

Trajectory Analysis Applied to Crime Incidence

Rodrigo Tapia-McClung[1(✉)], Rodrigo Lopez-Farias[2,3], and Camilo Caudillo-Cos[1]

[1] Centro de Investigación en Ciencias de Información Geoespacial CDMX, Contoy 137, 14240 Lomas de Padierna, Tlalpan, Mexico
{rtapia,ccaudillo}@centrogeo.edu.mx

[2] Centro de Investigación en Ciencias de Información Geoespacial Querétaro, Pedro Escobedo, 76709 Querétaro, Mexico
rlopez@centrogeo.edu.mx

[3] Secretaría de Ciencia, Humanidades, Tecnología e Innovación, Insurgentes Sur 1582, Crédito Constructor, 03940 Benito Juárez, Mexico

Abstract. In the study of criminal incidence, the lack of graphical specialized tools makes it difficult to properly communicate in a consistent way, and with simple terms, its spatio-temporal evolution. In this work, we propose the application of a count-based trajectory time series analysis to study the spatio-temporal evolution of larceny theft in Mexico City. This framework helps describing the spatio-temporal evolution of an event by identifying and grouping the individual location dynamics that contribute in the description of the types and magnitude of different changes in space. These changes can be expressed in terms of gains, losses, and possible alternations that generate exchange of criminal incidence without observing significant gross loss or gain in incidence. We highlight the characteristics that are beneficial to the interpretation of the spatio-temporal evolution of crime incidence and the drawbacks of using this technique. We conclude that while there is a general decreasing trend of crime incidence, the graphical method is sensitive to show internal patterns such as gains, losses and alternation, that are important to analyze key spatio-temporal patterns throughout the years. This approach is intended to be useful for evaluating strategies for the reduction of criminal incidence.

Keywords: Trajectory analysis · Crime incidence · Spatio-temporal time series

1 Introduction

The main objectives of criminology are understanding, communicating, and delineating strategies to minimize criminal incidence. Understanding criminal phenomena requires the articulation of diverse areas of work like sociology, such as psychology, anthropology, political science law [3], and any other discipline

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
O. Gervasi et al. (Eds.): ICCSA 2025 Workshops, LNCS 15890, pp. 358–370, 2026.
https://doi.org/10.1007/978-3-031-97606-3_24

that contributes to study the phenomena such as statistics and geospatial sciences. In this work we propose applying a Trajectory Analysis, that is descriptive statistics method oriented to visualize the magnitude of criminal incidence in terms of its quantity and spatial change.

A trajectory, in the broad sense, is defined as "the path that a moving object follows through space as a function of time" [14]. In geospatial sciences, the trajectory analysis of criminal incidence is the study of its patterns and describing its characteristics of importance in the spatio-temporal domain.

A count-based Trajectory Analysis approach for binary cover maps series put forward in [2] is applied in this work for the study of the evolution of criminal events through space and time in terms of gain and loss components. This approach is based on the classification of every location of the map according to the type of temporal trajectory of a binary variable taken throughout time. Then, the size of the gain and loss component is obtained as the contribution of different trajectories.

We posit that this kind of description would allow the classification of the type of change in criminal incidence in an area of study. In this sense, it would be possible to identify whether crime incidence shows a general increase, decrease, if it alternates in space or if it is only being exchanged from one place to another without observing an apparent loss. Specifically, we carry out this analysis and find it useful in recognizing crime patterns in Mexico City.

This paper is organized as follows. We first present the background and previous work that justifies the connection between the study of larceny theft and the count-based trajectory analysis. We then introduce the basic concepts of the trajectory analysis of spatio-temporal data used in this contribution along with the study area and data used. Then, we present the processed data used to make conclusions regarding the dynamics of the crime incidence in Mexico City. Lastly, we present some constructive criticism highlighting the areas for improvements for this kind of application and future work.

2 Background and Related Work

Time series analysis, longitudinal analysis, and trajectory analysis are all methods for analyzing data over time, but they each have different purposes. In a time series analysis, data points are collected and analyzed at regular intervals over time to identify patterns and trends. Usually, one can expect to predict future values. In a longitudinal analysis, the interest is on studying how variables change over time for a group of subjects. The goal is commonly identify cause-and-effect relationships. In a trajectory analysis, the objective is usually to describe how a variable changes over time or age. It can be used to study how covariates affect the shape of a variable [8].

