

Violence and the Limits of Urban Resilience Capacity: A Spatial and Spatio-temporal Analysis of Homicides in the Metropolitan Region of Campinas, Brazil (2017-2022)

João Pedro Toledo Tricoli de Lucas

pedrottlucas@gmail.com

Pontifícia Universidade Católica de Campinas

Juan Arturo Castañeda Ayarza

Pontifícia Universidade Católica de Campinas

Research Article

Keywords: Crime Determinants, Spatial Econometrics, Bayesian Hierarchical Modeling, Exploratory Spatial Data Analysis, Social Vulnerability, Urban Resilience

Posted Date: August 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10769155/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Urban violence remains one of the principal obstacles to building resilient, sustainable cities, particularly in unequal metropolitan regions. This study analyzes the spatial and spatio-temporal distribution of homicides in the Metropolitan Region of Campinas (RMC), Brazil, between 2017 and 2022, relating it to socioeconomic conditions measured at the census-tract level from the 2022 Demographic Census. The analysis implements Anselin's (2005) four-stage spatial econometric protocol on a complete panel of 1,150 georeferenced homicides: OLS regression on an empirical-Bayes-smoothed rate, global and local diagnostics of spatial dependence (Moran's I and LISA), and Spatial Lag and Spatial Error models selected via Lagrange Multiplier tests. Because local empirical Bayes smoothing proved undefined for nearly half of the eligible tracts, a Bayesian hierarchical model (BYM) estimated directly on raw counts with a population offset was also fitted, together with a space-time extension comparing 2017–2019 and 2020–2022. Residual spatial autocorrelation was strong (Moran's I = 0.138, $p < .001$), and the Spatial Lag Model outperformed OLS and the Spatial Error Model (AIC = 14,394.4 vs. 14,465.7 and 14,405.2). The Bayesian model, estimated over a larger sample ($n = 3,644$ tracts), identified sex ratio, dependency ratio, and household income as robust correlates of homicide risk, with temporal persistence in risk confirmed by the space-time extension ($\rho_t = 0.42$, 95% CI [0.29, 0.55]). High-High risk clusters concentrated in southern Campinas near Viracopos, while Low-Low clusters characterized the region's northwestern and eastern fringes. These findings support territory-sensitive public security policy in the RMC.

INTRODUCTION

Urban violence is one of the most persistent and complex phenomena of contemporary societies, constituting a structural obstacle to sustainable development and the consolidation of resilient cities. As Naef (2022) observes, urban resilience must be understood as a political and social process aimed at strengthening cities' capacity to face shocks and stresses, including those arising from inequality and violence. In this context, homicides stand out as an indicator of social fragility and institutional vulnerability, making them a relevant dimension of urban resilience.

The international literature has shown that resilience is not limited to reacting in the face of crises, but involves the continuous adaptation of the social, economic, and spatial structures that sustain urban life, broadening the debate beyond the environmental dimension to incorporate aspects of security, social cohesion, and governance. Coaffee and O'Hare (2008) highlight that the resilient-cities agenda has come to guide urban policies aimed at mitigating risks and reorganizing spaces in response to contemporary threats, whether natural disasters or manifestations of insecurity.

In Latin America, these discussions gain relevance in the face of rapid urbanization, socio-spatial inequality, and institutional fragility, and violence must be understood as a structural element that compromises urban resilience, affecting the balance between its environmental, social, and institutional dimensions.

In Brazil, the relationship between inequality, urbanization, and crime was analyzed by Gaulez and Maciel (2010), who identified significant effects of income and demographic density on crime incidence, reinforcing the importance of a spatial approach. This perspective aligns with the methodology of Anselin (2005), whose spatial analysis framework allows the identification of autocorrelation patterns across territorial units and provides support for territory-sensitive public policies.

The literature on violence tends to be anchored at the municipal level of aggregation, as demonstrated by analyses conducted in Brazil by Barros et al. (2020), who identify socioeconomic determinants of crime using municipalities as the unit of analysis, and by studies applied in Rio Grande do Sul, where the spatial diagnosis of crime is also carried out at the municipal scale (Fochezatto et al., 2021). However, intramunicipal-scale approaches reveal important nuances that are lost at more aggregated levels. The analysis developed by Dirago et al. (2024) for Chicago, for example, demonstrates that social vulnerability and armed violence show strong spatial cluster concentration at the census-tract level, reinforcing the importance of fine-grained units of analysis for territorial diagnosis.

This scale enables more precise diagnoses of urban vulnerabilities and provides public managers with a foundation for security and local planning policies aligned with the perspectives of resilient cities, incorporating violence as one of its key dimensions.

The Metropolitan Region of Campinas (RMC) constitutes a relevant field for this analysis. With strong economic dynamism and significant socio-spatial heterogeneity, the RMC comprises twenty municipalities of distinct characteristics, with Campinas as its university, industrial, and technological hub. This centrality, combined with the territory's internal inequalities, makes the study of violence distribution relevant for understanding the relationships between urbanization and vulnerability.

Accordingly, the general objective of this study is to examine the relationship between socioeconomic conditions at the census-tract level and the spatial distribution of homicides in the Metropolitan Region of Campinas between 2017 and 2022. More specifically, the study aims to: (i) map the spatial density of homicides; (ii) construct socioeconomic and demographic indicators by census tract; (iii) estimate an ordinary least squares (OLS) linear regression model and test spatial autocorrelation; (iv) identify local patterns of spatial association (LISA); (v) specify and estimate formal spatial econometric models (Spatial Lag and Spatial Error) following the Lagrange Multiplier decision protocol; and (vi) estimate a Bayesian hierarchical spatial model, extended to a space-time formulation, that models raw homicide counts directly and is therefore more robust to the sparsity of rare events than pre-smoothed rate approaches, completing the four-stage analytical sequence proposed by Anselin (2005).

The underlying premise is that urban violence reflects structural inequalities and, at the same time, limits cities' adaptive capacity. As argued by Naef (2022), strengthening urban resilience requires integrating security, equity, and spatial planning, reducing vulnerabilities and broadening social cohesion. Thus, the territorial analysis of crime, particularly at the census-tract scale, constitutes a strategic instrument for understanding local fragilities and guiding public policies aimed at building safer and more sustainable cities.

THEORETICAL FRAMEWORK

2.1 Violence and Territorial Inequality

Urban violence, especially when observed through the lens of homicides, is associated with socioeconomic inequalities and spatial fragmentation in the urban environment. Gaulez and Maciel (2010) demonstrate that structural factors, such as income and demographic density, exert a positive influence on crime in the state of São Paulo, evidencing the role of socioeconomic conditions and uneven urbanization as determinants of violence, consistent with the broader finding that urban scale and density amplify crime incidence (Glaeser & Sacerdote, 1999).

Beato Filho and Reis (2000) highlight a paradox. Despite continuous improvements in societal well-being, including higher literacy levels, democratization, income growth, and inflation control, crime rates have persisted and continued to rise. These authors emphasize that understanding crime is enriched by a spatial approach capable of identifying territorial patterns and concentrations of vulnerability, moving beyond purely aggregate analyses.

In the same direction, Ye and Wu (2011) show, through the analysis of homicides in Chicago, that the spatial distribution of crime tends to follow persistent structures over time associated with residential segregation and social marginalization. According to these authors, spatial panel models enable the understanding of the temporal persistence of urban violence and the influence of local contexts on its reproduction. Territorial inequality thus manifests not only in income differences, but also in the unequal distribution of risk and exposure to violence.

The literature on resilience and urban planning reinforces that violence constitutes an indicator of social and institutional fragility whose territorial incidence reflects failures in urban governance (Naef, 2022). Crime therefore acts as a barrier to the construction of sustainable and equitable cities, as it amplifies feelings of insecurity, weakens the social fabric, and limits equitable access to urban spaces.

Understanding violence from a territorial perspective means recognizing that urban inequalities are also expressed spatially through the concentration of risks and opportunities. Spatial analysis, in this sense, makes these asymmetries visible, identifying areas where social vulnerabilities translate into greater exposure to violence.

2.2 Urban Resilience

Holling (1973) introduced the concept of resilience from ecology, describing a system's capacity to reorganize after a disturbance while maintaining its essential functions. This framework was gradually incorporated into urban sciences, which began to understand resilient cities as those that not only resist shocks, whether physical, social, or economic, but also learn and reorganize despite their occurrence and persistence.

Coaffee and O'Hare (2008) argue that urban planning policies have come to incorporate security as a strategic axis, reorganizing city design in the face of contemporary risks, which range from environmental threats to manifestations of urban violence.

Similarly, Naef (2022), in analyzing the 100 Resilient Cities initiative, draws attention to the role of violence as a structural vulnerability that weakens governance and compromises the pursuit of equitable cities. This reading aligns with the approach of Becker (1968), according to whom crime results from a rational calculation of expected costs and rewards. In contexts marked by inequality and institutional fragility, the perceived risk tends to diminish while the potential gains from criminal activity expand, favoring its occurrence.

Promoting resilient cities therefore requires confronting violence and inequality as structural components of urban planning and social sustainability.

Within the operational framework proposed by Naef (2022) for the 100 Resilient Cities initiative, urban resilience is assessed across multiple interdependent dimensions, including health and wellbeing, economy and society, infrastructure and environment, and leadership and strategy, of which public safety constitutes a core component of the health and wellbeing pillar. Homicide incidence is among the most direct and internationally comparable indicators used to operationalize this security dimension, since it captures both the lethality of violence and its territorial concentration, without requiring the compound weighting schemes used for broader multidimensional indices. The present study accordingly restricts its empirical claims to this single, well-measured dimension of resilience, territorial exposure to lethal violence, rather than to the construct in its entirety. The relationship between homicide risk and other resilience dimensions, such as institutional trust or economic diversification, falls outside the scope of the spatial analysis conducted here and is noted as a boundary condition for interpretation in Section 5.

