

*Empresa Brasileira de Pesquisa Agropecuária
Embrapa Cerrados
Ministério da Agricultura, Pecuária e Abastecimento*

ANÁLISE ESPACIAL DE DADOS GEOGRÁFICOS

Editores Técnicos

Suzana Druck
Marília Sá Carvalho
Gilberto Câmara
Antônio Miguel Vieira Monteiro

Planaltina, DF
2004

APRESENTAÇÃO

Um dos maiores desafios para todos aqueles que estudam os fenômenos sociais, econômicos e ambientais no Brasil é a necessidade de combinar a compreensão histórica e qualitativa desses processos com o tratamento quantitativo de dados, para subsidiar a tomada de decisões em bases mais objetivas. Essa tendência vem sendo fortalecida com a redução de custo da tecnologia de informação que permite o acesso a técnicas de tratamento analítico de dados sofisticadas e o uso de sistemas de informação geográfica (SIG), como forma de organizar os dados em áreas como Saúde Coletiva, Demografia, Políticas Públicas, Planejamento Urbano, Serviço Social, Criminalidade, Geologia, Ciência dos Solos e Zoneamento Ambiental.

No entanto, a publicação de material didático de qualidade e, em português, não tem acompanhado a velocidade da disponibilização dos métodos e tecnologias que lidam com esse tipo de dado. Para completar o processo de democratização no tratamento da informação espacial, é necessário ampliar a oferta de textos que apresentem questões de análise de dados georreferenciados, mostrando que se pode ir muito além da geração dos velhos mapas coloridos.

Com essa motivação, a Embrapa Cerrados apresenta este livro. Procurou-se, com essa publicação, preencher o vazio editorial em língua portuguesa, na área de tratamento da informação espacial. Este trabalho foi desenhado como um livro-texto introdutório, para embasamento a alunos de graduação e pós-graduação nas diferentes disciplinas que fazem uso direto de dados de natureza geográfica. Ao longo de seus cinco capítulos, o leitor poderá dispor de uma série de métodos e técnicas aplicados às mais diversas áreas e, em particular, vai observar o uso de vários exemplos com dados de diferente natureza e de diferentes regiões do País. O estilo simples, direto e ilustrativo transforma esta obra em uma ferramenta

importante para a ampliação da rede de especialistas que possam procurar usar o “lugar” dos dados, o espaço geográfico no qual eles existem, como uma linguagem comum, cada vez mais necessária à compreensão e a apreensão da complexidade dos problemas com os quais precisamos lidar.

Roberto Teixeira Alves
Chefe-Geral da Embrapa Cerrados

PREFÁCIO

Neste livro, apresenta-se um conjunto de técnicas estatísticas que faz uso explícito da localização dos fenômenos. Tais técnicas partem do princípio que “há algo de especial com dados espaciais”, e que “o local faz a diferença”. Originalmente desenvolvidas em ambientes especializados, essas técnicas estão se tornando disponíveis a um público cada vez mais amplo, seja pela inclusão de funções de estatística espacial nos SIGs, seja pelo desenvolvimento de interfaces eficientes para combinar a capacidade de visualização dos SIGs com o poder analítico dos pacotes de estatística e também pela disseminação dos software de domínio público, democratizando o acesso à ferramenta.

Imaginamos nosso leitor como um especialista em disciplinas como Ciência dos Solos, Geologia, Oceanografia, Saúde Coletiva ou Políticas Públicas que possui noções básicas de estatística multivariada e deseja utilizar as técnicas de análise espacial em seus dados georreferenciados. Buscamos produzir um livro-texto que pudesse ser acessível a esse leitor, limitando as questões teóricas de estatística ao mínimo indispensável. Para tanto, adotamos uma abordagem *prática*, focando nossa discussão em técnicas de eficácia comprovada e disponíveis de forma ampla. O caráter introdutório e (esperamos) didático deste livro também decorre do fato de ser utilizado como suporte a cursos de pós-graduação, ministrados pelos autores no Inpe, na USP e na Fiocruz. Esse material, também, é utilizado em tutorias e cursos de curta duração que vêm sendo ministrados pelos autores em diversas instituições no Brasil. A seleção dos métodos apresentados no livro foi relativamente arbitrária, baseada na experiência dos autores, no ensino, e em trabalhos de pesquisa, parte das quais utilizadas nos exemplos ao longo do livro. Sendo um livro introdutório, optamos por omitir algumas técnicas importantes, mas que iam requerer conhecimento diferenciado, como é o caso dos modelos que utilizam inferência bayesiana.

