

Indicadores do risco: análise dos homicídios por intermédio de variáveis preditoras no município de São Paulo em 2001¹

Marcelo Batista Nery
Eduardo Henrique Geraldi Araújo
Corina da Costa Freitas
Camilo Daleles Rennó
Antonio Miguel Vieira Monteiro

Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201-970 - São José dos Campos - SP, Brasil
{mbnery, araujo}@tid.inpe.br, {corina, camilo, miguel}@dpi.inpe.br

Abstract. This study emphasizes the urban spaces and searches its differences as a way to categorize the extension of the local variability of the violence inside a metropolitan environment. The plan is to create a multiple linear regression model that helps to better understand the occurrences of homicides using several predictor variables and observe the existence of patterns and identify atypical data about this subject in São Paulo city. Social and economic data obtained from the federal and state governments were used to build up a multivariate model which was reduced and improved by tests and previous knowledge. In addition, according with the overall objective of this paper, a discussion about the crimes geographic distribution was necessary in order to make the choice of the explanatory variables more accurate and most effective.

Palavras-chave: homicides, regression model, social studies, homicídios, modelo de regressão, estudos sociais.

1. Introdução

Centrando-se nos estudos urbanos, muitos pesquisadores já buscaram desenvolver perspectivas em relação à formação, expansão, ordenação das cidades e a influência desses fluxos na organização simbólica e nas relações sociais. Assim sendo, a vida comunitária das cidades (tomada em sentido genérico), com todos os seus problemas sociais, se torna um privilegiado objeto de estudo, influenciando a realização de pesquisas sobre as partes cidadinas.

Com a violência elevada à segunda causa de óbito no país a partir de 1989 (Souza e Minayo, 1995), o Brasil tem este fenômeno social como um de seus principais problemas, expressos no campo econômico, na saúde pública e nas inter-relações humanas. Entrementes, a crescente ascensão dos homicídios nos grandes centros, torna-o um elemento que merece ser observado de modo particular.

Atentando-se a isso, tem-se aqui a menção do problema que pode ser enunciado como a inspiração inicial deste trabalho que, ao questionar dados, indicadores e suas expressões territoriais, possíveis sinais de certos estados, qualidades ou atitudes sociais, propõem alcançar um melhor entendimento das ocorrências criminais no ambiente urbano. Nesse caso, este artigo tanto se integra nas pesquisas que exploram estratégias de controle do crime – para citar apenas alguns, os trabalhos de Beato et al. (1997), Fernandes e Carneiro (1996), Carneiro (1999) e Zalar (1993) –, quanto nos trabalhos que avaliam os processos dos quais resultam os dados registrados – Beato et al. (1997) e Vargas (1993) – ou buscam padrões de criminalidade – como ocorre em Câmara et al. (2000) e Silva (2001).

Portanto, a discussão que orienta esse estudo, o que interessa, está associado à hipótese de que áreas de risco para homicídios podem ser apontadas por um conjunto de indicadores associados às variações espaciais, possíveis de serem observadas e aferidas, evidenciando características locais deste processo violento.

¹ Este trabalho foi financiado pela FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), processo nº 02/10849-9 - tipo mestrado.

2. Metodologia e Descrição das variáveis

Ao estudar os homicídios, tendo em vista suas definições, sua fenomenologia e as idéias geradas a partir das mesmas, percebe-se que diferentes instrumentos de estudo podem se tornar imprescindíveis para o entendimento destas ocorrências. Um destes recursos é a análise de regressão, uma ferramenta estatística que utiliza as relações entre duas ou mais variáveis, de tal forma que uma variável possa ser explicada pela outra (ou outras). De mais a mais, um modelo de regressão é um modo formal de expressar a tendência que uma variável dependente “Y” tem em variar com as variáveis independentes de uma maneira sistemática.

O exame das variáveis permitirá construir um modelo da distribuição dos homicídios, implicando, posteriormente, em diferenciadas formas de abordagens da representação espacial. Destarte, os dados utilizados na pesquisa foram obtidos na Base de Dados do Programa de Aprimoramento em Informações sobre Mortalidade (PRO-AIM) do Município de São Paulo, para o ano de 2001, e no Censo Demográfico de 2000 do Instituto Brasileiro de Geografia e Estatística (IBGE).