In the social sciences, sequence analysis has been translated into a set of longitudinal data methods [1]. Surveys like this one, focus on highlighting the application of qualitative methods that study the population and their changes throughout time, also referred to as life course trajectories or event history trajectories [9]. Over time, computational approaches have been emphasized to

characterize complex successions of different states [12], finding that social organization derives from spatial, hierarchical, and other ordered phenomena.

Trajectory analysis is a sub-discipline commonly used for exploratory analysis and to develop hypotheses to explain differences between certain groups. In criminology, it has been used to examine how large samples of individuals could be grouped according to crimes throughout life [10]. More recently, longitudinal analyses have also been applied at the place scale to identify trends and clusters of street segments in terms of crime incidences [16] and microgeographic hotspots [15]. Cluster analysis methods have also been adapted to extend to spatio-temporal patterns, resulting in the characterization of trajectories in other longitudinal studies of crime [5]. This last example implements the classification of trajectories with a longitudinal k-means algorithm, which identifies groups of spatial units with similar behaviors. This work also strongly suggests to retain the resource allocation information regarding the time and place of occurrence and use count-based methods to measure the frequency of events in specific locations.

Examples of the application of the count-based Trajectory Analysis include measuring and tracking the trajectories of losses and gains of soybean cultivation [11] and to cross comparing the marsh evolution in three sites within the United States Long-Term Ecological Research Network [2]. Both use the proposed count-based methodology for a time series trajectory analysis to describe spatio-temporal changes in land cover to elucidate details and differences of the evolution of certain types of cover.

The methodology proposed in [2] is a relatively new approach and, thus, the diversity of applications is yet scarce. In this sense we propose using this methodology for a spatio-temporal graphical analysis of the dynamics of larceny theft. But first, we briefly explain this methodology in the next section.

3 Count-Based Trajectory Analysis of Spatio-Temporal Discrete Data

The count-based time series trajectory analysis consists of analyzing and classifying the sequences of all the individual locations in the area of interest of cover maps.

The characteristic of this approach is the preservation of both qualitative and quantitative information. The qualitative information is associated with the classification of the dynamic described by the sequence of presence and absences of a variable for each location. The quantitative information is associated with the total of gains and looses of a variable in the space for each time interval through the time. Other approaches, such as trajectory identification through unsupervised learning, have the disadvantage that the quality of the clusters will depend on an arbitrary number of classes defined a priori. The approach summarizes the possible trajectories (changes over time) in six classes, providing a synthetic view of the crime incidence behavior.

Initially, every local trajectory is a sequence of binary labels representing the presence or absence of a variable of interest at different moments in time. The method classifies trajectories into one of eight different types, depending on the number of time intervals considered. The possible trajectories defined with this approach appear when considering a minimum set of four maps [M_1, M_2, M_3, M_4] and are shown in Table 1.

Table 1. Schematic representation of the different types of space time trajectories. White and black cells represent the absence and presence of the variable of interest, respectively.

ID	M_1	→	M_2	→	M_3	→	M_4	Trajectory Classification
One time interval								
1	□	→	□					Stable Absence
2	□	→	■					Gain without Alternation
3	■	→	□					Loss without Alternation
4	■	→	■					Stable Presence
Two time intervals								
5	□	→	■	→	□			All Alternation Gain First
6	■	→	□	→	■			All Alternation Loss first
Three time intervals								
7	■	→	□	→	■	→	□	Loss with Alternation
8	□	→	■	→	□	→	■	Gain with Alternation

The first column shows the ID of the location, the second one indicates the map sequence, and the last one contains the corresponding trajectory classification.