Como material complementar, tornamos disponível os dados utilizados nos exemplos do livro, com sugestões de exercícios e procedimentos operacionais para que o leitor possa reproduzi-los

e ampliá-los, no sítio <www.dpi.inpe.br/gilberto/livro/analise>. A maior parte das técnicas pode ser executada pelo software Spring ou a partir da ligação entre o Spring e softwares especializados, como o SpaceStat. O Spring é um desenvolvimento do Inpe e está acessível livremente em <www.dpi.inpe.br/spring>. Além disso, o pacote estatístico R <www.r-project.org>, de domínio público, também permite aplicar a maior parte das técnicas apresentadas.

O que esperamos do leitor? Além de uma leitura participativa, recomendamos o empenho em reproduzir as técnicas apresentadas, tanto nos dados de exemplo, quanto em seus próprios problemas. Gostaríamos ainda de imaginar que, pelo menos, parte de nossos leitores venha a se interessar por questões mais avançadas de estatística espacial para as quais apontamos a bibliografia relevante.

Como leitura complementar a este livro, sugerimos ao leitor duas especiais: o livro *Interactive Spatial Data Analysis*, de Trevor Bailey e Anthony Gatrell, obra de grande valor didático que serviu de inspiração para os autores, e a monografia *Estatística Espacial com Aplicações em Epidemiologia, Economia e Sociologia*, de Renato Assunção (UFMG), disponível em sua página www.est.ufmg.br/-assuncao.

Passada mais de uma década da implantação dos primeiros sistemas de Geoprocessamento no Brasil, a comunidade precisa dar um “salto de qualidade” em sua capacidade de bem-utilizar os dados espaciais. Nossa intenção foi contribuir para esse necessário aprimoramento. Considerando a grande diversidade dos profissionais de Geoprocessamento, os autores têm consciência de que a leitura desse texto irá exigir esforço considerável por parte de grande parcela de seu público-alvo. Acreditamos que o conhecimento a ser adquirido será a plena recompensa do esforço.

Os Editores

SUMÁRIO

Capítulo 1

ANÁLISE ESPACIAL E GEOPROCESSAMENTO

1.1	INTRODUÇÃO	21
1.2	TIPOS DE DADOS EM ANÁLISE ESPACIAL	23
1.3	REPRESENTAÇÃO COMPUTACIONAL DE DADOS GEOGRÁFICOS	27
1.4	CONCEITOS BÁSICOS EM ANÁLISE ESPACIAL	33
	Dependência Espacial	33
	Autocorrelação Espacial	33
	Inferência Estatística para Dados Espaciais	35
	Estacionariedade e Isotropia	35
1.5	PROCESSO DA ANÁLISE ESPACIAL	36
	Modelos Inferenciais	39
	Processo Pontual	39
	Variação Contínua	40
	Variação Discreta	41
1.6	CONCLUSÕES	42
1.7	RESENHA	44
1.8	REFERÊNCIAS	45
	ANEXOS - SOFTWARES PARA ANÁLISE ESPACIAL	48
	Anexo 1. GSLIB – Biblioteca para Geoestatística	49
	Anexo 2. GSTAT – Software para Geoestatística	49
	Anexo 3. CLUSTERSEER – <i>Clustering</i> de Processos Pontuais	50
	Anexo 4. CRIMESTAT – Análise de Estatísticas Criminais	50
	Anexo 5. SPACESTAT – Análise Espacial de Áreas	51
	Anexo 6. SPRING	51
	Anexo 7. ARCGIS Geostatistical Analyst	52