O PRO-AIM tem como fonte de informação a declaração de óbito (DO), conseqüentemente, leva em conta as mortes dos moradores da cidade de São Paulo pelo local de residência da vítima e não de ocorrência do crime. Neste trabalho, o interesse é pelos homicídios registrados. A variável dependente (Y) foi composta pela divisão entre o valor dos homicídios nos distritos censitários e a população de cada setor censitário², visto que não havia valores daquela ocorrência criminal para esta unidade territorial. Essa alternativa foi escolhida porque compreende a noção de que a cidade constitui uma lógica interna que projeta formas de organização baseadas nas relações entre indivíduos, interferindo gravemente na conduta das pessoas. Ademais, sabe-se que o crescimento populacional pode estar ligado ao aparecimento de problemas sociais que influenciam a qualidade de vida dos cidadãos.

Por outro lado, o recenseamento de 2000 incluiu dez quesitos de características do domicílio, nove quesitos para a pessoa responsável pelo domicílio ou para o morador individual em domicílio coletivo e seis quesitos para cada um dos demais membros do domicílio – sendo utilizado, principalmente, informações dos domicílios particulares permanentes (DPP). Com estes dados, foram criados 25 indicadores (**Tabela 1**) que buscaram captar a infra-estrutura local (cristalizadas em serviços públicos, como saneamento básico e tratamento de lixo, habitação, congestionamento e ambiência) e avaliar fatores socioeconômicos (como renda, instrução, longevidade e condição do chefe de família)³. A maior parte dos indicadores selecionados para este trabalho foram constituídos tendo por base os indicadores presentes em Sposati (1996). O método consistiu na produção de índices simples e representativos resultantes da combinação de variáveis primitivas, todas tendo como valor original as variáveis do IBGE.

O cálculo do índice de chefes de família com rendimento superior a 20 salários mínimos estabelece um bom exemplo de indicador que satisfaz este objetivo. Sendo a maior porcentagem de chefes de famílias com esta característica, em certa região, um apontamento de maior capacidade de inclusão social (ISO), acredita-se que deva apresentar uma relação inversa com as ocorrências de homicídios. Outros índices em que se espera a mesma correspondência são: APR, LON, CS1, CS2, AAG, ISA e TLX. Os índices que devem ter uma relação direta com os homicídios são: DHA, HPR, CPR, DAM, ATA, CMC, MNA, CAL, CLP, JU1, JU2 e CNA. Além disso, as variáveis de instrução que devem variar seu comportamento mediante a elevação ou redução dos anos de estudo 1AE, 4AE, 8AE, 11AE e 15AE.

² No presente momento está sendo desenvolvido um trabalho que toma por base dados pontuais de criminalidade, de modo a permitir a real localização das ocorrências, tanto pela residência da vítima como pelo local de ocorrência do delito.

³ Neste trabalho, a diferenciação utilizada separa campos que estão verdadeiramente inter-relacionados. Tal condição ocorre por uma opção didática, pois se acredita que desta forma é possível apresentar as idéias com maior clareza.

Tabela 1 – Descrição das variáveis preditoras.

