received [January 31, 2024]; Revised [March 5, 2024]; Accepted [April 4, 2024]
The utilization of information technology in this study highlights the transformative role of information technology (IT) in contemporary law enforcement. Beyond just data management, IT facilitates real-time crime monitoring, information sharing across agencies, the development of predictive models [1], information sharing among law enforcement agencies, and the development of predictive models. The integration of IT empowers law enforcement not only to identify crime hotspots but also to respond promptly to emerging trends, allocate resources strategically, and implement targeted interventions. As technology continues to evolve, the synergy between criminology, clustering techniques, pattern recognition, and information technology becomes increasingly critical for creating a proactive and data-driven approach to addressing the complexities of crime in contemporary society [2-4].

Advances in survey and remote sensing technology over the past ten years have greatly increased the ability to gather large amounts of geographic data [5]. Even with this quantity, the trick is to synthesize and interpret this spatial data intelligently in order to extract relevant information. It is widely acknowledged that the Geographic Information System (GIS) is an essential tool for turning geographic data into insightful knowledge [6-7]. Geographic Information Systems (GIS), a potent instrument for examining geographical occurrences connected to crime [8], have transformed the domains of information technology, criminology, medicine, and other related professions. Finding trends, connections, and hotspots in the geographic patterns of crime incidences is part of the analysis of criminal spatial events in GIS [9]. This method offers a fresh viewpoint on comprehending criminal behavior by enabling the display, investigation, and interpretation of spatial data. Integrating data from various sources, such as crime statistics, demographic data, and environmental factors, can help create comprehensive spatial models [10-12]. These algorithms have a high degree of accuracy in predicting future hotspots and revealing intricate patterns of crime [13]. The use of sophisticated spatial statistical techniques further increases the predictive power of GIS, making it a crucial tool in modern criminology.

The study acknowledges the value of creating hotspot maps as an effective policing tool that can help with understanding various areas and the possible causes of crime occurrence. To fill this gap, the paper's crime pattern analyses focus on forecasting crime patterns by grouping criminal activity data from the PRO 6 Philippine National Police, which is the police department’s record of documented crimes. Using a crime dataset of consummated offenses in the province of Antique from 2018 to 2023, the study shows the risk indicators in order to help predict the hotspot pattern and show it on a map.

2. Research Methods

This research investigates the potential for predicting crime hotspots and coldspots within the Province of Antique. The study concentrates on four specific crimes: direct assault, homicide, murder, and parricide. Antique offers a fascinating opportunity due to its distinct mix of geographic features, demographics, and socioeconomic conditions. Located in Western Visayas, the province encompasses diverse landscapes, from coastal regions to mountainous areas, potentially impacting the way crimes occur across these varied locations.

Antique, a large, sharp edge that resembles a seahorse, is located on the western edge of Panay, a geographical mass that resembles a scarf with three corners. It is separated from the rest of Panay by towering mountain ranges that are 33 km wide at their widest point and 155 km long, which stretch between the West Philippine Sea to the west. There are roughly 2,798.57 square kilometers of land in Antique. Capiz borders it on the east, Iloilo borders it on the southwest, and Aklan borders it on the north and northeast. The Cuyo East Pass, which connects to the Sulu Sea and the larger China Sea, is located on the west side [14]. The province consists of 18 municipalities: one on each of the six islets, three inland, and fourteen along the shore. These are the coastal towns: Libertad, Pandan, Sebaste,
Culasi, Tibiao, Barbaza, Laua-an, Bugasong, Patnonong, Belison, San Jose de Buenavista, Hamtic, Tobias Fornier, and Anini-y. Sibalom, San Remigio, and Valderrama are the towns that are located inland.

The geographic coordinates of Antique are located at longitude 122.08333000 and latitude 11.16667000. The GPS coordinates for the Province of Antique are 11° 10' 0.012'' N and 122° 4' 59.988' E [15]. The Province of Antique landed in 61st place out of 82 Philippine provinces in the 2023 Cities and Municipalities Competitiveness Index (CMCI). This ranking is based on the combined scores of each province’s cities and municipalities across four key areas: economic dynamism, government efficiency, infrastructure, and resiliency [16]. It is divided into 18 municipalities and 420 barangays. The major industries in Antique are agriculture, fishing, and tourism.

Additionally, considering the province’s historical, cultural, and economic dynamics, a thorough analysis of crime distribution could provide valuable insights into the interconnected factors contributing to criminal occurrences. The selection of Antique as a study area enables a localized examination that is relevant to the specific challenges and opportunities present in this province, contributing to the development of targeted and effective crime prevention strategies tailored to its distinct context.

