Homicide in São Paulo, Brazil: Assessing spatial-temporal and weather variations

Vânia Ceccato*

Institute of Criminology, University of Cambridge, Sidgwick Avenue, CB3 9DT, Cambridge, UK
Available online 24 October 2005

Abstract

Although São Paulo is one of the most dangerous cities in the world, very little is known about the variations of levels of crime in this Brazilian city over time. This article begins by investigating whether or not homicides are seasonal in São Paulo. Then, hypotheses based on the principles of routine activities theory are tested to evaluate the influence of weather and temporal variations on violent behaviour expressed as cases of homicides. Finally, the geography of space–time clusters of high homicide areas are assessed using Geographical Information System (GIS) and Kulldorff’s scan test. The findings suggest that central and peripheral deprived areas show the highest number of killings over the year. Moreover, homicides take place when most people have time off: particularly during vacations (hot months of the year), evenings and weekends. Overall, the results show that temporal variables are far more powerful for explaining levels of homicide than weather covariates for the Brazilian case—a finding that lends weight to the suggested hypotheses derived from routine activity theory.

1. Introduction

There are several reasons for examining the impact of temporal and weather variations on homicide in São Paulo, Brazil. First, empirical evidence in this field is largely based on case studies from countries of a temperate climate, either Europe or the United States (e.g., Cohn, 1990, 1993; Cohn & Rotton, 2000, 2003a,b; Farrell & Pease, 1994; Field, 1992; Hakko, 2000; Harries & Stadler, 1983, 1988; Harries, Stadler, & Zdorkwski, 1984; Semmens, Dillane, & Ditton, 2002). Whilst the yearly average temperature in many European cities does not reach above 10°C, in São Paulo it is often above the yearly average, which is 18 degrees. It is quite possible that this difference in temperature between climate regimes has an effect on how, when and where human interactions occur. As Harries (1997) suggests, ‘there is a need to obtain empirical evidence from a variety of climatic regimes in order to have a better understating of mechanisms linking levels of violence and weather and temporal variations’ (p. 150).

Second, relatively little is known about temporal and weather variations relating to criminal behaviour in developing countries. São Paulo is an important example since it is the prime metropolitan area in South America and one of the most populated and dangerous cities in the world (Mercer Human Resources, 2004; UN, 2003). Compared to other international cities, São Paulo is a compact city with about 21,000 inhabitants per square mile, a population of over ten million and the larger metropolitan area nearly 18 million, as recorded in 2003. The metropolitan area of São Paulo had three of the most violent areas in Latin America during the 1990s with 140 homicides a year per 100,000 inhabitants in some districts (InterAmerican Development Bank, 2000). Although homicide rates have been falling over the last 5 years, the Police in São Paulo recorded a total of 24,242 homicides between 1999 and 2003, with an average of 13 a day.

Third, past studies assessing the variations in crime events by temporal and weather variables vary widely in
levels vary and the way in which these variations relate
researchers have found empirical evidence on how crime
and the fewest during winter. Since his seminal work,
crimes against a person is committed during summer
1984; Rotton & Cohn, 2004). Relatively little is known
about space–time variations of criminal events and,
particularly in developing countries, spatial analysis of
crime was, until recently, rare. One reason was that
offence data was either unavailable or of poor quality.
An important aspect of this particular study is that it is
based on a new and extensive geo-referenced crime
database was made feasible by the use of Geographic
Information System (GIS) technology.

This article makes three contributions. First, it assess
whether or not homicides are seasonal in São Paulo,
Brazil. Second, it reports the testing of hypotheses for
levels of homicides being affected by the suggested
effects of the weather and temporal variations on human
behaviour and activity patterns. Ordinary Least-Square
(OLS) regression modelling allows us to identify the
variables that most significantly contribute to the
variation in homicide levels by exploring three types of
modelling strategy (by hour, by day and by week).
Third, a contribution is made to the knowledge base
concerning seasonal changes in the geography of high
homicide areas in São Paulo using space–time cluster
techniques and GIS. In other words, this article
demonstrates what the geography of clusters of homi-
cides look like during different seasons and assesses the
possible reasons for these differences.

The structure of this paper is as follows. Section 2
contains a brief review of previous research in the field
of crime and how it is influenced by weather and
temporal variations. This is followed by a discussion of
the main important theories that link weather and
temporal variations to levels of crime. In Section 3, the
theories and hypotheses are applied to São Paulo.
Section 4 presents the process of data acquisition and
data quality used in the analysis. Levels of homicides
over time are analysed in Section 5. In Section 6, the
results of the OLS models are discussed followed by the
analysis of the space–time clusters in Section 7. Finally,
Section 8 summarizes the main findings and considers
directions for future work.

2. Previous research

Quetelet (1842) suggested that the greatest number of
crimes against a person is committed during summer
and the fewest during winter. Since his seminal work,
researchers have found empirical evidence on how crime
levels vary and the way in which these variations relate
to weather conditions. Findings have often been contra-
dictory since the time scale of these studies differed
widely, as did their methodology (for a comprehensive
review, see Cohn, 1990; Cohn & Rotton, 2000; Harries,
Stadler, & Zdorkwski, 1984). Studies of the relationship
between homicide and heat are no exception (Cohn,
1990; Cohn & Rotton, 2000; Hakko, 2000; Rotton,
Cohn, Peterson, & Tarr, 2004). There seems to be no
particular season for homicides according to evidence
from Cheatwood (1988) and Yan (2000). Michael and
Zumpe (1983), for instance, showed no clear links
between temperature and monthly number of homicides
in different geographic locations and neither did Maes,
Meltzer, Suy, and De Meyer, (1993) when they assessed
the effect of weather variables on homicides levels.
Using cross-sectional and time series analyses Rotton
and Cohn (2003) showed that temperature is associated
to many violent crimes, such as assault or rape but the
effect of temperature was not verified for cases of
homicide.