2.3 Spatial Econometrics in the Study of Urban Violence

The incorporation of spatial methods in the analysis of urban phenomena represents a relevant methodological contribution. Anselin (2005) established the foundations of spatial econometrics by demonstrating that the independence of observations, a central assumption of traditional models, is rarely verified in territorial phenomena. In urban contexts, spatial units tend to interact, generating autocorrelation patterns that may be positive (concentration) or negative (dispersion).

The Global Moran's I measures the degree of spatial autocorrelation across the full set of observations, while the Local Indicator of Spatial Association (LISA) identifies local clustering patterns of high or low incidence of a given phenomenon (Anselin, 2005).

Complementarily, spatial regression models, such as the Spatial Lag Model (SLM) and the Spatial Error Model (SEM), allow the estimation of the effects of spatial dependencies on the variable of interest, offering more consistent interpretations than traditional OLS linear regression (LeSage & Pace, 2009). Beyond the frequentist spatial-regression tradition, Bayesian hierarchical formulations such as the Besag-York-Mollié (BYM) model (Besag et al., 1991) decompose spatial dependence into structured and

unstructured random effects estimated directly from event counts, offering a further alternative particularly suited to rare, spatial sparse events such as homicides in small areal units. This class of models has recently been used to delineate violent-crime clusters at fine spatial resolution outside Brazil as well, for instance in an application to South Korea (Lee & Lee, 2020).

In the context of this study, spatial analysis is applied at the census-tract scale, enabling a more granular reading of intra-urban inequalities. The choice of this scale is grounded in its capacity to capture local heterogeneities that are lost at more aggregated levels. It also opens the possibility of governmental action at the intramunicipal scale, focusing resource allocation on specific localities within municipalities.

Taken together, these spatial techniques broaden the understanding of urban violence by articulating statistical data with territorial dynamics. This approach brings academic analysis closer to the concrete realities of urban planning, enabling the production of knowledge that translates into more targeted and effective public actions.

METHODOLOGY

The methodology adopted is quantitative, anchored in the framework proposed by Anselin (2005), which systematizes the process of spatial analysis into four main stages: (i) formulation of spatial hypotheses, (ii) estimation of an initial OLS linear regression model, (iii) diagnosis of spatial dependence, and (iv) specification of appropriate spatial models, as depicted in Fig. 1. The present study implements the complete four-stage protocol, estimating and comparing formal spatial econometric models (Section 3.4) and extending it with a Bayesian hierarchical alternative (Sections 3.5–3.6) suited to the sparsity of the event data documented in Section 4.1.

The analysis is applied to the Metropolitan Region of Campinas (RMC), composed of twenty municipalities, using homicide microdata from the São Paulo State Public Security Secretariat (SSP-SP) for the period 2017–2022. Socioeconomic indicators and census-tract boundaries are derived from the 2022 Demographic Census of the Brazilian Institute of Geography and Statistics (IBGE). The census tract is adopted as the minimum unit of observation, in line with the recommendation of Anselin (2005) for high-granularity territorial studies. This choice, like any areal partition of continuous urban space, is not neutral: results can be sensitive to the specific zoning and scale of the areal mesh, a limitation long recognized in spatial analysis as the Modifiable Areal Unit Problem (MAUP) and acknowledged here rather than resolved.

3.1 Data Source and Treatment

Homicide microdata was compiled from the São Paulo State Public Security Secretariat (SSP-SP) for 2017–2022, combining six annual data files into a single harmonized panel. Coordinate fields were standardized across years to accommodate differences in field naming between source files, and municipality names were matched using accent- and case-insensitive normalization to ensure consistent

identification of all twenty RMC municipalities. This process yields 1,150 georeferenced homicides across the six-year period, of which 1,116 (97.0%) were assigned to one of 6,547 census tracts across all twenty RMC municipalities through a point-in-polygon spatial join. Twenty-eight points falling exactly on a shared tract boundary were assigned to a single tract to avoid double counting. Census tract boundaries were obtained directly by IBGE municipality code and downloaded individually by municipality, improving robustness to name-matching errors.

A separate computational constraint affected the treatment of the census socioeconomic source files. The five tables used to construct the independent variables (income, demography, household characteristics, literacy, and race/ethnicity) are distributed by IBGE at the national level, together covering approximately 458,000 census tracts. Loading these files directly, in full, was computationally infeasible on standard hardware, since the underlying worksheet markup decompresses to several gigabytes per file. A dedicated preprocessing routine therefore streams each source file row by row, retaining only roughly 6,500 tracts within the RMC's twenty municipalities before any further processing. This reduces peak memory usage without altering the underlying census values, which were cross-checked against the official IBGE data dictionary.

3.2 Dependent and Independent Variables

Two complementary specifications of the dependent variable are used across the models presented in Section 4. For the ordinary least squares baseline (Section 4.1) and the formal spatial regression models (Section 4.4), the dependent variable is the census-tract homicide rate adjusted by local empirical Bayesian smoothing (EBL), applied following Anselin (2005) and Marshall (1991), building on the earlier disease-mapping tradition of Clayton et al. (1987), to reduce rate instability in tracts with small populations by combining each tract's observed rate with the expected mean of its neighbors. The logarithmic transformation of the smoothed rate is used for regression, which helps approximate residual normality and attenuate the influence of extreme values. For the Bayesian hierarchical models (Sections 4.5–4.6), the dependent variable is instead the raw homicide count at each tract, modeled directly via a Poisson likelihood with an offset for resident population (Section 3.5). This specification avoids the pre-smoothing step altogether and, as shown in Section 4.1, proves substantially more robust under the sparsity of the event data analyzed here.

The independent variables are organized into five thematic axes derived from the 2022 IBGE Census: (i) income, (ii) demography, (iii) housing characteristics, (iv) literacy, and (v) race/ethnicity. Each axis aggregates proportional indicators, ratios, or synthetic indices constructed at the census-tract level, as shown in Table 1 and detailed, with their exact construction formulas, in Appendix Table A1. These five axes reflect what the Demographic Census can measure at census-tract resolution; environmental-criminology covariates commonly used elsewhere, such as policing intensity, informal-economy activity, or drug-market presence, are not routinely collected at this spatial granularity in Brazilian public data and could therefore not be included, a data-availability constraint rather than a theoretical omission.

Table 1
Independent variables remained in the final specification (Section 3.2).

Source: Authors' own elaboration based on the 2022 Demographic Census.

Axis	Variable	Expected sign
Income	Log of average monthly income of household head	negative
Demography	Proportion of males aged 15 to 29 (p_homens_15_29)	positive
Demography	Sex ratio, males/females (sex_ratio)	positive
Demography	Dependency ratio, (0–14 + 60+)/15–59 (dependency_ratio)	positive
Housing	Mean household size (hh_size_mean)	mixed
Housing	Presence of penitentiary/detention center (0/1)	positive
Literacy	Illiteracy rate among males aged 15 to 29 (p_analf_m_15_29)	positive

The seven-variable specifications in Table 1 results from a stepwise backward selection procedure applied to a pool of twenty-six candidate indicators spanning the five thematic axes described above. As a robustness check, this specification was also evaluated against the full candidate pool directly. Because several literacy and race-composition candidates exhibit substantial census confidentiality suppression, requiring simultaneous non-missing values across all twenty-six candidates before selection reduces the usable sample to fewer than forty tracts. Excluding the four candidates with the most severe suppression restores a workable sample of 1,046 tracts. Under this alternative pool, however, stepwise selection retains only illiteracy among residents aged 60 and older and two mutually collinear racial-diversity indices (variance inflation factors near 18), none of which remain statistically distinguishable from zero once re-estimated within the Bayesian hierarchical framework of Section 3.5. Given this instability and the weaker correspondence of the alternative specification to the income, demographic, and educational determinants established in the literature reviewed in Section 2, the seven-variable specification in Table 1 is retained as the primary model throughout, and the comparison is reported as a limitation in Section 5.

Subsequently, spatial dependence in the OLS residuals was diagnosed using Moran's I and LISA analysis, in accordance with the methodological protocol of Anselin (2005). Based on these results, formal spatial econometric models (Section 3.4) and Bayesian hierarchical models (Sections 3.5–3.6) were estimated for substantive interpretation.

3.3 Spatial Dependencies: Global Moran's I and LISA

The measurement of spatial dependencies is a central step in the analytical framework proposed by Anselin (2005). Socioeconomic and territorial phenomena rarely distribute independently in space; instead, they exhibit spatial autocorrelation, the tendency of nearby areas to show similar (positive autocorrelation) or contrasting (negative autocorrelation) values. Diagnosing this dependence verifies

whether the spatial structure of the data violates the independence assumption of OLS residuals and whether spatial models are warranted.

The first step is the construction of a spatial weights matrix (W), which defines the neighborhood relationship between units of analysis. This study uses the Queen contiguity criterion, in which two census tracts are considered neighbors if they share a boundary point or edge. The matrix is row-standardized so that the sum of weights for each unit equals 1, following the standard procedure recommended by Anselin (2005) to ensure comparability across observations with different numbers of neighbors. The same weights matrix, computed once over the final analytic sample, is reused consistently across the empirical Bayes smoothing, the OLS model, the Moran's I test, and the LISA analysis, avoiding the internal inconsistency that would arise from recomputing the neighbor structure over different row subsets at different stages.

With the weights matrix defined, the Global Moran's I is calculated, expressing the degree of spatial autocorrelation across all census tracts. Positive values indicate similar spatial correlation (High-High or Low-Low), while negative values reflect dispersed patterns (High-Low or Low-High).