Capítulo 2

ANÁLISE ESPACIAL DE EVENTOS

2.1	INTRODUÇÃO	55
2.2	CARACTERIZAÇÃO DE DISTRIBUIÇÕES DE PONTOS	57
2.3	ESTIMADOR DE INTENSIDADE (“ <i>KERNEL ESTIMATION</i> ”)	60
2.4	ESTIMADORES DE DEPENDÊNCIA ESPACIAL	63
	Método do Vizinho mais Próximo	63
	Função K	66
2.5	PONTOS NO ESPAÇO-TEMPO	68
2.6	PROCESSO PONTUAL MARCADO	69
2.7	ESTUDOS CASO-CONTROLE	71
2.8	COMENTÁRIOS FINAIS	75
2.9	REFERÊNCIAS	75

Capítulo 3

ANÁLISE ESPACIAL DE SUPERFÍCIES

3.1	INTRODUÇÃO	79
3.2	MODELOS DETERMINÍSTICOS LOCAIS	83
3.3	SUPERFÍCIES DE TENDÊNCIA	87
3.4	MODELOS ESTATÍSTICOS DE EFEITOS GLOBAL E LOCAL: KRIGEAGEM	90
	3.4.1 Fundamentação Teórica	90
	3.4.2 Determinação Experimental do Semivariograma	93
	3.4.3 Modelos Teóricos	96

3.5 KRIGEAGEM	103
3.5.1 Krigeagem Ordinária	103
3.6 ESTUDO DE CASO	105
3.7 CONCLUSÕES	108
3.8 REFERÊNCIAS	108
APÊNDICE: MODELAGEM DA ANISOTROPIA	110
Tipos de Anisotropia	110
Detecção da Anisotropia	111
Modelagem da Anisotropia	113
Modelagem da anisotropia geométrica ou zonal	113
Modelagem da anisotropia combinada	117

Capítulo 4

ANÁLISE ESPACIAL DE SUPERFÍCIES POR GEOESTATÍSTICA: O ENFOQUE POR INDICAÇÃO

4.1 INTRODUÇÃO	125
4.2 INCERTEZAS LOCAIS	126
4.3 O ENFOQUE POR INDICAÇÃO	129
4.3.1 A Codificação por Indicação dos Dados Amostrais	129
4.3.2 A Variografia por Indicação	130
4.3.3 A Estimação dos Valores por Indicação	131
Krigeagem por indicação simples	131
Krigeagem por indicação ordinária	132
4.3.4 Correção dos Desvios de Ordem	133
4.4 ESTIMATIVA DE INCERTEZAS LOCAIS	137
Intervalos de Probabilidade	137
Distância Interquantil	139
4.5 MEDIDAS DO ESPALHAMENTO DE UMA DISTRIBUIÇÃO	140
Variância Condicional	140
Entropia de Shannon	142

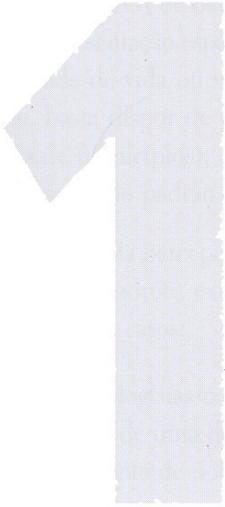
4.6	ESTIMADORES ÓTIMOS PARA AS SUPERFÍCIES INTERPOLADAS	142
	Estimativa do Valor Esperado	143
	Estimativa da Mediana	144
	Estimativa de Quantis	145
4.7	INCERTEZAS LOCAIS PARA ATRIBUTOS CATEGÓRICOS	145
	O Enfoque por Indicação para Atributos Categóricos	146
4.8	CLASSIFICADORES PARA ATRIBUTOS CATEGÓRICOS	148
4.9	MEDIDAS DE INCERTEZA PARA ATRIBUTOS CATEGÓRICOS	148
	A Incerteza do Classificador de Moda	149
	Incerteza por Entropia de Shannon	150
4.10	CONCLUSÕES	152
4.11	REFERÊNCIAS	154