### Table 1. Crime Dataset Cases Registered during the 2018-2023

<table>
<thead>
<tr>
<th>OFFENSE</th>
<th>2023</th>
<th>2022</th>
<th>2021</th>
<th>2020</th>
<th>2019</th>
<th>2018</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIRECT ASSAULTS</td>
<td>15</td>
<td>20</td>
<td>29</td>
<td>40</td>
<td>37</td>
<td>34</td>
<td>175</td>
</tr>
<tr>
<td>HOMICIDE</td>
<td>40</td>
<td>30</td>
<td>43</td>
<td>44</td>
<td>81</td>
<td>19</td>
<td>257</td>
</tr>
<tr>
<td>MURDER</td>
<td>18</td>
<td>22</td>
<td>31</td>
<td>31</td>
<td>33</td>
<td>42</td>
<td>177</td>
</tr>
<tr>
<td>PARRICIDE</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

The study’s crime dataset, which is shown in Table 1, comes from the Police Regional Office (PRO) 6 Philippine National Police (PNP), which was requested via foi.gov.ph website. Numerous characteristics are present in the datasets, such as temporal and spatial crime data from 2018 to 2023 on direct assaults, homicides, murders, and parricides.

### Table 2. Raw Spreadsheet Containing Crime Dataset

<table>
<thead>
<tr>
<th>YEAR</th>
<th>OFFENSE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>DIRECT ASSAULTS</td>
<td>175</td>
</tr>
<tr>
<td>2019</td>
<td>HOMICIDE</td>
<td>257</td>
</tr>
<tr>
<td>2020</td>
<td>MURDER</td>
<td>177</td>
</tr>
<tr>
<td>2021</td>
<td>PARRICIDE</td>
<td>17</td>
</tr>
</tbody>
</table>

The columns represent the years from 2018 to 2021, and the rows represent different offenses, totaling the number of cases for each.
The crime dataset comprised 6,649 cases of consummated crimes. These crimes fell into a number of categories, such as rape, murder, parricide, direct assaults, serious physical injuries, and reckless imprudence that led to homicide, among many others. A geographic information system (GIS) was used to geocode the data in order to get average positions. The data was then linked to particular addresses with matching longitude and latitude values. The study, which focused on spatial analysis, looked at crucial aspects of crime data: the location of crimes, date committed, their description, latitude, and longitude, as shown in Table 2.

The first phase in dealing with the intricacies of crime analysis entails displaying crime incidents on a map through the utilization of OpenStreetMap (OSM) and employing Quantum Geographic Information System (QGIS) software for a clear and effective representation. This procedure is complemented by thorough data preprocessing aimed at improving the quality and significance of the information.

Figure 1 shows the steps that need to be taken in order to use Kernel Density Estimation (KDE) for crime analysis on the QGIS platform. It commences with loading crime data into QGIS, a crucial initial step where point layers representing crime incidents are added to the map. Subsequently, the Processing Toolbox is installed and enabled, ensuring the availability of necessary tools for spatial analysis. Accessing the Processing Toolbox leads to a search for the "Kernel Density Estimation" tool, a pivotal algorithm for estimating spatial distributions. Following this, the KDE algorithm is executed by selecting the crime point layer, configuring parameters like cell size and search radius, and initiating the process. The result is then visualized by adding the resulting raster layer to the map and symbolizing it to convey crime density, employing techniques such as color gradients.
The latter part of the flowchart introduces optional steps to enhance the analysis. Overlaying other layers, such as administrative boundaries or socio-economic data, provides additional context for interpreting crime patterns. The final step involves the interpretation of the KDE map, where users can analyze the spatial distribution to identify areas with varying crime densities.

Kernel Density Estimation (KDE) is a statistical technique that allows for the estimation of the distribution of a set of points in space [17]. In the context of crime analysis using QGIS, KDE can be used to visualize and analyze the spatial distribution of crime incidents. It is a non-parametric way to estimate the probability density function of a random variable. The basic formula for KDE involves placing a kernel (a smooth, symmetric, and positive function) at each data point and summing them up to create a smooth estimate of the underlying distribution. The formula for the one-dimensional KDE is depicted by Equation 1:

$$f(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right)$$

where:

- $f(x)$ is the estimated probability density function (PDF) at the point $x$.
- $n$ is the number of data points.
- $x_i$ represents each individual data point.
- $K$ is the kernel function.
- $h$ is the bandwidth, which controls the width of the kernel and, consequently, the smoothness of the estimated PDF.