Variável	Sigla	Descrição
Abastecimento de Água	AAG	Domicílios particulares permanentes – abastecimento de água rede geral
Alfabetização Precoce	APR	Pessoas residentes – alfabetizadas - 5 a 9 anos de idade por Pessoas residentes
Alfabetização Tardia	ATA	Pessoas residentes – não alfabetizadas - 10 a 14 anos de idade por Pessoas residentes
Chefe de família na linha de Pobreza	CLP	Pessoas responsáveis pelos DPP – rendimento nominal mensal - até 2 salários mínimos por Pessoas responsáveis pelos DPP
Chefes de família abaixo da linha de Pobreza	CAL	Pessoas responsáveis pelos DPP – rendimento nominal mensal - sem rendimento por Pessoas responsáveis pelos DPP
Concentração de Mulheres Chefes de Família	CMC	Mulheres responsáveis pelos DPP por Pessoas responsáveis pelos DPP
Condições de Privacidade	CPR	Pessoas residentes – domicílios coletivos por Pessoas residentes
Conforto Sanitário	CS1	Número médio de Banheiros por Domicílio
Conforto Sanitário2	CS2	Número médio de Pessoas por Banheiro por Domicílio
Chefes de Família não Alfabetizados	CNA	Pessoas responsáveis pelos DPP - com menos de 1 ano de estudo por Pessoas responsáveis pelos DPP
De 1 à 3 anos de estudo	1AE	Pessoas responsáveis pelos DPP - 1 a 3 anos de estudo por Pessoas responsáveis pelos DPP
De 4 a 7 anos de estudo	4AE	Pessoas responsáveis pelos DPP - 4 a 7 anos de estudo por Pessoas responsáveis pelos DPP
De 8 a 10 anos de estudo	8AE	Pessoas responsáveis pelos DPP - 8 a 10 anos de estudo por Pessoas responsáveis pelos DPP
De 11 a 14 anos de estudo	11AE	Pessoas responsáveis pelos DPP - 11 a 14 anos de estudo por Pessoas responsáveis pelos DPP
Mais de 15 anos de estudo	15AE	Pessoas responsáveis pelos DPP - 15 e mais anos de estudo por Pessoas responsáveis pelos DPP
Densidade habitacional	DHA	Pessoas residentes pelos Domicílios
Desrespeito ambiental	DAM	DPP – destino do lixo – jogado em terreno baldio, logradouro, em rio, lago ou mar pelos DPP
Habitação precária	HPR	Pessoas residentes - domicílios particulares improvisados por Pessoas residentes - domicílios particulares
Inclusão social	ISO	Pessoas responsáveis pelos DPP – rendimento nominal mensal - mais de 20 salários mínimos por Pessoas responsáveis pelos DPP
Instalação sanitária	ISA	Domicílios particulares permanentes – com banheiro ou sanitário – esgotamento sanitário – rede geral de esgoto ou pluvial
Juventude	JU1	Pessoas residentes – 7 a 14 anos de idade por Pessoas residentes
Juventude2	JU2	Pessoas residentes – 15 a 24 anos de idade por Pessoas residentes
Longevidade	LON	Pessoas residentes – 70 anos de idade ou mais por Pessoas residentes
Mulheres não Alfabetizadas	MNA	Mulheres residentes – não alfabetizadas por Mulheres residentes
Tratamento do Lixo	TLX	DPP – destino do lixo – coletado pelos DPP

3. Construção do Modelo

O primeiro passo na construção do modelo de regressão é o estudo das variáveis a serem utilizadas. Assim, a análise exploratória dos dados indicou a necessidade de algumas transformações para a diminuição na correlação entre as variáveis preditoras.

A observação da variável dependente contra a variável explicativa demonstrou que os homicídios se relacionavam exponencialmente com boa parte das variáveis. Não obstante, aplicou-se uma transformação logarítmica na base dez para este dado (representado por LHOM), ocasionando uma sensível melhora na maior parte das relações. De forma similar, bons resultados foram conseguidos ao aplicar o logaritmo neperiano aos seguintes indicadores: conforto sanitário 2, juventude e longevidade, transformando-os em LCS2, LJU1, LLON.

Selecionou-se aleatoriamente 75% dos dados para realizar novas análises, deixando o restante para a validação do modelo. Com estes, as análises voltaram a confirmar a melhora gerada pelas transformações. Contudo, algumas variáveis ainda apresentavam informações nitidamente redundantes ou de pouca contribuição na inferência dos homicídios. Deste modo, foram observadas aquelas que possuíam alta correlação com outra variável preditora e baixa com a variável dependente, o que poderia motivar a eliminação de algumas delas. Feito isto, realizou-se o teste *stepwise regression*⁴ (Neter et al., 1996) que auxiliou na escolha das variáveis pertinentes ao modelo.