Location with IDs $\{1, 2, 3, 4\}$ only have one time interval available and their representation is limited to four possible trajectories: stable absence, stable presence, gain and loss without alternation. For locations with IDs $\{5, 6\}$, having two time intervals, the representation is extended in addition to the previous trajectories, with presence or absence at the initial and final time intervals. However, these additional trajectories exhibit alternation. They are aptly named: all alternation gain first and all alternation loss first. Lastly, for location IDs $\{7, 8\}$, with three time intervals, apart from the previous trajectories also being represented, it is possible to include one for gains or losses with alternation in the entire time interval. For practical and descriptive purposes, in the case of four or more time intervals, the method considers the same classes with the disadvantage of loosing details regarding the temporality and number of the changes of the events.

While it is possible to map every trajectory class, it only gives a simplified idea regarding the description of the dynamics of the phenomena because it hides temporal information details. It is here that the need of a Trajectory Plot arises.

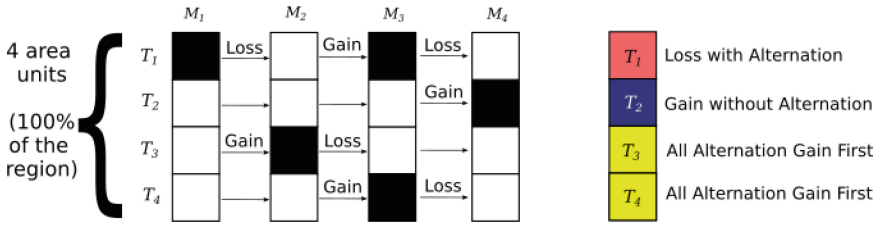

Fig. 1. Example of binary maps with trajectories and their classification.

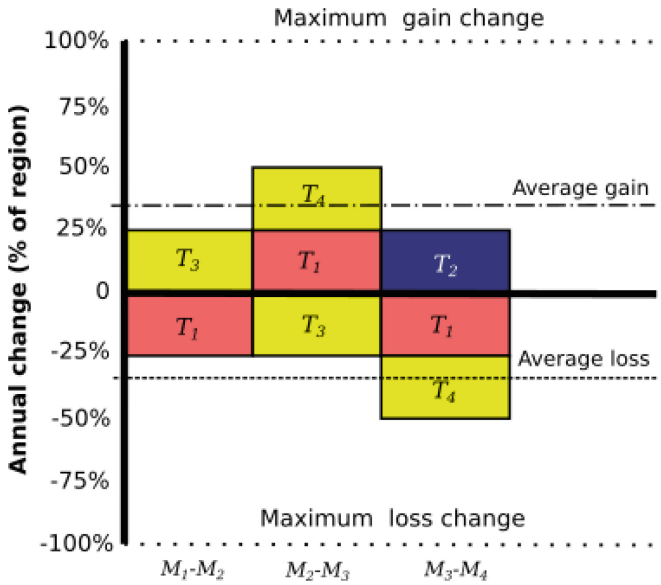

Fig. 2. Example of a trajectory plot.

This plot essentially records the contribution of the changes to the magnitude of gains or losses of change for every time interval.

These trajectories contribute to the overall gains and looses for every time interval. A trajectory plot is a way to graphically and visually represent these contributions and easily and quickly convey how these overall changes behave through time. The only trajectories that do no contribute to changes in the trajectory plot are "Stable Presence" and "Stable Absence".

Figure 1 shows an example of the evolution of four maps M_1, M_2, M_3, and M_4 with an extent of four areal units located implicitly in space. Each unit i has a corresponding trajectory T_i. In this example we see that trajectory T_1 is of the kind "Loss with Alternation", trajectory T_2 is "Gain without Alternation", and trajectories T_3 and T_4 are of type "All Alternation Gain First".

The gross gains and losses for each time interval in this example are depicted in the trajectory plot in Fig. 2.

The trajectory plot shows change in terms of gross gains and losses for each time interval. It presents the contribution source associated with each trajectory observed in Fig. 1. The horizontal axis represents the time intervals, while the vertical one indicates the gross gains and losses for each corresponding time interval as an annual percentage of the area of study. The dash-dotted line shows the average gain, while the dotted one shows the average loss over the full time interval. This graphical representation summarizes all the trajectories and their respective changes with an idea of the size and the year of occurrence.