Capítulo 5

ANÁLISE ESPACIAL DE ÁREAS

5.1	INTRODUÇÃO	157
5.2	MODELOS DE DISTRIBUIÇÃO DE DADOS EM ÁREAS	159
5.3	PROBLEMAS DE ESCALA E A RELAÇÃO ÁREA-INDIVÍDUO	160
5.4	ANÁLISE EXPLORATÓRIA	164
	Visualização de Dados	164
	Gráficos de Médias e Medianas	167
	Análise de Autocorrelação Espacial	169
	Matrizes de Proximidade Espacial	169
	Média móvel Espacial	171

Indicadores Globais de Autocorrelação Espacial: Índices de Moran e Geary	172
Diagrama de Espalhamento de Moran	179
Indicadores Locais de Associação Espacial	182
5.5 ESTIMAÇÃO DE INDICADORES	183
5.6 MODELOS DE REGRESSÃO	189
Modelos com Efeitos Espaciais Globais	191
Modelos de Regressão com Efeitos Espaciais Locais	193
(a) Caso discreto - modelos de regressão com regimes espaciais	193
Modelos de Regressão com Efeitos Espaciais Locais	194
(b) Modelos de regressão com efeitos espaciais contínuos	194
Diagnóstico de Modelos com Efeitos Espaciais	196
5.7 ESTIMAÇÃO DE MODELOS CONTÍNUOS A PARTIR DE DADOS DE ÁREA	200
Estimador de Intensidade não Paramétrico	201
Uso de Interpoladores Geoestatísticos	203
5.8 COMENTÁRIOS FINAIS	205
5.9 REFERÊNCIAS	206



Análise Espacial e Geoprocessamento

Gilberto Câmara
Antônio Miguel Monteiro
Suzana Druck
Marília Sá Carvalho

1.1 INTRODUÇÃO

Compreender a distribuição espacial de fenômenos constitui hoje um desafio para a elucidação de questões importantes em diversas áreas do conhecimento, como saúde, ambiente, geologia, agronomia, entre outras. Tais estudos vêm-se tornando cada vez mais comuns, devido à disponibilidade de sistemas de informação geográfica (SIG) de baixo custo e com interfaces amigáveis. Esses sistemas permitem a apresentação espacial de variáveis como população de indivíduos, índices de qualidade de vida ou vendas de empresa numa região por meio de mapas. Para tanto, basta dispor de um banco de dados e de uma base geográfica (como um mapa de municípios), e o SIG torna possível apresentar um mapa que permite a visualização do padrão espacial determinado pelo fenômeno em observação.

Além da percepção visual da distribuição espacial do problema, é muito útil traduzir os padrões existentes com considerações objetivas e mensuráveis, como nos seguintes casos:

- Epidemiologistas coletam dados sobre ocorrência de doenças. A distribuição dos casos de uma doença forma um padrão no espaço? Existe associação com alguma fonte de poluição? Há evidência de contágio? Variou no tempo?
- Deseja-se investigar se existe alguma concentração espacial na distribuição de roubos. Roubos que ocorrem em determinadas áreas estão correlacionados com características socioeconômicas dessas áreas?
- Geólogos desejam estimar a extensão de um depósito mineral em uma região a partir de amostras. Pode-se usar essas amostras para estimar a distribuição do mineral na região?
- Deseja-se analisar uma região para fins de zoneamento agrícola. Como escolher as variáveis explicativas – solo, vegetação, geomorfologia – e determinar qual a contribuição de cada uma delas para definir, para cada local, o tipo de cultura mais adequado.