Inicialmente, descartaram-se as variáveis AAG, ISA, TLX, DAM e HPR. O que pode justificar a exclusão dos três primeiros, é o fato destas atividades serem praticamente universalizadas na cidade de São Paulo, o que torna os índices oriundos das mesmas insuficiente para expressar desigualdades entre os setores. No que concerne aos indicadores ligados à consciência ecológica e infra-estrutura, os baixos valores geraram variáveis dotadas de pequenas porcentagens, tornando-os pouco relevantes já no princípio da análise.

Retornando às análises por meio do *stepwise regression*, utilizando parâmetros mais restritos⁵ com relação aos utilizados anteriormente, foi possível eliminar outros cinco indicadores dos vinte restantes: ISO, APR, CLP, 1AE e 4AE. Suprimiu-se ISO devido à alta correlação desta com 15AE (cerca de 0,93), que demonstrou ser mais importante para o modelo. De certa forma, o mesmo ocorre com as duas últimas variáveis indicadas, que compõe com de 8AE, de 11AE e 15AE todas as possibilidades relacionadas ao nível de instrução. No decorrer dos passos do *stepwise*, CLP foi colocado entre as selecionadas, porém quando CAL foi integrado ao modelo, aquela foi excluída, revelando que CLP perde muito de seu poder explicativo com a presença de CAL. Por outro lado, APR não chega a fazer parte do modelo em nenhum momento, o que indica que mesmo apresentando correlação com “Y” de 0,38, maior que ATA (0,26) e CMC (0,37), por exemplo, seu valor não é significativo quando em um modelo múltiplo. Com as variáveis remanescentes executou-se às relações percebidas como as mais pertinentes no esforço de criar um modelo mais fácil de compreender. Para reduzir o número de indicadores efetuaram-se análises das possíveis combinações entre as variáveis independentes. Este método permitiu eleger um modelo com seis variáveis (CMC, CS1, LCS2, 15AE, DHA e LLON), apresentando um bom ajuste e um $R^2=0,735$, enquanto o modelo com 15 variáveis (ATA, CAL, CMC, CPR, CS1, LCS2, CNA, 8AE, 11AE, 15AE, DHA, LJU1, JU2, LLON, MNA) exibiu um $R^2=0,744$.

Outrossim, foi necessário estabelecer se havia alguma interação relevante entre as variáveis presentes no modelo. Estabelecidas todas as interações possíveis para cada uma das seis variáveis, observou-se que, especificamente, interagindo CMC e 15AE conseguia-se uma melhora

⁴ Especificou-se limites para adição de 3,8 e para exclusão de 3,7, correspondendo a um nível de significância de 0,05, para um teste com a tolerância mínima aceitável de 0,01.

⁵ Os limites para adição foram de 10,8 e para exclusão 10,7 correspondendo a um nível de significância de 0,01, com a tolerância mínima aceitável de mesmo valor.

significativa no coeficiente de determinação. Contudo, permanecer com estas e sua interação acarretaria em um dificultador, pois existe uma alta correlação da interação com a variável 15AE, causando a necessidade de transformação das variáveis para remediar o problema de colinearidade. Antes disso, porém, tentou-se substituí-las por um indicador que fosse uma expressão da interação entre CMC e 15AE. Com sucesso gerou-se a variável denominada “mulheres chefes de família com mais de 15 anos de estudo” (MC15)⁶ – a qual acabou se tornando a variável com maior correlação (negativa) com os homicídios. Em vista disso, elaborou-se um modelo final simples e sem interações, com cinco variáveis ($R^2=0,739$):

$$\text{LHOM} = -2,597 -0,338 \text{ CS1} -0,398 \text{ LCS2} -2,203 \text{ MC15} +0,283 \text{ DHA} -0,081 \text{ LLON}.$$

4. Análise da adequabilidade do modelo

Para averiguar o ajuste do modelo, um conjunto de testes foi efetuado. As análises gráficas revelaram não haver indícios de alta assimetria nos resíduos, aproximando-se de uma distribuição normal. Por intermédio do teste Kolmogorov-Smirnov (Morcillo, 2004) confirmou-se a normalidade dos resíduos.