4 Study Area and Data

We apply the time series trajectory analysis methodology in Mexico City. Located in the south-central portion of the country, it is the capital, and covers an area of about 1495 km^2 (roughly 0.1% of the country), making it the smallest state. It is also divided into 16 municipalities and has a population of about 9.2 million, representing 7.3% of the total population of the country.

According to results from the most recent National Survey on Victimization and Public Safety Perception carried out by the National Institute of Statistics and Geography (Instituto Nacional de Estadística y Geografía - INEGI), roughly one third of households in the city reported being victims of crime incidence [7]. Additionally, crime incidence has been consistently reported with decreasing trends in the last five years.

It, then, becomes interesting to explore the spatio-temporal pattern of this reported decrease. Because not all crime types behave the same way and some have higher "dark" or "hidden figures" (the gap between reported and unreported crimes), we focus on those that have the lowest dark figures.

We use yearly larceny incidence in Mexico City from 2019 to 2023 that contain little over 55,000 data points distributed in the city reported by the Office of the Attorney General of Mexico City [4].

Mexico City's Secretariat of Citizen Security divides the city in 71 territorial units (equivalently to coordinations or jurisdictions in other countries) to monitor and serve policing. They have varying sizes and areas that range from 1.08 to 73.43 km^2, with a mean area of 12.70 km^2 and a standard deviation of 9.68 km^2. This variation can hinder an easy and straightforward comparison of larceny incidence with respect to population at risk and police areas of attention.

For that reason, a regular tessellation was preferred over the original jurisdiction geometries. The H3 hierarchical tessellation of level 7 was deemed a reasonable choice, with an average area of 5.16 km^2 and average edge length of 1.41 km (https://h3geo.org/). Mexico City is covered by 301 hexagons and we count the number of observed incidences in each hexagon for each year. It is important to note that using cells of the same size gives the advantage of avoiding the modifiable areal unit problem, and that they allow for efficient data aggregation and classification of similar groups [13].

Figure 3 shows the locations of Mexico City, its public safety territorial units together with the extend of the city's administrative boundary, and the data

points considered for 2019–2023 on top of the H3 units. The number of data points considered for each year is shown in Table 2.

5 Results

This section describes the count-based Trajectory Analysis applied to larceny data in Mexico City from 2019 to 2023. First, the H3 level 7 tessellation that covers the study area was created and the number of observation inside each cell was summarized by year. Next, a raster stack with a 50 m spatial resolution was created. The result is a series of images depicting the distribution of larceny in