Other studies do show however that weather vari-
ables, especially temperature, are correlated with serious
and lethal violence (Anderson & Anderson, 1984;
Anderson, Anderson, Dorr, DeNeve, & Flanagan,
2000; Anderson & Bushman, & Groom 2002; Anderson,
Bushman & Groom, 1997; Dexter, 1899; Rotton & Frey,
1985). In Hakko (2000) the seasonal pattern of homicides
shows that in Finland, from 1957 to 1995, there was a statistically significant peak in summer and a
trough in winter. Findings from preliminary studies in
Brazil also show some evidence of seasonal variations of
homicides. Lima (1998) suggests that the greatest
number of homicides in São Paulo state took place
during the first 4 months of the year, with the exception
of February, which shows lower rates. Beato Filho,
Assunção, Santos, Santo, Sapori, and Batitucci (1999)
found that the highest quantity of violent crimes in
Minas Gerais state occurred between January and
March and October and December. A similar variation
in São Paulo state between 1996 and 2002 is also
suggested by Kahn (2003). In the next section the main
important theories that link weather and temporal
variations to levels of crime are discussed. This is
followed by a synthesis of these theories applied to the
case study of São Paulo as well as the hypotheses which
are proposed in section three.

2.1. Theories on the effect of weather and temporal
variations on crime

Studies often assess whether or not changes in crime
levels are affected by weather conditions in two
ways. Either directly, by examining how people’s
behaviour is influenced by weather (e.g. when heat
effects violent behaviour), or indirectly, by exploring
the changes in human activity patterns determined by
weather variations that make people more inclined to commit a crime or be victimized (e.g. people are more willing to engage in outdoor activities in the summer than in the winter).

When this is assessed directly, assumptions are based on the concept that changes in the weather changes or extremes of weather function as ‘stresses’. For example, individuals who are highly sensitive to changes in the weather might exhibit behavioural or mood changes, leading to a criminal act (Cohn, 1990). The general aggression model or also called GAM theory (Anderson et al., 2000) suggests that weather variables, but particularly temperature, heightens physiological arousal and leads to aggressive thoughts and, in certain cases, violence. Although most of the literature indicates that more violent crimes occur on hot days (e.g. Dexter, 1899; Hakko, 2000; Rotton and Frey, 1985), there might be a temperature threshold that triggers the reverse behaviour. In other words, people’s motivation to engage in aggressive behaviour is reduced as a result of a need to avoid the heat. This is consistent with Baron and Bell’s (1976) negative affect escape (NAE) model that suggests that moderately high temperatures cause negative affect (which leads individuals to behave more aggressively) whilst very high temperatures result in an attempt to escape the situation and engage in activities that reduce discomfort. The precise point at which temperature becomes uncomfortable is not clear (Hipp, Bauer, Curran, & Bollen, 2004) and certainly depends on the yearly average temperature of the place. The relationship between temperature and violence would therefore be curvilinear and not linear as stated in previous models. Rotton and Cohn (2004) suggest however that the relationship between temperature and violence is highly dependent on location and time. There are times when this relationship is curvilinear as suggested by the NAE model, whilst at other times it is either linear (e.g. early evening hours) or nonexistent, when temperature and violence do not appear to be related (e.g. early morning hours).

When conducting an indirect assessment, analyses are based on routine activity theory (Cohen & Felson, 1979). This theory suggests that an individual’s activities and daily habits are rhythmic and consist of patterns that are constantly repeated. Such behaviour is influenced by changes in the environment, such as climate. Most crimes depend on the interrelation of space and time: offenders’ motivation, suitable targets and absence of responsible guardians. During periods when people are more often outside there is a greater risk of victimization. This is because there is a greater chance of potential victims being in the same place at the same time as motivated offenders. This is the basis of the explanation of the mechanisms behind seasonal (summer–winter) and weekly (weekend–weekday) variations in levels of certain types of violent crimes, particularly assault and rape (e.g. Rotton & Cohn, 2003).

The effect of weather overlaps the impact of temporal variations on crime since weather determines the probability and intensity of routine activities. There is evidence from time-budget studies (see Petland, Harvey, Lawton, & McColl, 1999 for a review of time-budget studies) that human activities, including criminal acts, relate significantly to temporal variations such as, weekends–weekdays, holidays (e.g. see Cohn & Rotton, 2003a,b). It has been suggested in the recent literature (e.g. Rotton, Cohn, Peterson, & Tarr, 2004) that routine activity theory and the NAE model should be seen as complementary approaches rather than as competitors in the attempt to explain crime levels. This because the principles of routine activity theory can be used to explain nonlinear patterns of violent crimes if it can be assumed that individuals shift the locus of their activities from outdoor (e.g. streets) to indoor (e.g. home) locations when temperatures are either very high or low. The difference is, according to Rotton and Cohn (2004) that, whereas the NAE model attributes decline in offending at high temperatures to a decrease in the number of offenders, routine activity theory suggests that it is also due to a decrease in the number of potential victims. The need to gain a better understanding of the mechanisms linking crime and weather and temporal variations is also pointed out by Hipp, Bauer, Curran, and Bollen (2004) who show evidence in support of both aggression and routine activity theories being able to explain seasonal variations in levels of violent crime.

3. Framing São Paulo as case study

This section sets out a framework for the case of São Paulo. Firstly, it examines how weather and temporal variations influence levels of homicides over time. Second, it demonstrates what the geography of clusters of homicide look like in different seasons and suggests possible reasons behind such seasonal differences.