While the Global Moran's I evaluates dependence at the aggregate scale, LISA identifies local autocorrelation patterns, revealing specific spatial clusters and their significance levels. According to Anselin (2005), LISA helps understand the heterogeneous structure of urban phenomena, since autocorrelation may vary substantially within the same territory. Each spatial unit receives a local autocorrelation value and is classified into one of four cluster types: High-High (HH), tracts with high homicide values surrounded by similarly high areas; Low-Low (LL), tracts with low homicide values surrounded by similarly low areas; High-Low (HL), tracts with high homicide values surrounded by low-value neighbors; and Low-High (LH), tracts with low homicide values surrounded by high-value neighbors.

3.4 Formal Spatial Econometric Models: SLM and SEM

Beyond diagnosing the presence of residual spatial dependence, the protocol of Anselin (2005) prescribes a formal decision rule for choosing between the two principal spatial regression alternatives once dependence is confirmed. The Spatial Lag Model (SLM) augments the OLS specification with a spatially lagged dependent variable, appropriate when the dependence reflects genuine spillover in the outcome itself (for example, diffusion of violence into neighboring tracts), while the Spatial Error Model (SEM) instead models spatial dependence in the disturbance term, appropriate when the dependence stems from unobserved, spatially correlated factors omitted from the regressors (LeSage & Pace, 2009). Following Anselin's (2005) decision protocol, both the standard and robust Lagrange Multiplier (LM) tests for spatial lag and spatial error dependence were computed on the OLS residuals. When both standard tests are significant, the model corresponding to the larger robust (adjusted) LM statistic is preferred, since the robust tests remain valid under local misspecification of the alternative form. Both SLM and SEM were estimated by maximum likelihood using the same row-standardized Queen

contiguity weights matrix as the Moran's I and LISA diagnostics and compared to the OLS baseline via the Akaike Information Criterion (AIC) and log-likelihood.

3.5 Bayesian Hierarchical Spatial Model (BYM)

The local empirical Bayes smoothing used to construct the dependent variable in Sections 3.2–3.4 is a moment-based estimator (Marshall, 1991) that becomes numerically undefined when the local neighborhood used to compute it contains too few observed events relative to its expected variance, a condition that, as shown in Section 4.1, arises for nearly half of the eligible tracts under the sparse event counts analyzed here. As an alternative that avoids this limitation, a Besag-York-Mollié (BYM) Bayesian hierarchical model (Besag et al., 1991) was additionally fitted. The model specifies homicide counts directly via a Poisson likelihood with an offset for the (logged) resident population, and decomposes the log relative risk into a fixed-effects component (the same seven socioeconomic predictors used above, less any predictor with zero variance in the estimation sample) and two random effects: a spatially structured component, modeled with a conditional autoregressive (CAR) prior over the Queen contiguity graph, and an unstructured, tract-level heterogeneity component. Because the BYM specification operates on raw counts rather than a pre-smoothed rate, it does not inherit the empirical Bayes instability described above and can be estimated over a correspondingly larger analytic sample. The model was fit via Markov Chain Monte Carlo (MCMC) using the CARBayes package (Lee, 2013) in R, with a 20,000-iteration burn-in followed by 100,000 further iterations thinned by a factor of ten, yielding 8,000 retained posterior samples; convergence was assessed via the Geweke (1992) diagnostic. Tracts with no remaining neighbor within the analytic sample (a structural consequence of the covariate-missingness pattern rather than of the underlying census geometry, since no such isolated tracts exist when the neighbor structure is computed over the full, unfiltered tract universe) cannot be accommodated by the CAR prior and were excluded prior to model fitting (245 of 3,889 eligible tracts).

3.6 Space-Time Extension (BYM-ST)

To incorporate the temporal dimension of the six-year panel, a first-order autoregressive space-time extension of the BYM model was estimated using the CARBayesST package (Lee et al., 2018), which adds a temporally autoregressive structure to the spatial random effect, allowing risk to evolve across periods while retaining the spatial dependence structure described in Section 3.5. Given the extreme sparsity of the underlying annual event counts, the model was first estimated at annual resolution as a convergence-screening step. Convergence was judged inadequate (maximum Geweke $|z|$ statistic across parameters well above conventional thresholds), and the model was re-estimated over two triennial periods (2017–2019 and 2020–2022) with a full posterior sample of 8,000 draws (20,000-iteration burn-in, 100,000 further iterations thinned by ten), which converged satisfactorily (Section 4.6).

3.7 Use of AI Tools

An AI coding assistant (Claude, Anthropic) was used during manuscript preparation to assist with debugging and verifying the R data-processing and statistical pipeline, cross-checking reported statistics against the underlying script output, drafting and revising portions of the manuscript text and generating

Fig. 1 as an original schematic diagram. All AI-assisted content was reviewed, verified, and approved by the authors, who take full responsibility for the accuracy and integrity of the final manuscript.

RESULTS

4.1 Linear Regression Model (OLS + EBL)

Of the 6,547 census tracts in the RMC, 3,889 had complete data across the seven predictors retained in Table 1 (2,658 tracts were excluded, predominantly due to census confidentiality suppression of small counts). Within this subset, the local empirical Bayes rate could not be reliably estimated for 1,862 tracts (47.9%). The underlying moment-based estimator (Marshall, 1991) returns an undefined value whenever the local neighborhood used to compute it has too little observed variance in event counts relative to what the model expects, a condition met throughout much of the RMC given that only 615 homicides fall within this 3,889-tract subset across six years. This is a documented property of the empirical Bayes estimator under sparse counts rather than an artifact of the estimation code, and it motivates the Bayesian hierarchical alternative introduced in Section 3.5. The final OLS/EBL analytic sample comprises $n = 2,027$ census tracts.

The model specification includes the seven variables in Table 1: log of average household-head income (log_rend_resp_med), proportion of males aged 15 to 29 (p_homens_15_29), sex ratio (sex_ratio), dependency ratio (dependency_ratio), mean household size (hh_size_mean), presence of a penitentiary (has_penitentiary), and proportion of illiterate males aged 15 to 29 (p_analf_m_15_29). The presence-of-penitentiary indicator has zero variance in the final analytic sample: only one census tract in the entire RMC records any population in a penitentiary-type collective dwelling, and that tract was excluded for missing literacy data. Its coefficient is consequently not identified (Table 2).

Table 2

OLS regression results, dependent variable log(EBL-smoothed homicide rate). $n = 2,027$; $R^2 = 0.034$; Adj. $R^2 = 0.031$; $F(6, 2020) = 11.96$, $p < 0.001$; has_penitentiary listed but not estimated due to zero variance in the final analytic sample.

Source: Authors' own elaboration.

Variable	Estimate	Std. Error	t	p
(Intercept)	13.352	5.414	2.466	0.014
log_rend_resp_med	-2.444	0.461	-5.304	< 0.001
p_homens_15_29	21.184	10.820	1.958	0.050
sex_ratio	2.232	2.200	1.015	0.310
dependency_ratio	2.460	1.615	1.523	0.128
hh_size_mean	-0.511	0.665	-0.769	0.442
has_penitentiary	n/a	n/a	n/a	n/a
p_analf_m_15_29	16.447	16.449	1.000	0.317

Household-head income is negative and highly significant, consistent with the expectation that higher-income tracts experience lower homicide risk. The proportion of young males is positive and approaches, without clearly crossing, the conventional significance threshold (21.184, $p = .050$), broadly consistent with the theoretical expectation that a higher concentration of young men is associated with greater homicide risk. Mean household size, sex ratio, and dependency ratio are not statistically significant in this specification. The illiteracy coefficient for young males is positive and not significant here; this coefficient proves unstable across specifications, a pattern examined further using the Bayesian models in Sections 4.5–4.6 and discussed in Section 5.

Table 3
Variance Inflation Factor (VIF)
results. All values well below
the threshold of 5
recommended by Hair et al.
(2010), confirming no relevant
multicollinearity.

Source: Authors' own
elaboration.

Variable	VIF
log_rend_resp_med	1.60
p_homens_15_29	2.01
sex_ratio	1.33
dependency_ratio	1.37
hh_size_mean	1.39
p_analf_m_15_29	1.07

4.2 Spatial Dependencies: Global Moran's I

The Global Moran's I statistic is positive and highly significant ($I = 0.138$, standard deviate = 6.97, $p < .001$), indicating strong positive spatial autocorrelation in the OLS residuals (Table 4) and providing clear empirical grounds for the formal spatial models estimated in Section 4.4.

Table 4
Global Moran's I for OLS
regression residuals.

Source: Authors' own elaboration.

Statistic	Value
Observed Moran's I	0.1383
Expected value	-0.0014
Variance	0.0004
Standard deviate	6.973
p-value	< 0.001

4.3 Exploratory and Confirmatory Spatial Patterns: KDE and LISA

Two complementary, but methodologically distinct, spatial visualizations are presented. The kernel density estimate (KDE, Fig. 3) is an exploratory tool (Medina Ariza & Solymosi, 2023). It summarizes the

raw geographic concentration of homicide events without reference to the resident population at risk or to any test of statistical significance and should be read as descriptive rather than confirmatory. The Local Indicator of Spatial Association (LISA, Fig. 5), by contrast, is confirmatory. It is computed on the population-adjusted, empirical-Bayes-smoothed rate used throughout Section 4.1 and formally tests each tract's local autocorrelation for statistical significance against a reference distribution. The two are complementary rather than interchangeable, and the discussion in Section 5 relates their agreement, where it occurs, without treating spatial coincidence in the KDE surface as itself evidence of significant clustering.

The highest concentrations occur in the extreme south of the municipality of Campinas, near Viracopos airport, and in localities such as Campo Belo, the Industrial District, and the Monsenhor Housing Estate, areas already identified as having higher vulnerability levels by the PUC-Campinas Observatory and Fundação FEAC (2021). Cosmópolis and Paulínia also show elevated concentrations in their most urbanized areas.