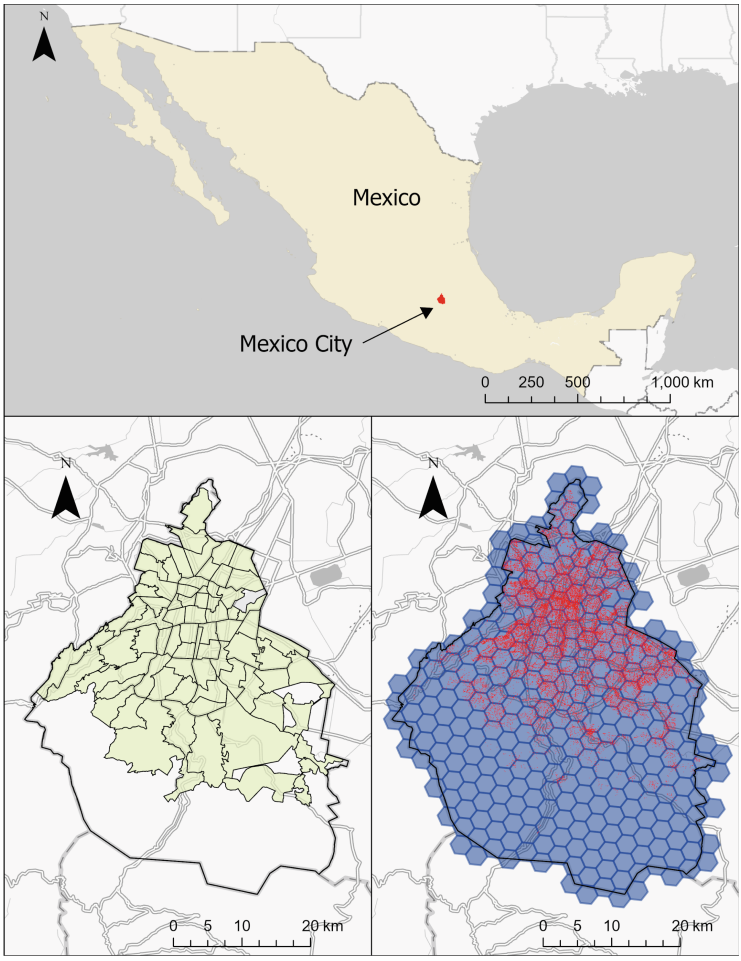

Fig. 3. Area of study and data: The location of Mexico City, its public safety territorial units, H3 level 7 hexagons and data points.

Table 2. Number of data points considered for each year.

Year	Data points
2019	15,440
2020	10,122
2021	10,762
2022	10,273
2023	8,600
Total	55,197

space and time. It is important to note that, while the original methodology is carried out in terms of presence or absence of a given class (binary data), here we apply the methodology to integer (counts) data.

Yearly hex binned incidence count is rasterized and stacked together as input for the `timeseriesTrajectories` R package [6]. By analyzing the values in each cell for consecutive years, it is possible to quantify the magnitude and direction

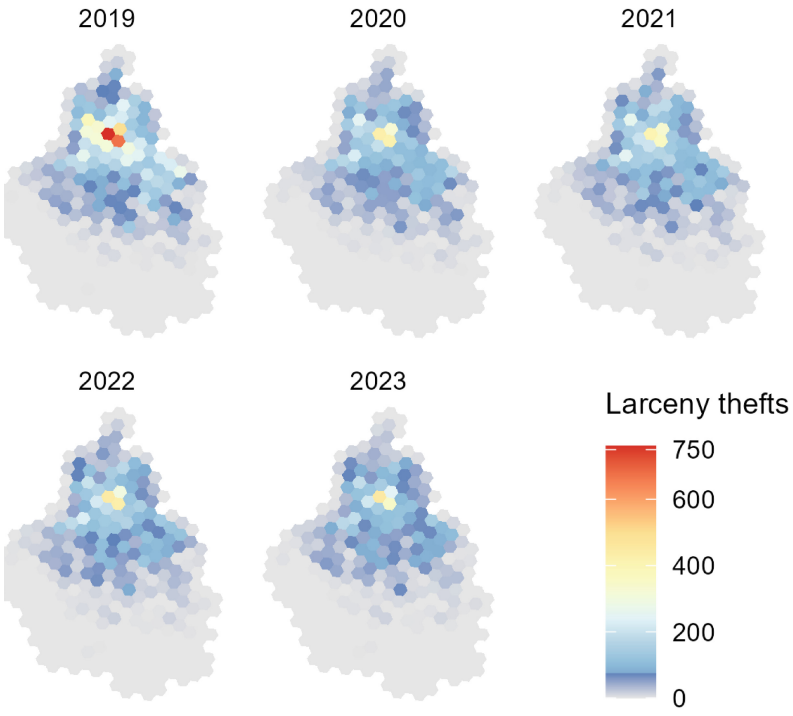

Fig. 4. Spatial and temporal distribution of larceny counts in H3 hexagons covering Mexico City.

of changes in larceny. Figure 4 shows the yearly distribution of counts inside the tessellation.

It can be seen that 2019 has a significant drop in counts. Two cells show high values compared to the rest of that year and all of the subsequent years. Of the 301 hexagons that cover the study area, only 189 have reported incidences and the remaining 112 have no observations. Because it can be overwhelming to try and follow the changes for any given hexagon, trajectories are calculated and classified into one of the categories introduced in Sect. 3. The number of hexagons classified in each type of trajectory are shown in Table 3.

Table 3. Number of hexagons classified in each trajectory.