3.1. Variations of levels of homicides

The Tropic of Capricorn passes through the city of São Paulo (latitude: 23°32’S, longitude: 46°37’W), Brazil’s financial and economic centre. The climate is tropical, with an average temperature of 18.2°C, and rain all through the year (between 1250 and 2000 mm). São Paulo has a relatively high average temperature all year round. It is unlikely therefore that an increase in temperature alone would be the stressor responsible for significant increases in people’s propensity to commit acts of violence in the hot months of the year (Fig. 1A) as advocated by the general aggression model.
It is likely that heat would be perceived more as a stressor in countries with a temperate climate (Fig. 1B) than in tropical countries (Fig. 1A). Heat would have a deterrent effect on lethal violence after reaching a peak in countries with large oscillations in temperature (Fig. 1D). In tropical countries, a rise in levels of violence would be imposed by changes in human activity patterns that coincide with heat periods (Fig. 1C).

In São Paulo, the changes in people’s routine activity over time are more important for explaining variations in levels of homicides. This is created by people’s patterns of activities during their free time, particularly in the hot months of the year when most people take vacation or have periods away from work (Christmas, New Year’s Eve, school breaks). Remember that as many as 72% of homicides in São Paulo take place outdoors (see Appendix A). In December, for instance, people go out on the streets carrying money and stay outdoors until the late evening since commercial areas stay open until 10 pm or later. Moreover, in late summer the most popular festivities in Brazil, the Carnaval, generates gatherings that commonly engender friction and an increase in cases of violent incidents. It is expected therefore that people’s activities during their free time are responsible for generating more violent encounters during long, hot days than at any other time of the year (Fig. 1C).

**Hypothesis 1.** Assuming principles of routine activity theory, it is likely that homicides in São Paulo will occur mostly during people’s free time. Since long holidays and vacations are concentrated during the hot months of the year, more people are expected to be killed during these months. Pursuits outside work or school at the evenings, weekends and vacation will create conditions conducive to violence.

It is very difficult to disentangle the explanations of aggression models from routine activities theory when patterns of free time coincide with periods of high temperature. While each of these theories suggests a positive relationship between seasonal temperature changes and oscillations in violence, it is possible to track differences in their predictions by looking closely to the case of São Paulo. The general aggression theory suggests that the higher the temperature, the greater the number of killings. Thus, extra cases of homicides should occur from December to February. However, since the average temperature in São Paulo is high in early spring (minimum temperature 18.2°C) or late autumn (minimum temperature 20.5°C), an increase in temperature in the summer should have little effect on homicide levels because people are already used to the heat. When modelling, the variable temperature should therefore not be significant as a predictor of homicides. However, it is possible that the principles of aggression model theory alone cannot explain these findings. One strategy for assessing which of these theories contributes most to the explanation of high homicide levels is to:

1. Check if more cases of homicides take place during the hot season (late spring, summer, and early autumn) rather than only during the summer. The first evidence would corroborate the routine activity principles (since there being more people on the streets would increase the probability of having violent encounters), whilst the second one would support the general stress theory, in which only the extremely high temperatures of summer would lead to more stress and more violence.

2. Verify the frequency of killings by hour (e.g. morning, evening, day, week) and week (e.g. first week of each month is payday). If homicides occur mostly during people’s free time (e.g. evenings, weekends and vacation time), there would be strong evidence in favour of routine activity theory. However, the general stress theory/NAE model could be reinforced by São Paulo evidence if it were evidenced that more killings occurred during weekday evenings, when the temperature is lower than afternoons, for instance.

3. See if there is coherence within the significant variables when modelling levels of homicide. This would require evidence of significant temporal variables (such as holidays, paydays) that support routine activity theory, whilst the significance of weather covariates alone would be the key to support the general aggression model, particularly heat.

4. Test the contribution of each group of variables (temporal and weather) to the model separately. If the contribution of weather variables is minimal to the model when controlling for temporal variables,
then very little support can be given to the general stress theory and vice versa.

(5) Test interaction between weather and temporal covariates. If the interaction term is significant, this suggests the combined effect of both theories.

3.2. The geography of homicides

The literature that deals with explanations for intra-urban differences of homicide rates often neglects any effect of weather and temporal covariates. The pioneer study by Harries, Stadler and Zdorkwski (1984) however shows that the relationship between temperature and assaults is stronger in poor than rich neighbourhoods in Dallas, Texas. This study suggests that intra-urban variations in assaults are related to the fact that weather extremes heighten the stress levels of individuals who already live in highly stressful conditions. It was stated that in poor areas individuals have fewer resources (e.g. air conditioning, swimming pools) to avoid weather extremes than those living in rich neighbourhoods; a suggestion that was later contested by Harries and Stadler, (1988) and Rotton and Cohn (2004). As these previous studies illustrate, any attempt to explain how homicides are affected by interactions between a city’s structure and weather/temporal variations becomes a difficult task for two reasons: First, weather variables are not collected using the same geographical framework as demographic and socio-economic data. Even assuming that the effect might be different from area to area, it does not mean that all individuals living in the same geographical area would be equally affected by the same weather conditions. Second, according to Harries (1997), ‘weather conditions vary on a regional basis, and are insufficiently extreme in many areas to warrant the suggestion that deviant behaviour could be explained to any degree by...’

European and American criminology research has revealed strong associations between the geography of homicides and poverty, the presence of marginalized subcultures, drugs and Police practices (e.g. Baumer, Lauritsen, Rosenfeld, & Wright, 1998; Heimer, 1997; Krivo & Peterson, 1996). Findings from pioneering studies in Brazil corroborate the importance of social disorganization risk factors as determinants of homicide rates at intra urban level (Cardia, Adorno, & Poleto, 2003; Carneiro, 1999). Ceccato, Haining, and Kahn (2005) found an ‘inverted doughnut-pattern’ for homicides in São Paulo. The concentration of killings was found in the most central areas of the city and in a set of surrounding clusters scattered in the peripheral areas. High homicide rates however were explained as a function of differences in the city’s land use, socio-economic structure and other criminogenic conditions triggered by other crimes (such as drug-related offences and availability of weapons on the streets).