Figure 4 maps the EBL-smoothed rate underlying the OLS model of Section 4.1. Its coverage, 2,027 of 6,547 tracts, or 31.0 percent of the RMC, is visibly incomplete, a direct visualization of the sparsity limitation quantified in Section 4.1; large contiguous areas of the region simply have no defined value under this specification. Figure 6 (Section 4.5) presents the corresponding Bayesian relative-risk surface, which covers 3,644 tracts (55.7 percent of the RMC) without this gap.

High-High (HH) significant clusters concentrate in areas of greater socioeconomic vulnerability, particularly the extreme south of Campinas near Viracopos airport, and in adjoining sectors of Hortolândia. High-Low (HL) transition tracts appear at the margins of these clusters. No Low-Low (LL) clusters reach significance within the smaller, EBL-defined subsample used for LISA (Table 5), a further consequence of the coverage gap discussed above.

Table 5
Mean covariate values by significant LISA cluster ($p < 0.05$), EBL-smoothed rate per 100,000 inhabitants. No Low-High (LH) tract reached significance.

Source: Authors' own elaboration.

Cluster	n	EBL rate	Income (log)	% young males	Sex ratio	Dependency	HH size
High-High (HH)	25	238.4	8.19	0.101	0.964	0.535	2.70
High-Low (HL)	26	25.0	8.33	0.099	0.940	0.551	2.69
Low-Low (LL)	5	0.0	8.60	0.087	0.813	0.680	2.37

4.4 Formal Spatial Econometric Models: SLM and SEM

Lagrange Multiplier tests computed on the OLS residuals favor a spatial lag specification over a spatial error specification. Both standard tests are significant (LM-error = 48.47, $p < .001$; LM-lag = 56.26, $p < .001$), and the robust (adjusted) LM-lag statistic (23.56, $p < .001$) exceeds the robust LM-error statistic (15.78, $p < .001$), indicating that the residual spatial dependence is better explained by genuine spillover in the outcome than by omitted, spatially correlated covariates (Table 6).

Table 6
Lagrange Multiplier diagnostics for spatial dependence (Anselin, 2005 protocol).

Source: Authors' own elaboration.

Test	Statistic	p
LM-error	48.47	< 0.001
LM-lag	56.26	< 0.001
Robust LM-error	15.78	< 0.001
Robust LM-lag	23.56	< 0.001
SARMA	72.04	< 0.001

Consistent with this diagnosis, the Spatial Lag Model (SLM) outperforms both the OLS baseline and the Spatial Error Model (SEM) on AIC (Table 7). The estimated spatial autoregressive coefficient is moderate and highly significant ($\rho = 0.234$, LR test = 73.4, $p < .001$), confirming that homicide risk in each tract is associated with risk in its immediate neighbors, net of the socioeconomic covariates. Under the SLM specification, only household-head income remains significant among the individual predictors (-2.029, $p < .001$); the coefficients for young males, dependency ratio, and the remaining variables retain the same sign as the OLS estimates in Table 2 but lose significance once spatial spillover is accounted for directly, suggesting that part of what OLS attributed to these covariates reflects spatially correlated risk rather than a purely compositional effect. Even so, a Lagrange Multiplier test for residual autocorrelation applied to the SLM residuals remains significant (test value = 21.815, $p < .001$), indicating that the spatial lag term alone does not fully absorb the dependence structure present in the data and motivating the Bayesian hierarchical alternative estimated in Section 4.5.

Table 7
Model comparison: OLS vs. Spatial Lag (SLM) vs. Spatial Error (SEM). Lower AIC indicates better fit.

Source: Authors' own elaboration.

Model	df	AIC	log-likelihood
OLS	8	14,465.7	-7,224.9
SLM	9	14,394.4	-7,188.2
SEM	9	14,405.2	-7,193.6

4.5 Bayesian Hierarchical Spatial Model (BYM)

The BYM model was estimated over $n = 3,644$ census tracts, 1,617 more than the EBL-based specifications of Sections 4.1–4.4, after excluding 245 tracts left without a remaining neighbor by the covariate-missingness pattern (Section 3.5). Table 8 reports posterior means and 95% credible intervals. Household-head income (-0.744, 95% CI [-1.103, -0.398]), sex ratio (2.933, 95% CI [1.565, 4.322]), and dependency ratio (2.140, 95% CI [1.065, 3.196]) are robustly associated with homicide risk, consistent in direction with the OLS and SLM estimates above. The proportion of young males and mean household size are not distinguishable from zero under this specification. Male illiteracy among young adults is negative and significant here (-12.769, 95% CI [-23.753, -2.406]), opposite in sign to its non-significant OLS estimate in Table 2; this discrepancy is addressed directly in Section 5.

Table 8
BYM spatial model, posterior means and 95% credible intervals. $n = 3,644$; DIC = 2,671.5; effective parameters = 466.6.

Source: Authors' own elaboration.

Variable	Mean	2.5%	97.5%
(Intercept)	-7.819	-11.603	-3.890
log_rend_resp_med	-0.744	-1.103	-0.398
p_homens_15_29	3.234	-4.130	10.405
sex_ratio	2.933	1.565	4.322
dependency_ratio	2.140	1.065	3.196
hh_size_mean	-0.265	-0.741	0.224
p_analf_m_15_29	-12.769	-23.753	-2.406

4.6 Space-Time Extension (BYM-ST)

The annual-resolution BYM-ST model did not converge (maximum Geweke $|z| = 12.8$ across parameters, well above the conventional threshold of approximately 2), consistent with the sparsity documented in Section 4.1. The annual panel comprises only 585 homicides across 3,644 tracts and six years. Aggregating into two triennial periods (2017–2019 and 2020–2022) resolved the convergence problem (maximum Geweke $|z| = 2.8$; Table 9). The temporal autoregressive coefficient is significant ($\rho_t = 0.421$, 95% CI [0.285, 0.553]), indicating that homicide risk in a tract during 2020–2022 is substantially predictable from its risk during 2017–2019, net of the spatial structure, providing direct evidence of temporal persistence in homicide risk across the study period. The spatial autoregressive coefficient is small ($\rho_s = 0.006$) once the temporal structure is included, suggesting that most of the spatial pattern identified in Sections 4.2–4.5 is absorbed by the model's spatial random-effect variance term ($\tau^2 = 2.303$) rather than by the first-order spatial autoregressive parameter specific to this space-time formulation. Household-head income, sex ratio, dependency ratio, and young-male illiteracy retain the same signs and approximate magnitudes as in the cross-sectional BYM model (Table 8).

Table 9
BYM-ST space-time model, triennial resolution
(2017–2019 vs. 2020–2022), posterior means and
95% credible intervals. $n = 3,644$ tracts $\times$ 2 periods;
585 homicides; DIC = 3,376.4.

Source: Authors' own elaboration.

Variable	Mean	2.5%	97.5%
(Intercept)	-4.005	-7.330	-0.529
log_rend_resp_med	-1.122	-1.426	-0.823
p_homens_15_29	4.063	-2.003	9.851
sex_ratio	2.513	1.294	3.704
dependency_ratio	1.720	0.785	2.641
hh_size_mean	-0.242	-0.665	0.186
p_analf_m_15_29	-9.923	-20.042	-0.504
ρ (spatial, AR)	0.006	0.000	0.023
ρ_t (temporal, AR)	0.421	0.285	0.553

DISCUSSION

The interpretation of the results obtained for the RMC allows for advancing the understanding of the relationship between social vulnerability, violence, and urban resilience. As operationalized in Section 2.2, homicide incidence is treated here as an indicator of the security dimension of urban resilience (Naef, 2022). Observing its spatialization in the RMC allows an empirical discussion of how structural elements of social fragility affect this specific dimension of adaptive capacity at the intramunicipal level, without extending the claim to resilience as a multidimensional construct.

Across every specification estimated (OLS, SLM, and the BYM cross-sectional and space-time models), household-head income is negatively and significantly associated with homicide risk, and sex ratio and dependency ratio are positively associated wherever they reach significance. Because the OLS and SLM specifications model a log-transformed, empirical-Bayes-smoothed rate under a Gaussian likelihood while the BYM specifications model raw counts under a Poisson likelihood, coefficient magnitudes are not directly comparable across these two families of models; only the direction and significance of association are compared here. This pattern is consistent with recent evidence. In Chicago, Dirago et al. (2024) find that census tracts with high social vulnerability present armed violence rates up to seven times higher than less vulnerable areas, revealing how territorial disadvantage undermines the collective capacity to respond to violence. The same pattern emerges in Fortaleza, where neighbourhoods with high vulnerability show disproportionate impacts of poverty on homicides (Fórum Brasileiro de Segurança Pública & Governo do Estado do Ceará, 2017, pp. 110–115). Barros et al. (2020) argue that poverty, family instability, and youth unemployment weaken community mechanisms of social control, amplifying violence risk. The High-High LISA clusters identified in the poorest RMC tracts (Section 4.3) can thus be read not merely as a statistical pattern but as an expression of low local social resilience.

The kernel density surface (Fig. 3) and the LISA clusters (Fig. 5) show a high degree of visual correspondence in the southern periphery of Campinas, but this correspondence should not be over-read. The KDE is, by construction, an unweighted, non-inferential description of where events happened to occur, while the LISA classification is obtained from a population-adjusted rate and a formal significance test. Their agreement in this case strengthens confidence in the substantive finding, but the two analyses answer different questions: where events are concentrated versus where the population-adjusted risk is significantly clustered. The distinction matters specifically in the tracts where they diverge, which are not highlighted by an exploratory density surface alone.