Trajectory	Number of hexagons
Loss without Alternation	21
Loss with Alternation	117
Gain without Alternation	3
Gain with Alternation	31
All Alternation Loss First	6
All Alternation Gain First	11
Stable Absence	112

In contrast to [2], here we do not observe the "Stable Presence" class because observing exactly the same counts every year in the time interval is unlikely, therefore is expected to observe variation in counts in a cell from one year to the next. The "Stable Absence" value means there are no counts in a given cell for every year (which correspond to the 112 hexagons with no observations).

The trajectory plot in Fig. 5 shows that while there certainly are gains and losses at the end of the interval, there are also variations in between. For instance, the significant change that had already been identified for the 2019–2020 time interval in Fig. 4 can be quantified in terms of the annual change expressed as a percentage of the region of study. In this case, the overall loss is greater than 30%, and the overall gain is less than 1%. Then, 2020–2021 shows an important gain of larceny theft accompanied of a slightly decrement. Then, from 2020 to 2023, a consistent drop trend for gains and a consistent increase for loss is observed. This change is much harder to appreciate in Fig. 4. Also, the gross gain (about 3.5%) is much less than the gross loss (about 14.3%), confirming the initial reports of a general decreasing trend of crime incidence, which can also be partially inferred from the number of data points reported for each year in Table 2.

The trajectory plot also describes that the major source of alternation comes from loss with alternation trajectories. This means that, in general, larceny thefts decrease but there are fluctuations. Another important component is "Loss without Alternation", which shows a solid decrement in this regard in approximately 5% of the region for the time interval 2019–2020. Then, the contribution of this

trajectory class for subsequent time intervals is remarkably less. Then the "All Alternation Gain First", "All Alternation Loss First", and "Gain with Alternation" trajectories have smaller contributions in general to gains and losses in the plot.

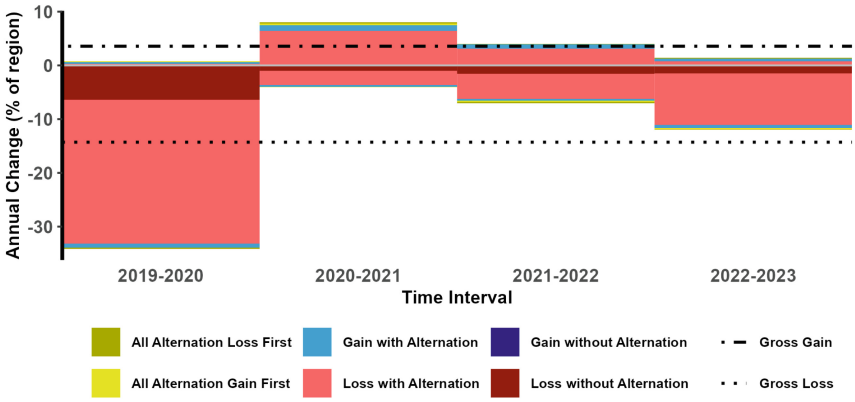

Fig. 5. Results of time series trajectory analysis with different types of gains and losses for each year.

Figure 6 shows the trajectories in space and it is possible to discern which cells had an overall gain, loss, or an alternating behavior. As mentioned before, cells that had no observations are classified as "Stable Absence". It is interesting to note that the region of the city where reports are consistently higher shows a decreasing trend over time, mostly classified as losses with and without alternation. Equally interesting is to find that towards the south there are more gains with alternation rather than losses. This could be an indication that due to more operatives in the central region of the city, some activity may be shifting towards other locations.

Figure 7 shows the trajectories for each cell. The lines depict the yearly behavior, clearly showing if a particular cell had gains, losses or an alternating behavior. It can be seen that there are few cells with pure gains, and the most populated one correspond to "Loss with Alternation" followed by "Gain with Alternation" and "Loss without Alternation".