Rotton and Cohn (2004) show that frequency of assaults increases linearly with temperature in environments that had some form climate control (and decline after peaking at moderately high temperatures in settings that lack climate control). There are few reasons to believe that this would apply to São Paulo, where only a minority of households have access to climate-controlled conditions.

Hypothesis 2. Although weather and temporal variations may find different expressions in different parts of the city, high homicide areas will show stable clusters of killings regardless of the season.

Routine activity theory suggests that the more outdoor activities by individuals participated in the more criminal opportunities are presented. Therefore, areas that concentrate attractors such as meeting places provide more opportunities for criminal acts (Miethe, Hughes, & McDowall, 1991). Hence clusters of homicides include the CBD—central business districts, and major regional commercial centres. Places close to local bars, bingo halls or other meeting places are known by the Police to be hot spots for homicides (only 10% of the killings analysed in this study take place at home). These killings often happen not far from people’s home regardless of the time of year (a sample of 2671 homicide records from Instituto Médico Legal shows, for instance, that in São Paulo victims were either killed in the district where they live (51%) or close by (24%).

Hypothesis 3. According to routine activities theory, areas with a concentration of entertainment and commercial establishments will become hotspots for homicides. Since these areas are more exposed to violent atmospheric anomalies’ (p.140). Therefore, with the restrictions of the data available to this study and by using a space–time cluster technique, this analysis is limited to hypotheses that deal only with the exploratory detection of seasonal clusters of homicides in São Paulo. It is expected that high homicide areas, particularly in the periphery, will show stable clusters of homicides throughout the year. This means that the number of intra-urban units which are homicides hot spots is similar regardless of the season. This pattern is expected for two reasons. First, São Paulo experiences relatively little variation in temperature from month to month (average temperature above 18.2 °C) and therefore little seasonal change in the geography of homicides is expected. Second, part of high homicide areas are constituted by places of extreme poverty, where social mechanisms that discourage violence are lacking (Heimer, 1997; Krivo & Peterson, 1996; Sampson & Wilson, 1995) and have very few chances of improving in a short period of time, from summer to winter, for instance. Moreover, people who live in these areas experience constantly stressful conditions regardless of the season (e.g. Harries, Stadler, & Zdorkwski, 1984). This does not necessarily mean that poverty or inequality is a source of stress that leads to aggression as suggested by supporters of general aggression theory. It is likely that inequality is a proxy for greater numbers of possible offenders, and hence such as measure could also capture the effects suggested by routine activity theory (Hipp et al., 2004).
encounters during people’s free time, these areas will have seasonal fluctuations in homicides (hotspots for homicides will increase in size during vacation time and shrink afterwards).

Neighbouring polygons to high homicide areas will be incorporated in the existent hot spots because they share similar criminogenic conditions at certain times of the year, in this case during the hot months of the year. Changes in patterns of human activity during the winter (particularly the unstructured ones) will make people living in these neighbouring areas less vulnerable to violence and therefore, the size of the clusters will decrease during this time of the year.

4. Definitions and data quality

The definition of homicide, in accordance with the Brazilian Penal code, is ‘to kill someone’ (art. 121). In this study, homicide is defined as an intentional killing that is documented as homicidio doloso. Rape, robbery, deaths that take place in confrontation with the police, death from a traffic accident and so-called ‘found bodies’ are all classified separately from homicidio doloso. There have been cases where it has been difficult to determine the circumstances that led to the death of a victim, and therefore there may be cases wrongly categorized as homicidio doloso. This includes expressive crime, which is an act carried out with the sole intention of being violent. However, there may also be cases of instrumental homicide, that is, when a violent act is carried out with a specific aim, such as robbery. Another factor is that police records may record the location of where a victim is found, however it does not necessarily follow that this is the murder scene but could simply be where the body was dumped.

Compared to other offence statistics, homicide is one of the most accurate because there is no significant variation in the way it is classified and it is cross-reported by several authorities. Despite this fact, such statistics are far from being problem free. Although evidence shows that the Police homicide data is class and gender-biased (Caldeira, 2000) a main problem is that it underestimates the number of incidents. One reason is that the data recorded by the Police refers to the event, and not the total number of victims, yet multiple murders (chacinas) constitute 3% of total homicides a year in São Paulo. Another reason is that victims murdered following a robbery or similar violent crime (e.g. rape followed by death) or after confrontation with the Police are not recorded as homicides. A further difficulty occurs when the cause of death is unclear, as in the case of ‘found bodies’. In these cases, it is difficult to determine if the individual was killed by natural causes or as a result of violence. For a review on quality of crime data in Brazil, see Carneiro (1999) and Ceccato, Haining, and Kahn (2005).

Despite these structural distortions, the way that the statistics are systematized by the Police in São Paulo has improved during recent years. One example is the implementation in 2000 of INFOCRIM—an automated system of criminal records at coordinate level run by the Public Security Secretary of São Paulo – which is the source of the data used in this study (Appendix B). In the case of homicide, the INFOCRIM database contains data on the nature of the offence, the name of the victim and where the victim was found. INFOCRIM also includes temporal data: the approximate hour, day and year the murder occurred as well as the time the offence was reported. However, the estimated time of death is often similar to that recorded for when the offence occurred. The reason being that it is difficult to ascertain an exact time for a murder without a close examination of the body. Such inconsistencies in the data are identified by a group of experts that run the data control every day by sampling 10% of all crimes. As a result of such improvements in quality control procedures, data mismatching has fallen from around 70% to approximately 18%, since INFOCRIM’s implementation, generating a robust and reliable basis for any analysis.

Whilst temporal variables were created based on available calendars, the weather data comes from the most reliable source of information on weather in the country, INMET—Instituto Nacional de Meteorologia (Appendix B). Data is from the station 83781, Latitude: 23:30S, Longitude: 46:37W. Two entries of wind directions were missing for year 2001 and were therefore replaced by the corresponding days from the following year.