The instability of the young-male illiteracy coefficient across specifications warrants explicit discussion rather than selective reporting. Under OLS and SLM, the coefficient is positive and not significant; under both BYM specifications, estimated on a substantially larger and non-overlapping sample that does not depend on the empirical Bayes smoothing step, it is negative and significant. Three non-exclusive explanations are plausible. First, the EBL-based models and the BYM models are estimated to be different, non-random subsets of tracts (Section 4.1), so a genuine composition effect, rather than mere estimation noise, could drive part of the discrepancy. Second, the outcome differs in scale and link function. OLS and SLM models have a log-transformed smoothed rate under a Gaussian likelihood, while BYM models raw counts under a Poisson likelihood with a spatial random effect that can absorb correlated variation the linear models leave in the residual. Third, given that illiteracy is itself spatially clustered, some of the association may be confounded with the spatially structured random effect that only the BYM specification estimates explicitly. This pattern is not considered resolved by the present analysis, and any policy inference drawn from this specific coefficient should be treated as provisional pending replication on an independent panel.

Finally, the low R squared of the baseline OLS model (3.4%, Table 2) should be read within the four-stage protocol it is designed to initiate, not as a final verdict on the explanatory power of the covariates. The Lagrange Multiplier tests (Table 6) and the sizeable improvement in AIC from OLS to SLM (Table 7) indicate that a large share of the pattern in the outcome that OLS cannot explain is spatially structured rather than random noise. The BYM deviance information criterion (DIC = 2,671.5, Table 8) offers a further, non-nested measure of fit appropriate to the hierarchical specification. Read together, the four models trace a coherent progression, from a weak linear baseline, through confirmation of strong residual spatial dependence, to spatial and space-time models that substantially outperform it, rather than four independent claims of equal strength, and the discussion above and the conclusion below moderate causal language accordingly. This caution is substantive rather than merely rhetorical. None of the models estimated here can rule out reverse causality or endogenous sorting, for example wealthier or more educated households selectively relocating away from historically violent tracts, which could generate the observed income and education associations even without a direct causal effect of local conditions on homicide risk.

Thus, in elucidating the results of this study alongside similar research, a central point of convergence emerges. Unequal territories tend to exhibit lower urban resilience, expressed by the inability to mitigate and interrupt violent dynamics. This pattern appears in Chicago, Fortaleza, and Brazilian municipalities analyzed by Barros et al. (2020) and Fochezatto et al. (2021), and it reproduces itself in the RMC. The significant temporal persistence identified by the BYM-ST model ($\rho_t = 0.42$, Section 4.6) further indicates that this territorial disadvantage is not a transient feature of any single year but a structural condition that persists across the 2017–2022 period.

Therefore, the results indicate that strengthening urban resilience in the RMC, from the perspective of social sustainability, requires not only security policies but structural interventions aimed at reducing inequalities, broadening educational opportunities, strengthening community networks, and territory-sensitive urban planning that accounts for the temporal persistence, and not only the spatial concentration, of risk.

CONCLUSION

The present study aimed to analyze the spatial and spatio-temporal distribution of homicides in the Metropolitan Region of Campinas (RMC) between 2017 and 2022, evaluating their association with socioeconomic and demographic conditions at the census-tract scale. A quantitative approach was applied, integrating homicide microdata from the São Paulo State Public Security Secretariat (SSP-SP) with socioeconomic and demographic indicators from the 2022 IBGE Demographic Census across the complete six-year panel.

The results confirm the existence of significant spatial dependence in the distribution of homicides, evidenced by a positive and statistically significant Moran's I (0.138). The present study completes the four-stage protocol proposed by Anselin (2005). The Spatial Lag Model, selected via Lagrange Multiplier

tests, outperforms both OLS and the Spatial Error Model, confirming that homicide risk in a census tract is strongly associated with risk in its immediate neighbors. A Bayesian hierarchical extension (BYM), motivated by the empirical Bayes instability documented in Section 4.1, corroborates the role of household income, sex ratio, and dependency ratio over a substantially larger analytic sample, and its space-time extension (BYM-ST) provides direct evidence of temporal persistence in risk across the study period.

The local analysis, through the LISA indicator and the BYM relative-risk surface, reinforced these findings by identifying significant spatial clusters, particularly in the southern periphery of Campinas near Viracopos airport and in adjoining sectors of Hortolândia, characterized by high homicide risk and equally critical neighboring tracts (High-High).

This research contributes to three aspects. First, by offering an intramunicipal analysis that uses the census tract as the unit of observation to reveal internal heterogeneities frequently invisible in more aggregated analyses, subsidizing public policies aimed at reducing territorial inequalities. Second, by documenting a memory-efficient, streaming approach to processing nationwide census microdata that generalizes to other fine-grained administrative-data pipelines facing similar computational constraints. Third, and most centrally, by completing the four-stage spatial econometric protocol proposed by Anselin (2005) and extending it with a Bayesian hierarchical formulation better suited to the sparsity of homicide counts at fine spatial resolution.

Three limitations should guide the interpretation of these results and the design of future research. First, the resident-population denominator used throughout is drawn from the single cross-section of the 2022 Census and applied uniformly across the 2017–2022 panel. To the extent that population changed differentially across tracts over this period, rates and risks in earlier years are approximated rather than measured directly. Second, the sign instability of the young-male illiteracy coefficient across model specifications, discussed in Section 5, remains an open question rather than a settled finding; future work should replicate this specific association on an independent panel, ideally with population denominators available at each survey year, before it is used to inform policy targeting. Third, the associations reported throughout are not identified as causal effects. Socioeconomic composition and homicide risk are plausibly jointly determined, for instance through selective migration into or out of high-risk tracts, and the current data do not support an instrumental-variable or panel fixed-effects strategy capable of separating this reverse pathway from a direct effect of local conditions on violence. Beyond these three points, subsequent studies may extend the space-time modeling to a finer temporal grain if future data collection reduces event sparsity and may explore the mechanisms of violence diffusion implied by the significant spatial-lag coefficient identified in Section 4.4.

Declarations

Funding

The authors received institutional scholarship support for this research. Full acknowledgment is provided on the separate title page, in accordance with GeoJournal's double-blind review policy.

Competing Interests

The authors declare no competing interests.

Author Contributions

Per GeoJournal's editorial policy, author contribution statements are submitted directly through the journal's online submission system and are not included in the manuscript file.

Ethics Approval

This study uses publicly available, aggregate, and anonymised data from the São Paulo State Public Security Secretariat (SSP-SP) and the Brazilian Institute of Geography and Statistics (IBGE). No individual-level or identifiable personal data were used. Accordingly, no formal ethics approval was required for this research.

Data Availability

Homicide data are publicly available from the São Paulo State Public Security Secretariat (SSP-SP): <http://www.ssp.sp.gov.br>. Socioeconomic and demographic data are publicly available from the 2022 Brazilian Demographic Census (IBGE): <https://www.ibge.gov.br/censo2022>. All spatial and statistical analysis was conducted in R (packages sf, spdep, spatialreg, CARBayes, and CARBayesST).

References

1. Anselin, L. (2005). *Exploring Spatial Data with GeoDa (TM): A Workbook*. Center for Spatially Integrated Social Science.
2. Barros, P. H. B., Baggio, H. S., & Baggio, I. S. (2020). Os determinantes socioeconômicos da criminalidade no Brasil: O papel das externalidades e heterogeneidades espaciais. *Revista Brasileira de Segurança Pública*, 14(2), 188–209. <https://doi.org/10.31060/rbsp.2020.v14.n2.1091>
3. Beato Filho, C. C., & Reis, I. A. (2000). Desigualdade, desenvolvimento socioeconômico e crime. In R. Henriques (Ed.), *Desigualdade e pobreza no Brasil* (pp. 385–405). IPEA.
4. Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217. <https://doi.org/10.1086/259394>
5. Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1), 1–20. <https://doi.org/10.1007/BF00116466>
6. Clayton, D., & Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43(3), 671–681. <https://doi.org/10.2307/2532003>

7. Coaffee, J., & O'Hare, P. (2008). Urban resilience and national security: The role for planning. *Proceedings of the Institution of Civil Engineers – Urban Design and Planning*, 161(4), 173–182. <https://doi.org/10.1680/udap.2008.161.4.173>
8. Dirago, C., Poulson, M., Hatchimonji, J., Byrne, J., & Scantling, D. (2024). Geospatial analysis of social vulnerability, race, and firearm violence in Chicago. *Journal of Surgical Research*, 294, 66–72.
9. Fochezatto, A., Petry, G., Braatz, J., & Marcondes, H. R. (2021). Investimento em educação diminui a criminalidade: Análise dos municípios do Rio Grande do Sul usando econometria espacial. *Administração Pública e Gestão Social*, 13(4). <https://doi.org/10.21118/apgs.v13i4.11239>
10. Fórum (2017). Brasileiro de Segurança Pública, & Governo do Estado do Ceará. *O cenário da violência e da criminalidade no Brasil e no Ceará: Análise comparativa* (Vol. 2). Secretaria da Segurança Pública e Defesa Social do Ceará.
11. Gaulez, M. P., & Maciel, V. F. (2010). Determinantes da criminalidade no Estado de São Paulo: Uma análise espacial de dados em cross-section. *Revista Vianna Sapiens*, 1(2), 91–120.
12. Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian Statistics 4* (pp. 169–193). Oxford University Press.
13. Glaeser, E. L., & Sacerdote, B. (1999). Why is there more crime in cities? *Journal of Political Economy*, 107(S6), S225–S258. <https://doi.org/10.1086/250109>
14. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). Prentice Hall.
15. Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
16. Lee, D. (2013). CARBayes: An R package for Bayesian spatial modeling with conditional autoregressive priors. *Journal of Statistical Software*, 55(13), 1–24. <https://doi.org/10.18637/jss.v055.i13>
17. Lee, D., & Lee, D. (2020). Analysis of influential factors of violent crimes and building a spatial cluster in South Korea. *Applied Spatial Analysis and Policy*, 13, 759–776. <https://doi.org/10.1007/s12061-019-09327-1>
18. Lee, D., Rushworth, A., & Napier, G. (2018). Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package. *Journal of Statistical Software*, 84(9), 1–39. <https://doi.org/10.18637/jss.v084.i09>
19. LeSage, J., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC.
20. Marshall, R. J. (1991). Mapping disease and mortality rates using Empirical Bayes Estimators. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 40(2), 283–294. <https://doi.org/10.2307/2347593>
21. Medina Ariza, J., & Solymosi, R. (2023). *Crime Mapping and Spatial Data Analysis Using R*. CRC Press / Taylor & Francis.