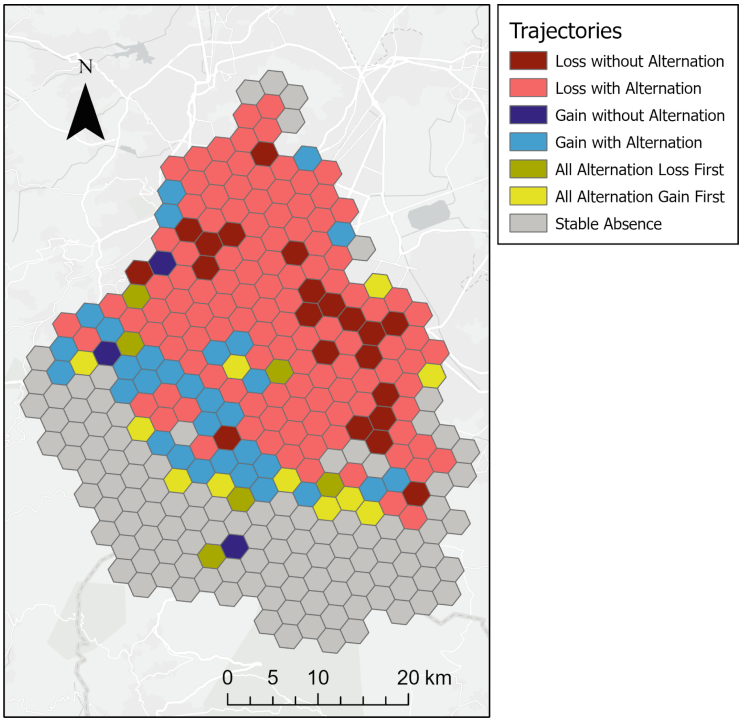

Fig. 6. Trajectories for larceny in Mexico City from 2019 to 2023.

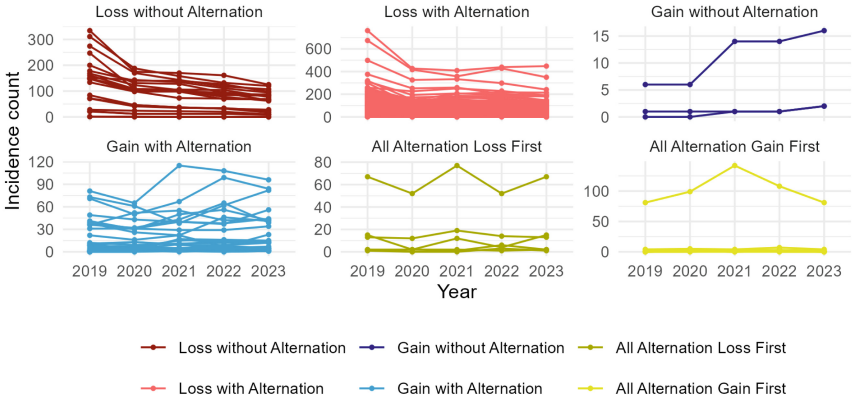

Fig. 7. Incidence count for each type of trajectory and year. Each line corresponds to one cell from the H3 tessellation used.

6 Conclusions and Future Work

We applied a time series trajectory analysis to larceny data in Mexico City and conclude that, indeed, there is a descending trend in general the last five years, but with slight area with increments of this crime. While it is complicated to visually assess these changes both through space and time with regular maps (like in Fig. 4), a trajectory plot provides additional details on the yearly dynamics of the incidence under study.

By identifying which areas are classified under each type of trajectory, it is possible to start asking questions about how is larceny incidence behaving at a more local scale. Some regions are fluctuating, but the trajectories alone are not enough to show the magnitude. By considering the incidence count for each type of trajectory it is then possible to put these changes into perspective, both in terms of other regions with the same trajectory and the rest of the cells; that is, it gives a first glimpse about the local and global behavior of the trajectories.

We have only scratched the surface of this methodology applied to studying crime incidence. It will be worth carrying out a temporal sensitivity study to assess how significant are these results in terms of their permanence or fluctuation through coarser or finer periods. In this sense it will be interesting to aggregate larceny data on a monthly or weekly basis instead of a yearly one and compare the resulting trajectory plots.

On the same sensitivity note, it could be possible to explore how changing the cell and time interval sizes affects these results. That is, study the dependence on their spatial and temporal scale. It is expected that there will be a critical cell size that will reveal interesting shifts in the distribution of reported crimes. Unfortunately, H3 tessellations' sizes are fixed in terms of their hierarchical structure and thus may not be the best option for this sensitivity analysis. It may be beneficial to use square grids for this purpose.