O gráfico que relaciona os valores estimados e observados foi utilizado para a identificação de pontos discrepantes (*outliers*) que uma vez reconhecidos foram examinados cuidadosamente. Decidiu-se que só seriam removidos caso fossem percebidos como fortes produtores de leituras não confiáveis dos dados. Não ocorrendo isso, estes pontos foram mantidos. Igualmente, observando os resíduos em relação às variáveis independentes utilizadas, o comportamento aparentou aleatoriedade e, em geral, não foi identificado um padrão indicativo de inconstância da variância do erro. Ainda assim, realizou-se o teste de Breusch-Pagan e Levene (Neter et al., 1996) e estes indicaram a não constância da variância para as variáveis indicadas⁷.

A partir da constatação de que a variância do modelo poderia não ser constante, optou-se por aplicar os mínimos quadrados ponderados para solucionar a questão. Segundo Neter et al. (1996), transformações em “Y” seriam úteis para se reduzir ou eliminar este problema, mas poderiam criar relações impróprias no modelo de regressão. Assim, se uma boa relação de regressão foi alcançada, mas existe uma condição de heterocedasticidade, uma alternativa é realizar este procedimento baseando-se na generalização do modelo de regressão múltipla. Para prosseguir com este procedimento foi necessário o uso de estimativas das variâncias por meio do método da função da estimação de variância que se baseia em conhecimentos empíricos de que a magnitude da variância geralmente varia de forma regular com uma ou diversas variáveis “X”. Assim, foi gerada, a partir da regressão dos resíduos absolutos pela variável LCS2 (esta variável mostrou-se mais correlacionada com os desvios absolutos), a função do desvio padrão estimado que permitiu encontrar o valor estimado do desvio padrão e então os pesos a serem aplicados.

O modelo alcançado com o uso dos pesos revelou-se semelhante ao obtido anteriormente, implicando em não ser necessário reestimar a função do desvio padrão e os pesos. Deste modo, os novos testes realizados evidenciaram uma melhora, mas não o suficiente para que se pudesse afirmar ser a variância constante. Vale a pena ressaltar as conseqüências da inconstância da variância. A pressuposição de homocedasticidade é fundamental para a construção dos intervalos de confiança e para os testes de hipóteses. Na presença de heterocedasticidade, os estimadores de mínimos quadrados dos parâmetros do modelo são lineares e não-tendenciosos, mas os estimadores das variâncias dos parâmetros são tendenciosos. Segundo Vasconcellos e Portela (2001), aceitar a existência heterocedasticidade significativa que, uma vez corrigida, não altera significativamente os resultados é uma situação no mínimo curiosa. De fato, pode ser uma indicação de que os testes

⁶ Assim calculado: Número de mulheres responsáveis pelos DPP com mais de 15 anos de estudo dividido pelo número total de pessoas responsáveis pelos DPP em determinado setor censitário, números coletados pelo IBGE.

⁷ Aqui, a abordagem espacial constitui-se uma importante maneira de tratamento dos dados, pois o modelo de correlação espacial, que descreve as tendências locais, permite a identificação de regiões por sua variabilidade.

aplicados foram excessivamente rigorosos, pois em amostras grandes, qualquer pequena variação tem alta probabilidade de ser considerada significativa pelos testes executados.

Diante disso, fez-se a validação do modelo utilizando os 25% dos dados reservados para esta atividade. Com esse grupo de elementos, conseguiu-se um coeficiente de determinação de 0,726, valor muito próximo ao encontrado anteriormente ($R^2=0,735$). Neste momento, um novo teste F apontou a relevância de todos os coeficientes de regressão. Outro indicativo da validade do modelo foi a pequena diferença entre o Quadrado Médio do Resíduo quando utilizada a amostra de 25% dos dados, 0,0438, e 75%, 0,0406. Ademais, executando passo a passo as mesmas análises de adequabilidade feitas previamente, nenhuma diferença expressiva foi observada.