22. Naef, P. (2022). 100 Resilient Cities: Addressing urban violence and creating a world of ordinary resilient cities. *Annals of the American Association of Geographers*, 112(7), 2012–2027.
<https://doi.org/10.1080/24694452.2022.2038069>
23. Observatório, P. U. C. C., & Fundação, F. E. A. C. (2021). *Proposta de índice de vulnerabilidade intramunicipal: Campinas*. PUC-Campinas.
24. Ye, X., & Wu, L. (2011). Analyzing the dynamics of homicide patterns in Chicago: ESDA and spatial panel approaches. *Applied Geography*, 31(2), 800–807.
<https://doi.org/10.1016/j.apgeog.2010.08.006>

Figures

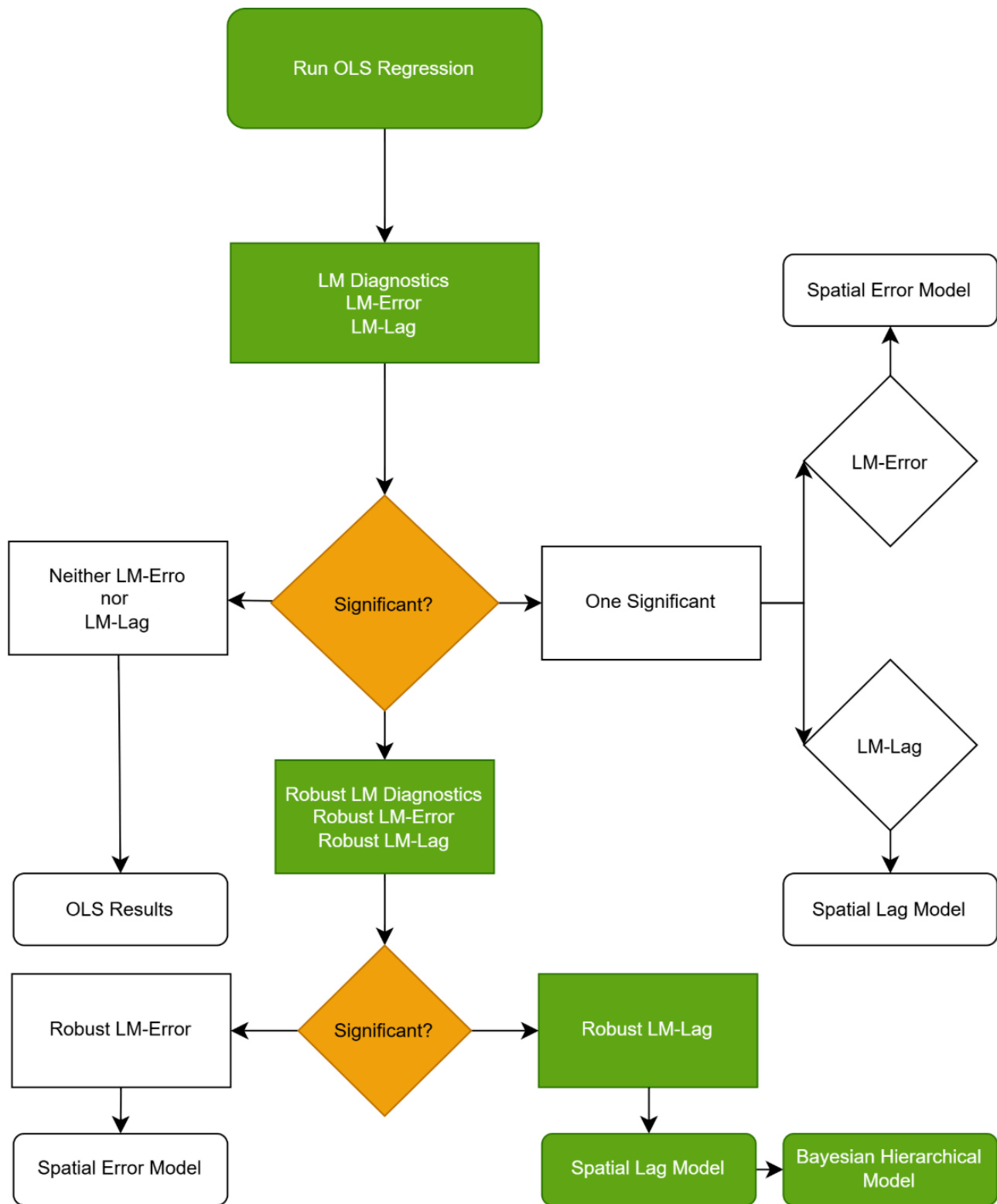


Figure 1

Spatial regression model selection protocol, following the four-stage decision framework described in Anselin (2005)

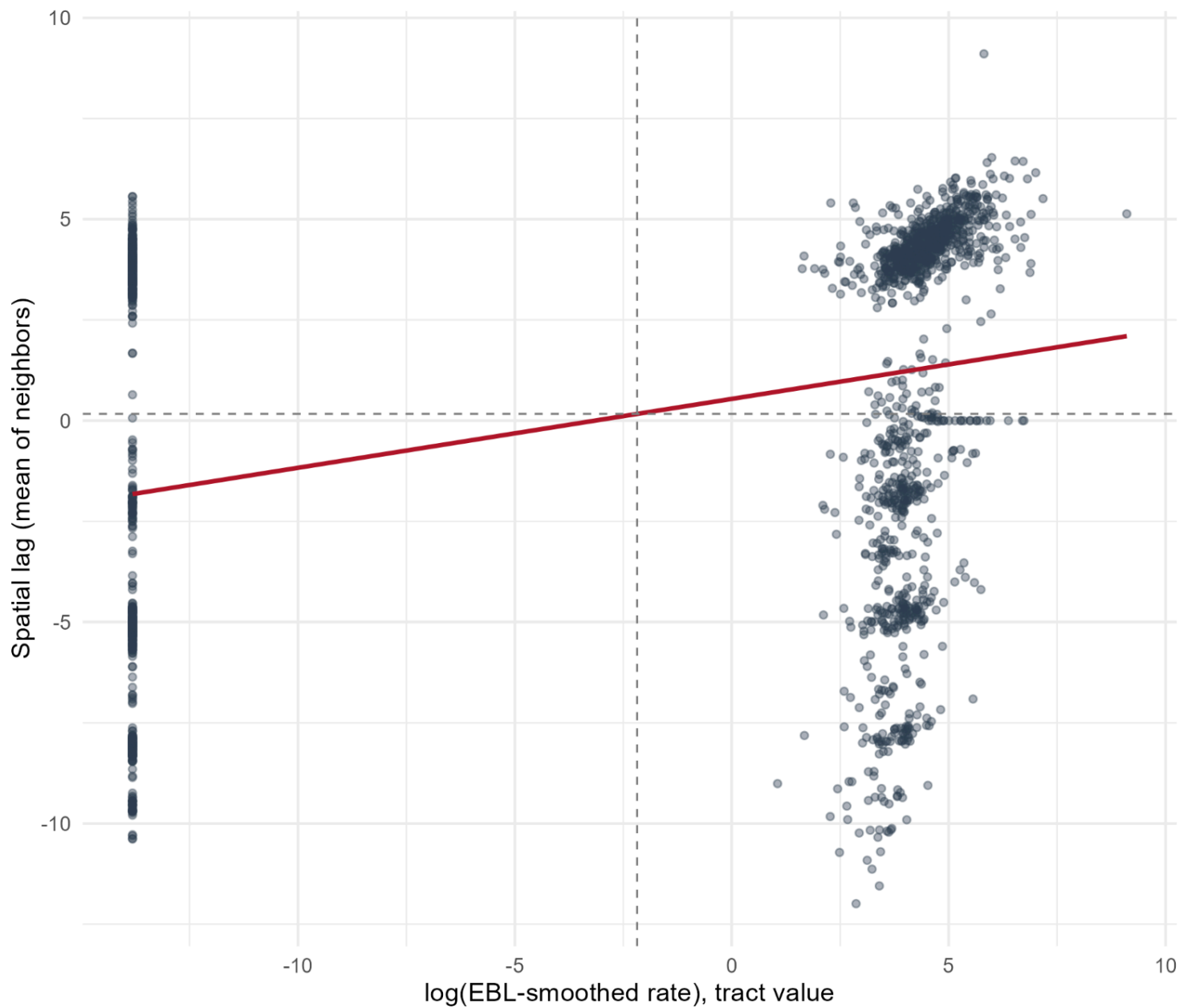


Figure 2

Moran's I scatterplot for the smoothed homicide rate (EBL, log) against its spatial lag; the positive slope confirms positive spatial autocorrelation