On the other hand, how valid could our conclusions be if data are noisy and there is a lot of alternation? It would then be difficult to guarantee that gains or losses are authentic. For that purpose, an initial exploration can be carried out by studying the behavior of counts shown in Fig. 7. This data sets does not seem to be too noisy as behavior within trajectories does not differ much. It would be, however, interesting to develop more refined tests of noise in data.

Lastly, it could also be interesting to confront these results and associated conclusions with alternate ways of accounting for the observed larceny. For instance, it could be relevant to study the spatial distribution of clusters of high or low values of observed counts and their associated trajectories to compare them with the ones presented here.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Abbott, A.: Sequences of social events: concepts and methods for the analysis of order in social processes. Hist. Methods **16**(4), 129–147 (1983). https://doi.org/10.1080/01615440.1983.10594107
2. Bilintoh, T.M., Jr, R.G.P., A.Z.: Methods to compare sites concerning a category's change during various time intervals. GISci. Remote Sens. **61**(1), 2409484 (2024). https://doi.org/10.1080/15481603.2024.2409484
3. Brazil, K.J., Whittingham, L.M.: Criminology. In: Shackelford, T.K., Weekes-Shackelford, V.A. (eds.) Encyclopedia of Evolutionary Psychological Science, pp. 1–15. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-16999-6_3193-1
4. Fiscalía General de Justicia: Carpetas de investigación FGJ. https://datos.cdmx.gob.mx/dataset/carpetas-de-investigacion-fgj-de-la-ciudad-de-mexico (2024), Accessed 20 Sep 2024
5. Harinam, V., Bavcevic, Z., Ariel, B.: Spatial distribution and developmental trajectories of crime versus crime severity: do not abandon the count-based model just yet. Crime Sci. **11**(1), 1–15 (2022). https://doi.org/10.1186/s40163-022-00176-x
6. timeseries Trajectories. https://github.com/bilintoh/timeseriesTrajectories (2024), Accessed 15 Jan 2025
7. Instituto Nacional de Estadística y Geografía: Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE). https://www.inegi.org.mx/programas/envipe/2024/ (2024), Accessed 20 Feb 2025
8. Longford, N.T.: Longitudinal and Time-Series Analysis, pp. 335–370. Springer, New York, NY (2008). https://doi.org/10.1007/978-0-387-73251-0_11
9. Mills, M.: Sequence Analysis. In: Introducing Survival and Event History Analysis, chap. 11, pp. 213–226. SAGE Publications Ltd, London (2012). https://doi.org/10.4135/9781446268360
10. Nagin, D.S., Tremblay, R.E.: Analyzing developmental trajectories of distinct but related behaviors: a group-based method. Psychol. Methods **6**(1), 18–34 (2001). https://doi.org/10.1037/1082-989X.6.1.18
11. Pontius Jr, R., Bilintoh, T., Oliveira, G., Shimbo, J.Z.: Trajectories of losses and gains of soybean cultivation during multiple time intervals in Western Bahia, Brazil. In: Anais da Space Week Nordeste 2023. Campinas. Galoá (2023)
12. Ritschard, G., Studer, M. (eds.): Sequence Analysis and Related Approaches. LCRSP, vol. 10. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-95420-2
13. Sadahiro, Y.: Analysis of the relations among spatial tessellations. J. Geog. Syst. **13**(4), 373–391 (2011). https://doi.org/10.1007/s10109-010-0127-3
14. Shoval, N.: Trajectories: analysis. In: International Encyclopedia of Geography, pp. 1–5. John Wiley & Sons, Ltd (2017). https://doi.org/10.1002/9781118786352.wbieg0611
15. Weisburd, D.: The law of crime concentration and the criminology of place. Criminology **53**(2), 133–157 (2015). https://doi.org/10.1111/1745-9125.12070
16. Weisburd, D., Bushway, S., Lum, C., Yang, S.M.: Trajectories of crime at places: a longitudinal study of street segments in the city of seattle. In: Quantitative Methods in Criminology, pp. 443–481. Routledge (2017). https://doi.org/10.4324/9781315089256-19