5. Is there a season for homicides in São Paulo?

The differences in levels of homicides by season is tested using one-way Anova with a post doc Scheffe test for 1 year data sample, from December 2000 to November 2001 (Table 1). The levels of homicides in the winter differ from the ones found for the hot months of the year (summer and autumn). The average of homicides by day drops from 16.86 in the summer to 13.90 in the winter, about 3 homicides a day less. A similar pattern is also found for the period 2001–2002. Unlike findings from countries in the Northern Hemisphere (e.g. Hakko, 2000), in São Paulo, summer alone does not contain the greatest number of homicides, which is confirmed by earlier studies in Brazil (Beato Filho et al.,1999; Kahn, 2003; Lima, 1998).
Table 1
Differences in homicides by season, from December 2000 to November 2001

<table>
<thead>
<tr>
<th>Season</th>
<th>Homicides mean</th>
<th>F-test</th>
<th>Scheffe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autumn (1)</td>
<td>16.46</td>
<td>4.96*</td>
<td>1.2/2.4</td>
</tr>
<tr>
<td>Winter (2)</td>
<td>13.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring (3)</td>
<td>14.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summer (4)</td>
<td>16.86</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at 99% level.

Most killings take place during the weekends, particularly in the evenings. As much as 54% of all homicides reported between 2000 and 2002 took place at weekends and 44% between 20.00 and 02.00 (Fig. 2).

Conflicts often reach a peak when people meet each other in their free time, at evenings or weekends. In the case of São Paulo, most homicides take place outdoors (Appendix B) but close to home. Harries (1997) suggests that there might be a ‘lag effect’ on people’s manifestation of stress. The stress is accumulated during the day and then ‘blows up’ later, for instance, when people go somewhere else after work. If this is true, this is an indication that it is not only heat that leads to stress and violence but the possibility of externalizing it by changing settings, for instance, from work to a bar.

The previous findings are important for two reasons. First, the fact that more killings take place during the hot months of the year (not only summer) corroborate even though only partially the suggestions made in Hypothesis 1. Second, most homicides take place during people’s free time (evenings, weekends and vacation: summer/early autumn), which lends weight to the argument that the increased number of homicides in the hot months of the year is more strongly related to variations in people’s activity patterns than changes in weather conditions. Aren’t weather conditions important to affect seasonal levels of homicides? In the next section, an attempt is made to refine the analysis by using three different models to test the importance of weather covariates to explain levels of homicide.

6. Modelling homicides in São Paulo

This analysis compares results from three regression models (Ordinary Least Squares) with number of homicides as the dependent variable, and weather (e.g. temperature, relative humidity) and temporal (e.g. weekend, nighttimes, payday) variables as explanatory variables using data from July 2001 to June 2002. Appendix B shows definitions of all the variables, their sources and an explanation of the coding process used in the analyses. The effects of these covariates are assessed by hours of the day, days of the week and weeks of the year.

In the first model, homicide data was calculated by intervals of 8 h. This interval was chosen because weather variables were available three times a day at 12.00, 18.00 and 24.00 h. The second model, homicide data was added up by day of the year (N = 365) whilst in the third model, the data was totalled by weeks (N = 52) and then regressed by weather variables (average of 7 days) and temporal covariates (Fig. 3).

Model 1– daily data—three times a day: Homicide data was recorded three times a day and later linked to weather measures in the model. Homicides that took place between 07.00 and 15.00 h are linked to the figures collected at 12.00 UTC, whilst data recorded between 15.00 and 22.00 h is linked to the figures taken at 18.00 UTC. Finally, homicides that occurred between 22.00 and 07.00 h are linked to the figures taken at 24.00 UTC. The number of hours of sunlight was dropped from the analysis because it was highly correlated with temperature. Multicollinearity in the model was controlled by the VIF number (a usual threshold is 4.0). The independent variables contained in the final data file included dummies for weekends, major holidays, payday, sheaf coefficients for month of the year, wind speed, temperature, cloud coverage, visibility, air pressure and relative humidity (Table 2).

First, the model without the temporal variables was used to check how much variance could be explained by weather variables only. At this stage, a sequence...
variable was added to the model to control for linear trend but was later excluded because it is not significant at 10%. Out of all the weather variables, only temperature is significant. Along with the inclusion of temporal variables, the Goodness of fit of the model increased and, with the exception of holidays, all temporal variables were significant. Hot evenings during warm weekends tended to show the most concentrated number of murders in São Paulo. The sign of the regression coefficient was only unexpected in the case of the payday variable, which was negative. The initial hypothesis based on previous evidence was that the arrival of pay cheques could be associated with an increase in certain types of violence (see, e.g., Cohn & Rotton, 2003a). The coefficient is not positive perhaps due to the fact that only some social benefits in Brazil are paid out during the first seven working days of the month and this may not affect significantly the flow of cash on the streets of São Paulo. Another explanation is that conflicts may arise precisely when monthly salaried people suffer money problems toward the end of each month.

Second, the order of entry for the variables was reversed, by inputting temporal before weather factors. In this model, payday is no longer significant and is excluded from the model as well as the sequence variable. Although weather variables slightly improved the model’s overall performance (from $R^2 = 48.1–48.8\%$), few are significant. The greater the temperature and cloud coverage, the higher was the number of homicides, particularly in the evenings during weekends. These findings illustrate the importance of temporal variables in relation to weather covariates to explain variation in levels of homicides, and consequently, the relevance of routine activity theory in detriment to principles of general aggression theory for the case of São Paulo.