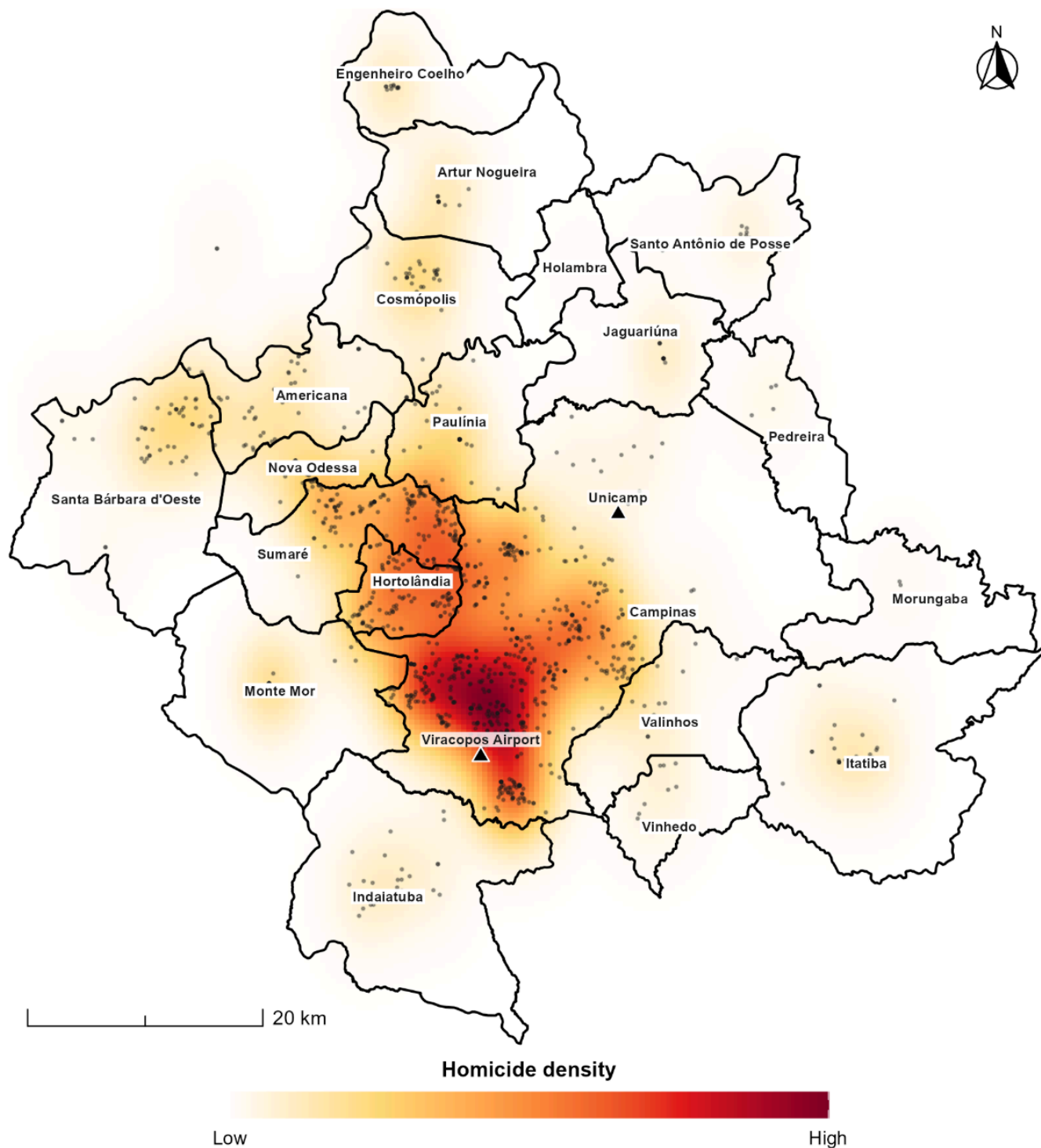


Figure 3

Kernel density of homicides in the Metropolitan Region of Campinas (RMC), 2017-2022 (1,150 georeferenced events; exploratory visualization)

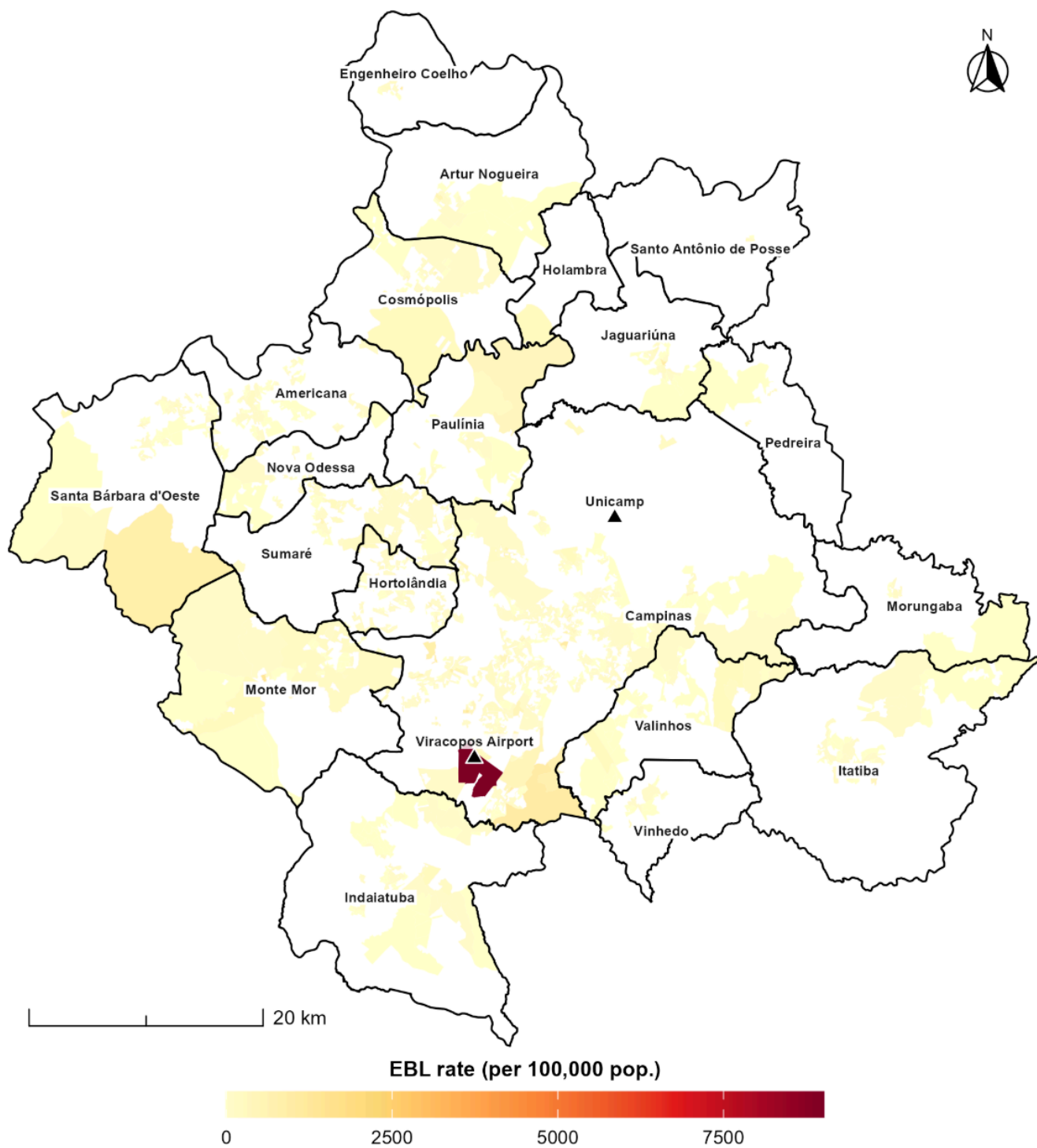


Figure 4

EBL-smoothed homicide rate by census tract; white areas indicate tracts without a defined estimate under this specification (Section 4.1)