The kernel function $K(u)$ is typically a symmetric probability density function, such as the Gaussian (normal) distribution. The choice of kernel function and bandwidth can affect the smoothness and accuracy of the KDE.

For a Gaussian kernel, the formula is depicted by Equation 2.

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

Figure 2 shows the Python code snippet of Kernel Density Estimation algorithm.
3. Results and Discussions

3.1 Direct Assault

Direct assault, a crime against public order, involves either using force or intimidation to achieve the goals of rebellion or sedition without a public uprising, or attacking, using force, or seriously intimidating a person in authority (or their agent) while they're performing their official duties [18].

Figure 3. Hot spots and cold zones of Direct Assault from 2018 to 2023
Figure 3 indicates a comprehensive approach to visualizing the distribution and density of direct assault cases across different municipal police stations, providing insights into spatial patterns and trends over the years. Examining the overall trend from 2018 to 2023, there has been a notable 13.32% decrease in direct assault cases, indicating a positive trajectory towards reduced violence over the six-year period. Analyzing the distribution of recorded cases across various police stations, there is a relatively even spread, as reflected by a standard deviation of 5.37%. This suggests that direct assault cases are not concentrated in a specific geographical area, emphasizing the importance of comprehensive strategies rather than targeted interventions in addressing this type of crime.

### 3.2 Homicide

![Figure 4](image)

**Figure 4.** Hot spots and cold zones of Homicide from 2018 to 2023

Homicide is the unlawful killing of another human being, typically with malicious intent. This legal term encompasses a range of situations, from intentional murder to manslaughter caused by recklessness or negligence [19]. An attempt to commit homicide shall be punished within level 4 [20]. Figure 4 illustrates the homicide data from the past six years and paints a contrasting picture compared to direct assaults. While direct assault cases show a fluctuating pattern with a peak in 2020 at 40 cases and a subsequent decrease in 2023, the spatial pattern of homicide cases from 2018 to 2023, distributed across various municipal police stations, reflects a dynamic landscape of violent incidents. The total number of homicides shows a peak in 2019 with 81 cases, followed by a subsequent decline, reaching 40 cases in 2023. Some police stations consistently report high homicide rates, contributing significantly to the overall numbers. The spatial distribution underscores the localized nature of homicides, indicating that certain areas bear a heavier burden of this violent crime.

### 3.3 Murder

![Figure 5](image)

**Figure 5.** Spatial analysis of murder cases from 2018 to 2023

Murder is the unlawful killing of another human being with malice aforethought. This means it's not accidental or justified, but done with intent and premeditation. It's a serious crime with grave consequences, and understanding its definition is crucial for upholding the value of human life and ensuring justice [21]. Figure 5 illustrates the spatial analysis of murder cases from 2018 to 2023 across
various municipal police stations and reveals distinct patterns in the distribution of this severe crime. The total number of murders decreased from 42 in 2018 to 18 in 2023, reflecting an overall declining trend. Notably, some police stations consistently report higher murder rates, while others have reported zero cases throughout the period. The data suggests that murder incidents are concentrated in a specific municipality, possibly influenced by local socio-economic factors, demographics, or law enforcement effectiveness.

Figure 5. Hot spots and cold zones of Murder from 2018 to 2023.

3.4 Parricide

Figure 6. Hot spots and cold zones of Parricide from 2018 to 2023
The data on parricide, the act of killing one's own parent [22], paints a concerning picture of relative consistency with occasional spikes. Figure 6 reveals a varying but generally low incidence of this severe crime. The recorded cases fluctuate from 3 cases in 2018 to 2 cases in both 2022 and 2023, with a peak of 6 cases in 2021. The spatial distribution of parricide cases across different municipal police stations further highlights the localized nature of these incidents. While some municipal police stations have reported cases consistently over the years, others show sporadic instances. The overall low frequency of parricide cases suggests that this type of crime is relatively rare, and the concentration in specific areas may be influenced by local factors such as socio-economic conditions, demographics, or law enforcement practices.

4. Conclusion and Recommendations

To effectively combat Antique's diverse crime landscape, a multi-layered approach is necessary. In areas burdened by high homicide and murder rates, targeted prevention programs must be implemented, addressing potential local factors like socio-economic conditions or demographics. Additionally, investigating the effectiveness of various community-based initiatives in reducing crime would provide valuable insights for optimizing social support programs. Furthermore, the development of a web-based decision support system could integrate crime data, demographic information, and social program metrics to guide resource allocation and program design. By continuously monitoring and analyzing crime trends through this system, policymakers can work towards a safer and more secure future for communities.

References