Model 2—daily data: Homicide data from July 2001 to June 2002 have been aggregated by day whilst daily averages have been calculated for the weather variables from the exact same period of time (exceptions are hours of sun light and precipitation that were originally daily

---

If the number of homicides goes up over 1 year, then an analysis of homicide might show a possibly spurious relationship between month and homicides. One way to control for this is to add a sequence variable to control for linear trend. For models 1 and 2, the variable has the value of 1 on July 1, 2001 and continues sequentially through a value of 365 on June 30, 2002. Similarly, in model 3, the model has a sequence variable for each week from 1 to 52.
measures). The correlations between the independent variables have been checked and only the variables of sunlight and temperature were correlated (hours of sunlight was excluded from the analysis). The independent variables contained in the final data file include dummies for weekends, major holidays, pay day and month of the year, daily average of wind speed, daily average temperature, daily average of cloud coverage, daily average visibility, daily average air pressure and daily average humidity. As in model 1, how much variance could be explained by weather variables only was checked. The temporal variables were then added and the results compared. There is no trace of a linear trend in this model either, which has been tested using a sequence variable. The goodness of fit of the model is 67% and variables for holidays, weekends and months of the year are significant. The procedure has been repeated inputting the temporal variables into the model first. Although the order of variables has changed, the results remain the same. The variable for holidays is significant in both models confirming the effect of changes in patterns of human activity during these breaks. Holidays may prompt situations of conflict that would not occur if people are engaged in structured settings (e.g. at work). This is consistent with predictions based on routine activity theory, which suggests that holidays, particularly major ones, are more likely to affect people’s activity patterns. Holidays, such as Christmas in Brazil, mean large gatherings of family and friends, which combined with alcohol consumption (an intervening factor here) can lead to an increase in situations conducive to conflict and violent behaviour.

<table>
<thead>
<tr>
<th>Model 1—data aggregated three times a day</th>
<th>0.14 + 0.16 Temp* + 2.16 Weekends* - 0.51 Payday + 2.73 Night + 0.25 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather variables → temporal variables</td>
<td>(0.25) (6.46) (10.25) (-1.95) (13.38) (2.57)</td>
</tr>
<tr>
<td>t-values in brackets</td>
<td>R² × 100 = 48.1% R² (adjusted) × 100 = 22.7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 1—data aggregated three times a day</th>
<th>1.70 + 0.12 Temp* + 0.11 Cloud coverage* - 0.02 Relat humidity** + 2.31 Weekends* + 2.91 Night + 0.30 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal variables → weather variables</td>
<td>(1.37) (3.40) (11.41) (-2.45) (-1.95) (13.78) (2.86)</td>
</tr>
<tr>
<td>t-values in brackets</td>
<td>R² × 100 = 48.8% R² (adjusted) × 100 = 23.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 2—data aggregated by day</th>
<th>12.73 + 3.08 Holidays** + 8.49 Weekends* + 1.37 Months*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather variables → temporal variables</td>
<td>(43.64) (2.49) (15.75) (5.58)</td>
</tr>
<tr>
<td>t-values in brackets</td>
<td>R² × 100 = 67.0% R² (adjusted) × 100 = 44.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 2—data aggregated by day</th>
<th>12.73 + 3.08 Holidays** + 8.49 Weekends* + 1.37 Months*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal variables → weather variables</td>
<td>(43.64) (2.49) (15.75) (5.58)</td>
</tr>
<tr>
<td>t-values in brackets</td>
<td>R² × 100 = 67.0% R² (adjusted) × 100 = 44.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 3—data aggregated by week</th>
<th>107.15 + 10.47 Months*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather variables → temporal variables</td>
<td>(97.20) (9.41)</td>
</tr>
<tr>
<td>t-values in brackets</td>
<td>R² × 100 = 79.9% R² (adjusted) × 100 = 63.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 3—data aggregated by week</th>
<th>126.24 + 9.90 Months* - 0.26 Relat humidity**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal variables → weather variables</td>
<td>(13.84) (8.91) (-2.10)</td>
</tr>
<tr>
<td>t-values in brackets</td>
<td>R² × 100 = 81.8% R² (adjusted) × 100 = 65.5%</td>
</tr>
</tbody>
</table>

*significant at 1% level  
**significant at 5% level
This largely explains acquaintance homicides that occur in a familiar environment but not necessarily at home.

Model 3—weekly data: In the third model, homicide data has been added by weeks \((N = 52)\) and then regressed by weather variables (average of seven days) and temporal covariates. Visibility has been dropped from the analysis since it correlates with humidity, maximum temperature and air pressure (greater than 0.5).

As in models 1 and 2, temporal variables (dummy for holidays and sheaf coefficients for months of the year) are included in the model after checking how much variance could be explained by weather variables. As much as 66% of variation in homicides is explained by a selected group of weather variables: cloud coverage, number of hours of sunlight and temperature. The temporal variables were included later and the results have been compared. The inclusion of temporal variables increases the goodness of fit from 66% to 80% but the variables for cloud coverage, hours of sunlight and maximum temperature are no longer significant. When weather variables are added to the model after the inclusion of the temporal ones, \(R^2\) increases slightly by 8%. As in model 1, relative humidity proves to be significant together with months of the year.

6.1. Implications of the results

In all three models, temporal variables are far more influential when explaining levels of homicide than are weather variables. The strategy of reversing the order of entry for variables into the models proves efficient to flag for evidence in favour of routine activity theory and to the detriment of general aggression model.

Based on these two contrasting theories, the coherence of the significant variables in the model has been assessed. Since few weather variables (cloud coverage, relative humidity and particularly temperature) are significant in half of the OLS models it is not possible to discard the possibility, although small, that level of homicides can be explained by both theories. If this is correct, the greater the temperature and cloud coverage, the higher will be the number of murders, particularly in hot evenings during weekends. Equally important to temperature is relative humidity, or lack of it as a source of stress. According to INMET (2003) São Paulo has periods of 10 to 15 days with relatively high temperature and low humidity, which, in combination with high pollution levels (coming mostly from traffic), could affect a person’s health. However, it is more likely that the significant weather covariates surrogate the behaviour of variables that are not present in the model and are correlated with the hottest months of the year. When the interaction effect between temperature and months of the year is tested in model 3, for instance, it turns out not to be significant at 5%.