Por conseguinte, ao permitir a identificação de uma relação estatística linear entre homicídios (variável dependente) e alguns indicadores intra-urbanos (variáveis independentes), o modelo de regressão proposto mostrou-se apropriado às pretensões deste trabalho, visto que, mesmo com a possível heterocedasticidade, não foram diagnosticadas características indesejáveis como alta correlação entre as variáveis predictoras, omissão de variáveis explicativas úteis, resíduos correlacionados, invalidação ou mau ajuste do modelo e não normalidade dos erros (**Figura 1**) e, outrossim, conseguiu-se um coeficiente de determinação significativo.

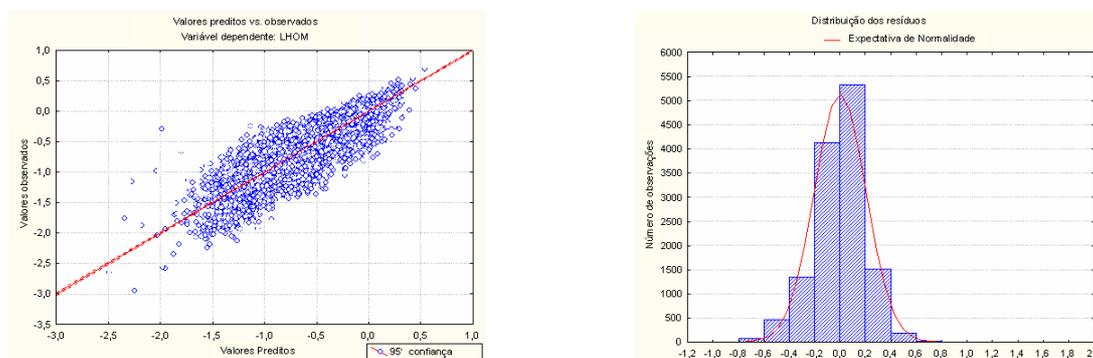


Figura 1 – Gráfico de dispersão entre observados e esperados e histograma dos resíduos.

5. Análise espacial do resíduo

A partir de então, os resíduos padronizados do modelo escolhido foram espacializados em um Sistema de Informações Geográficas (SIG)⁸ para que pudesse analisar a distribuição destes valores no município de São Paulo. Pela análise visual, nota-se que a espacialização dos resíduos parece não possuir uma dependência espacial. Afirmação sustentada pelo fato de não haver padrão perceptível na distribuição dos polígonos pertencentes a cada classe gerada (**Figura 2**).

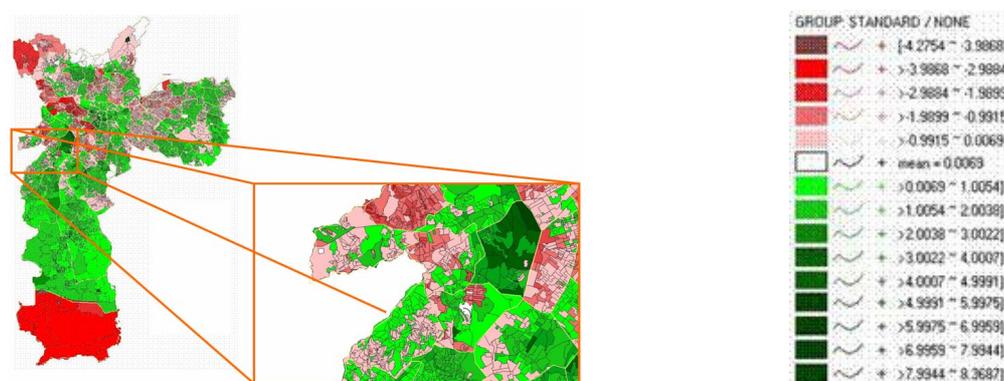


Figura 2 – Agrupamento do resíduo padronizado estatístico por faixas de desvio padrão.

⁸ Utilizou-se o Sistema de Processamento de Informações Georreferenciadas (SPRING), desenvolvido pelo INPE (Câmara et al., 1996).