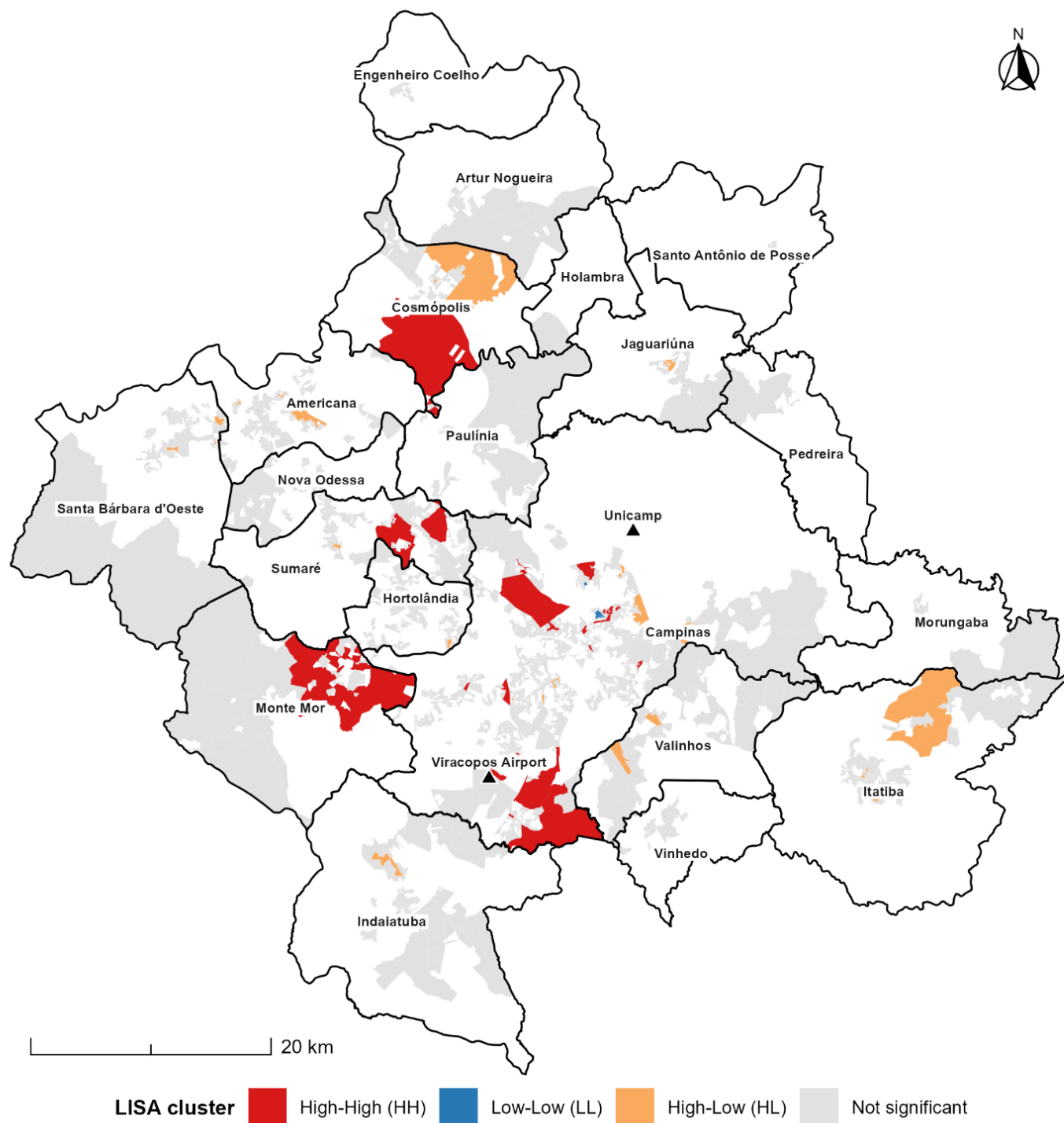


Figure 5

Local Indicator of Spatial Association (LISA) for census tracts, RMC

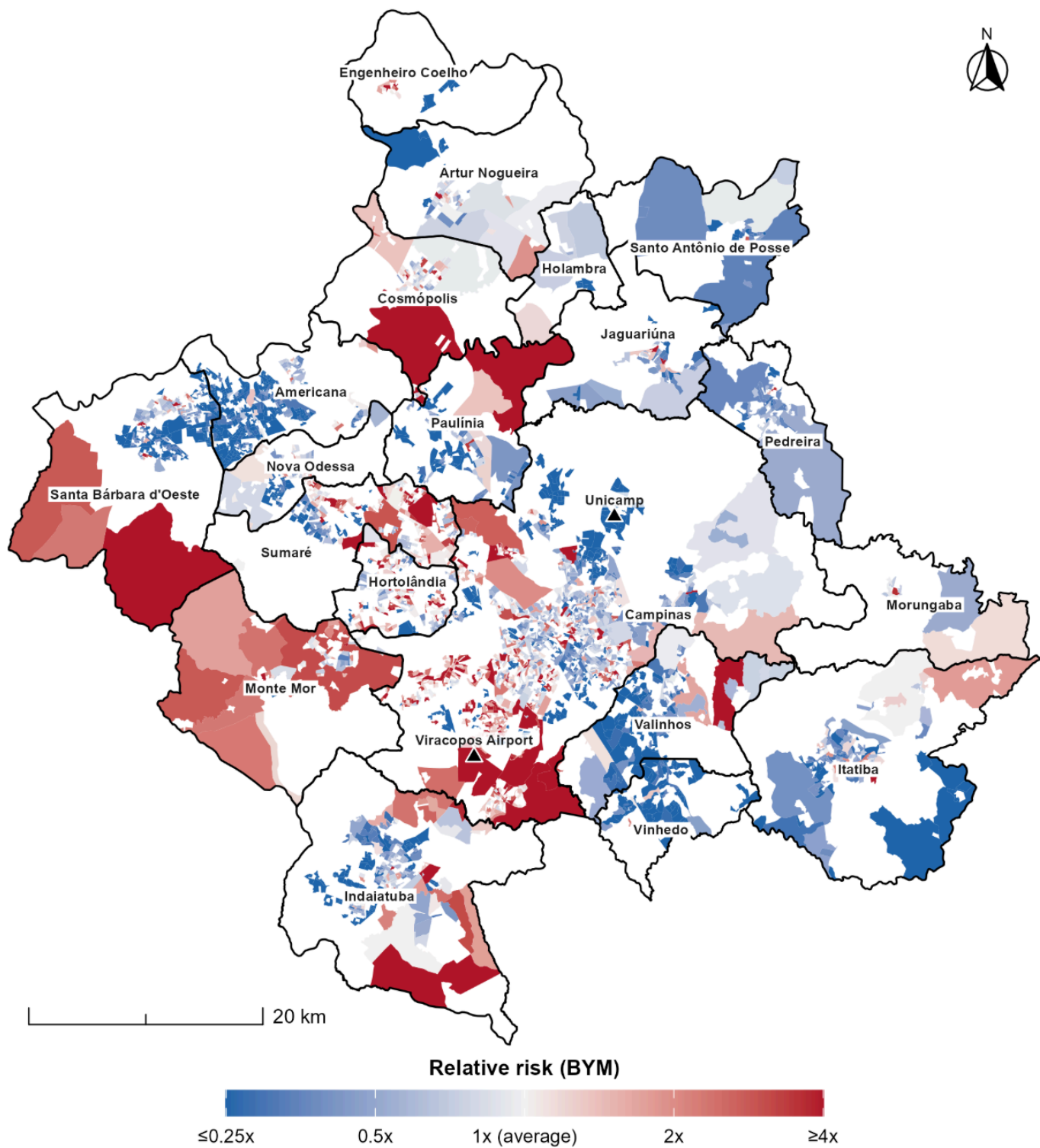


Figure 6

BYM-estimated relative risk of homicide by census tract (posterior mean, relative to the RMC-wide average)

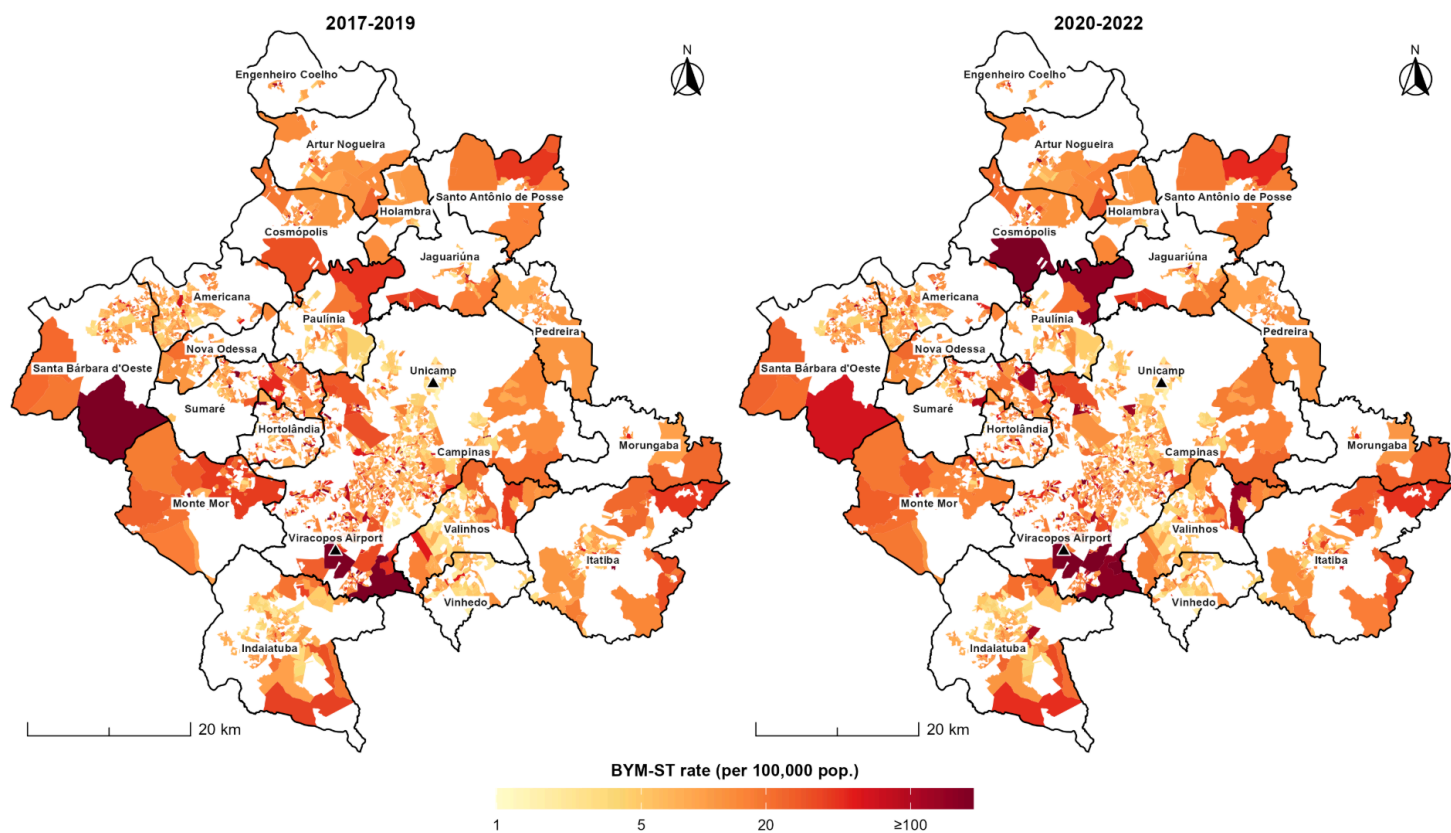


Figure 7

BYM-ST estimated homicide rate by census tract, 2017-2019 vs. 2020-2022

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [APPENDIXA.docx](#)