7. Spatial-time variations of homicide at intra-urban level

A cluster detection technique has been applied to investigate high rates of homicide. In order to detect changes over time and space in the geographical clustering of homicide, Kulldorff’s scan test has been used (SaTScan version 4.0.7. Kulldorff, 1997). This technique has a rigorous inference theory for identifying statistically significant clusters (Haining & Cliff, 2003). The test here uses the Poisson version of the scan test since under the null hypothesis of a random distribution of offences (with no area-specific effects) the number of events in any area is Poisson distributed. This test adjusts for heterogeneity in the background population. The space–time scan statistic is defined by a cylindrical window with a circular geographic base and with height corresponding to time. This cylindrical window is moved in space and time, so that, for each possible geographical location and size, it also visits each possible time period. An infinite number of overlapping cylinders of different size and shape are obtained, jointly covering the entire study region, where each cylinder reflects a possible cluster.

The space–time scan statistics are used in a single retrospective analysis using data from 1st January 2000 to 31st December 2002. Three years dataset has been collapsed into ‘one year’. All space and time dimensions of the data have been kept (by day and location) except ‘year’. This means that all homicides that took place, for example, on the 3rd April between 2000 and 2002 have been added to one year (e.g. 2000). This procedure has been used for two reasons. The first reason for adding 1 year on top of another is to look for seasonal clusters that occur let’s say every August or from February to March. If there are such clusters, there will be a higher power to pick them up with the collapsed data than 1 year dataset. Another reason is that homicide rates in São Paulo have been declining during the last 5 years and in order to have a more robust section of the data, 3 years have been collapsed into one (a common practice in classical spatial cluster analysis, see, e.g. Ratcliffe & McCullagh, 2001). The temporal scanning window is defined by setting a number of varying parameters in SaTScan. The first is the time interval, or the time aggregation level. In this case, the parameter is set to 7 days, so that the data is shown by weekly counts, and any clusters detected are a collection of consecutive weeks. In this case, the clusters are no longer than a week. Another parameter is the maximum temporal window size. This parameter is set to 28 days (and with a time interval of 7 days) so SaTScan has looked for clusters of 7, 14, 21 or 28 days. Fig. 4 shows that three significant clusters took place in the summer. The ‘most likely’ cluster, showing at 99%, is located in the South of

---

7Available at http://www.satscan.org/
São Paulo, whilst the secondary clusters are found in the northern and eastern parts of the city.

To test seasonal variations in homicides, the 3 years of dataset have been merged into one and an analysis for each season has been run. Whilst not as ideal as carrying out only one analysis without a cut-off, this allowed the detection clusters using short periods of time. Fig. 5 shows the significant clusters (‘most likely’ and ‘secondary’) for each season. There are considerable differences in patterns of homicides from season to season, particularly between summer and winter as already suggested in Section 5. However, the general pattern is similar for the ‘most likely’ clusters, which are composed of disadvantaged areas to the South. In winter and spring, clusters of homicides are concentrated in eastern, southern and city core. In summer and autumn, secondary clusters become much more scattered towards central and north-eastern areas. These results corroborate Hypothesis 3 that suggests that hotspots of homicides increase in size during vacation time and shrink afterwards, during winter and spring. These findings indicate that people are more likely to remain outdoors for longer during lengthy summer days in areas that are most conducive to criminal activities, typical of central or regional local centres. The scattered pattern in the summer and autumn months may indicate the co-existence with other criminogenic conditions that occur in the more central areas, such as those generated by commercial areas, places of entertainment and drug related markets that attract economically motivated crimes. Clusters of high homicide areas shrink during winter and spring, and human interactions resulting in violence are concentrated mostly in the poor areas of the periphery. As initially expected, the clusters of high homicide areas are stable over time.

8. Final considerations

Homicides take place mostly at evenings, during weekends, in the hot months of the year (summer and autumn). These findings are consistent with the paradigm of routine activity theory for violent crimes, which suggests that conflicts often peak during people’s free time. Long days during the hot season, for instance, mean that people are more in contact with each other and the likelihood of violent encounters is greater as suggested in Hypothesis 1.

Although the significance of few weather variables in the three models flags for the importance of the general aggression model, findings provide more support to the principles for the theory of routine activity in the Brazilian case. One reason for this is that, compared to the weather variables, the temporal ones make a more significant contribution to explaining the variance in the models (when temporal variables are added to the model first, the Goodness of fit is often better and the contribution of weather variables is marginal). Another reason is that temporal variables are significant in all six models whilst weather covariates are significant in only half of them. Finally, although not explored fully in this analysis, there might be traces of interaction between weather and temporal covariates (beyond the ones tested in this study), which could confirm the combined effect of both models. Future research should further deal with the problem of unravelling the effect weather and temporal covariates have on crime levels. The interaction between these covariates will perhaps require testing of new theoretical frameworks that go beyond the mechanisms suggested by general aggression model or routine activity theory.

As stated in Hypothesis 2, space–time clusters at intra-urban level show that most of the killings throughout the year occur in high homicide areas. Although their location is similar and stable, hotspots of homicides increase in size during vacation time and shrink afterwards, during winter and spring as suggested in Hypothesis 3. These findings indicate that people are more likely to remain outdoors for longer during lengthy, hot days in areas that are most conducive to criminal activities, typical of central areas, regional local centres but also disadvantaged areas. Poor neighbourhoods are more likely to have high levels of homicide, regardless of the time of the year. As in any other large city, disadvantaged areas in São Paulo have criminogenic conditions that
make conflicts and violence part of everyday life. These are long lasting, triggered by dispute over scarce resources (Cardia et al., 2003), repression by the Police (Chevigny, 1999), or simply by the coexistence with weapons on the streets and other crimes, such as drug related offences (Ceccato, Haining, & Kahn, 2005).