Para que esta percepção visual pudesse ser avaliada, realizou-se o teste para se chegar ao índice global de Moran (Anselin, 1996). Teste que supõe uma hipótese nula de que não há dependência espacial quando seu valor é zero ou resulta em um índice próximo a este número. Assim, por um processamento no SPRING, chegou-se a um valor I de 0,0505 com 99 permutações, resultado que sugere não haver dependência espacial dos resíduos padronizados.

Além disso, uma comparação sobre os resultados obtidos pela espacialização dos dados de homicídios observados e estimados também é importante para se compreender o comportamento do modelo. Optou-se em verificar parte da região sul por sua grande relevância no tema em questão. Como se nota na **Figura 3**, a distribuição dos dados observados e preditos são bastante próximas, principalmente quando se observa que os intervalos foram divididos de forma distinta pelo SIG para cada conjunto de dados. Assim, foi possível destacar a similaridade dos resultados, indicando a boa adequabilidade do modelo de regressão proposto.

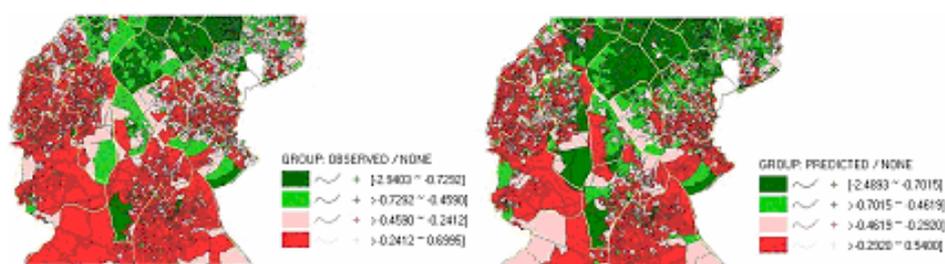


Figura 3 - Ilustração da distribuição espacial dos homicídios observados e estimados.

Desta maneira, o resultado apresentado pelo índice de Moran parece revelar não haver associação espacial para os homicídios, contradizendo a idéia proposta inicialmente de que características locais podem ser associadas a estes crimes. Porém, como na indicação da inconstância da variância, aqui não é possível corroborar tal condição, pois o referido índice mostrou-se pouco adequado à configuração dos dados abordados neste estudo, principalmente devido à forma de elaboração do indicador de homicídios e à heterogeneidade dos valores, relacionada ao elevado número de casos abrangidos 13.278 (total de setores censitários).

6. Conclusão

Neste trabalho, o uso de métodos estatísticos multivariados e ferramentas computacionais levou a um grupo de indicadores que se espera contribuir com a realização de análises multidimensionais, subsidiando a determinação das informações necessárias à consecução de políticas públicas ligadas ao combate à criminalidade.

O aumento da densidade populacional, por exemplo, pode denotar a elevação do risco da ocorrência de homicídios para uma dada localidade. Destarte, a importância do indicador elaborado, mulheres chefes de família com mais de 15 anos de estudo, ressalta o valor de estabelecer políticas públicas que facilitem o acesso de mulheres, muitas vezes mães e responsáveis pelo sustento familiar, ao ensino. Não obstante, críticas podem ser estabelecidas ao conjunto de termos escolhidos, porque considerar que locais com alta concentração de pessoas com mais de 70 anos serão dotados de menor criminalidade não expressa, certamente, uma verdade restrita, já que pode haver setores caracteristicamente jovens onde outros fatores podem, de certa forma, compensar uma baixa longevidade. Do mesmo modo, é interessante discutir a validade de indicadores como conforto sanitário, tendo por base a idéia de que a existência de domicílios dotados de elevado número de banheiros realmente retrata algum tipo de ganho social.

Portanto, à luz de tais considerações, metodologias que unem estatística e SIG reiteram a sua utilidade ao revelar características que possibilitam a melhor compreensão do problema da violência e a qualificação das ações preventivas de instituições de planejamento e intervenção.