This article makes contributions to the way relationships between weather/temporal variations and crime are analysed. This study is innovative in its exploration of different scales of analysis in the modelling section and in providing further evidence of the so-called Modifiable Area Unit Problem—MAUP in regression analysis (see, e.g. Fotheringham & Wong, 1991). Another important feature of this study is the creation of space–time clusters of homicides over a period of time using a new georeferenced crime database, Geographic

Fig. 5. Space–time significant clusters at 99% in São Paulo by season, 2000–2002.
Information System (GIS) technology and Kulldorff’s scan test, until now nonexistent in the literature. The analysis shares however limitations with other analyses of crime–weather/temporal relationships that are important to mention here. Although this paper recognizes that other factors could explain levels and geography of homicides in São Paulo, this study was only able to assess the effects of weather and temporal variations. Thus, any discussion that involves attempts to explain intra-urban differences in homicides in relation to weather/temporal variations and socio-demographic and criminal conditions can only be explorative. Another limitation is that the modelling section is based on a 1-year database, which is too narrow a time period for drawing final conclusions on the relationship between homicide levels and weather covariates. For future research, one of the main challenges is to elucidate the processes through which weather/temporal variables interact and influence levels of homicides using long-term data series. The inclusion of air pollution levels and life style indicators, such as alcohol consumption, would certainly shed light on the effect of these intervening factors on homicide levels. Despite these limitations, I believe the results from this study can enhance current research on relationships between homicides and weather/temporal variations by providing empirical evidence from a city in the Southern Hemisphere. In this context, it is important to report the findings from São Paulo which, until now, has been nonexistent in such research.

Acknowledgements

This research was undertaken while Vânia Ceccato was a visiting fellow at the Department of Geography at the University of Cambridge, England. The support of the Marie Curie Fellowship Scheme (Grant reference HPMF-CT-2001-01307) and STINT—The Swedish Foundation for International Cooperation in Research and Higher Education (Dnr PD2001-1045)—are gratefully acknowledged by the author. Vânia would also like to express her thanks to the Secretaria de Segurança Pública de São Paulo and the Instituto Nacional de Meteorologia for providing the dataset used in this analysis. The author wishes to thank Ellen Cohn and Dietrich Oberwittler for reading and making suggestions to an earlier version of this article as well as the suggestions made by the anonymous reviewers.

Appendix A. Homicides by selected places in São Paulo, 2000–2002 (Table A1)

<table>
<thead>
<tr>
<th></th>
<th>House</th>
<th>Bars, restaurants and recreation places</th>
<th>Streets, traffic lights</th>
<th>Transport nodes</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer</td>
<td>445</td>
<td>83</td>
<td>2898</td>
<td>54</td>
<td>579</td>
<td>4059</td>
</tr>
<tr>
<td>Autumn</td>
<td>453</td>
<td>112</td>
<td>3165</td>
<td>48</td>
<td>656</td>
<td>4434</td>
</tr>
<tr>
<td>Winter</td>
<td>408</td>
<td>104</td>
<td>2774</td>
<td>39</td>
<td>536</td>
<td>3861</td>
</tr>
<tr>
<td>Spring</td>
<td>398</td>
<td>104</td>
<td>2831</td>
<td>42</td>
<td>621</td>
<td>3996</td>
</tr>
<tr>
<td>Total</td>
<td>1704</td>
<td>403</td>
<td>11.668</td>
<td>183</td>
<td>2392</td>
<td>16.350</td>
</tr>
</tbody>
</table>

*As many as 2.3% of records have no information of place of occurrence.

Appendix B. Description of the data set (Table A2)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Crime variable</td>
<td>Homicide records from July 2001 to June 2002</td>
</tr>
<tr>
<td></td>
<td>Homicide records per day from July 2001 to June 2002</td>
</tr>
<tr>
<td></td>
<td>Source: INFOCRIM, Secretaria de Segurança Pública de São Paulo</td>
</tr>
<tr>
<td>2. Temporal variables</td>
<td>Month of the year (sheaf coefficient)</td>
</tr>
<tr>
<td></td>
<td>Dummy for major holidays (1 = New Year’s eve, New Year’s day, Christmas eve, Christmas day, Independency day, Republic proclamation day, Carnaval, Tiradentes, 0 = otherwise) models 1 and 2</td>
</tr>
<tr>
<td></td>
<td>Dummy for weekend (1 = Sunday, Saturday, 0 = otherwise) models 1 and 2</td>
</tr>
<tr>
<td></td>
<td>Dummy for pay check days (first seven office days of each month)</td>
</tr>
<tr>
<td></td>
<td>Dummy for night time (10 p.m. to 07 a.m.)</td>
</tr>
<tr>
<td>3. Weather variables</td>
<td>Temperature (°C) from July 2001 to June 2002</td>
</tr>
<tr>
<td></td>
<td>Wind speed (m/s)</td>
</tr>
<tr>
<td></td>
<td>Relative humidity (%)</td>
</tr>
<tr>
<td></td>
<td>Precipitation (rain, in mm)</td>
</tr>
<tr>
<td></td>
<td>Air pression (mb)</td>
</tr>
<tr>
<td></td>
<td>Sun light (hours)</td>
</tr>
<tr>
<td></td>
<td>Cloud coverage (dec)</td>
</tr>
<tr>
<td></td>
<td>Visibility (m)</td>
</tr>
<tr>
<td>4. Control variables</td>
<td>Variable containing 1–365 days—models 1 and 2</td>
</tr>
<tr>
<td></td>
<td>Variable containing 1–52 weeks—model 3</td>
</tr>
</tbody>
</table>
References