7. Agradecimentos

Aqui registra-se os mais sinceros agradecimentos ao Programa de Aprimoramento em Informações sobre Mortalidade do Município de São Paulo por fornecer o número de homicídios e à Coordenadoria de Análise e Pesquisa da Secretaria de Segurança Pública de São Paulo, na figura do Dr. Túlio Kahn, por ter gentilmente cedido a base digital da malha de setores censitários de 2000, assim como de seu banco de dados, tornando possível a realização deste artigo.

Referências

Anselin, L. The Moran scatterplot as ESDA tool to assess local instability in spatial association. In: Fisher, M.; Scholten, H. J.; Unwin, D. **Spatial Analytical Perspectives on GIS**. London: Taylor & Francis, 1996. p. 111-126.

Beato, C.; Assunção, R.; Santos M. C. **Análise da evolução temporal da criminalidade violenta em Minas Gerais (1986-1997)**. Mimeo, São Paulo, 1997.

Câmara, G.; Yi, J. R. L.; Felgueiras, C.; Moneiro, A. A. **Análise Espacial da Distribuição e dinâmica da violência na cidade de São Paulo entre os anos de 1996 e 1999**. São José dos Campos, INPE, 2000.

Câmara, G.; Freitas, U. M.; Souza, R. C. M.; Garrido, J. **Spring: Integrating remote sensing and GIS by object-oriented data modeling**. Computers & Graphics, v. 15, n. 6, p. 395-403, 1996.

Camargo, E. C. G.; Ortiz, J. O.; Moreira, R. C. **Relação da Taxa de Homicídios no Município de São Paulo com Variáveis Socioeconômicas**. Trabalho final de curso métodos estatísticos multivariados e algumas aplicações ao sensoriamento remoto, Inpe, São José dos Campos, 2003.

Carneiro, L. P. **Determinantes do Crime Na América Latina: Rio de Janeiro e São Paulo**. Relatório de Pesquisa, Departamento de Ciência Política, Universidade de São Paulo, 1999.

Fernandes, R. C.e Carneiro, L. P. Criminalidade, drogas e perdas econômicas no Rio de Janeiro. **Lei e liberdade**, Rio de Janeiro, ano 15, n. 47, 1996.

Morcillo, A. M. **Teste de Kolmogorov-Smirnov**. Disponível em: <<http://www.fcm.unicamp.br/centros/ciped/mp639/Teste%20de%20Kolmogorov.pdf>>. Acesso em: 13 de set. de 2004.

Neter, J.; Kutner, M. H.; Nachtsheim, C. J.; Wasserman, W. **Applied Linear Statistical Models**. Chicago: Irwin, 1996. 1408 p.

Silva, B. F. A. **Criminalidade urbana violenta: uma análise espaço-temporal dos homicídios em Belo Horizonte**. Fafich, UFMG, 2001.

Souza, E. R.; Minayo M. C. **O impacto da violência social na Saúde Pública do Brasil: década de 80**. In: Minayo, M. C. (Org.). Os muitos brasis: saúde e população na década de 80. São Paulo: HUCITEC, 1995.

Sposati, A. **Mapa da Exclusão/Inclusão da Cidade de São Paulo**. Editora da PUC-SP, São Paulo, 1996.

StatSoft, Inc. **Electronic Statistics Textbook**. Disponível em: <<http://www.statsoft.com/textbook/stathome.html>>. Acesso em: 13 set. 2004.

Vargas, J. H. **À espera do passado: as transformações recentes de São Paulo vistas de seu epicentro**. Campinas, Unicamp, dissertação de mestrado em Antropologia Social, 1993.

Vasconcellos, M. T. L.; Portela, M. C. **Índice de Massa Corporal e sua relação com variáveis nutricionais e sócio-econômicas: um exemplo de uso de regressão linear para um grupo de adultos brasileiros**. Disponível em: <http://www.scielo.br/scielo.php?script=sci_home&lng=pt&nrm=iso>. Acesso em: 13 set. 2004.

Zaluar, A. Relativismo cultural na cidade? In **Anuário antropológico 90**. Tempo Brasileiro, Rio de Janeiro, 1993.