

IPEAGEO

Alexandre Xavier Ywata de Carvalho*

Pedro Henrique Melo Albuquerque**

Waldery Rodrigues Júnior*

Rafael Dantas Guimarães**

Camilo Rey Laureto**

Marina Garcia Pena**

Gilberto Rezende de Almeida Júnior**

Fernanda Rocha Gomes da Silva**

1 INTRODUÇÃO

O IpeaGEO é um *software* desenvolvido pela assessoria de métodos quantitativos do Ipea e funciona como ferramenta de análises estatísticas, com foco em estudos espaciais. Foi criado com o objetivo de ser uma alternativa gratuita aos *softwares* já existentes no mercado. A inclusão de técnicas espaciais consolidadas e o foco no território nacional são alguns dos diferenciais que o programa apresenta.

O desenvolvimento do IpeaGEO partiu da ideia e da necessidade de um *software* que atendesse às demandas estatísticas e as geográficas e que não gerasse custos adicionais. A inclusão de técnicas presentes em recentes publicações e que ainda não foram implementadas em *softwares* concorrentes também foi ponto fundamental para idealização do programa.

A visualização, simulação e modelagem de processos espaciais no IpeaGEO contribuem para o melhor entendimento do contexto geográfico dos dados, permitindo observar relacionamentos e identificar padrões para assim dar suporte à tomada de decisão.

As técnicas foram implementadas de forma a facilitar a utilização do usuário final, permitindo que o trabalho seja feito intuitivamente, com todos os comandos ao alcance do mouse. Dessa forma, o IpeaGEO representa a oportunidade de desenvolver estudos econômicos e espaciais fundamentados em técnicas estatísticas avançadas e de fácil utilização.

* Pesquisador da Diretoria de Estudos e Políticas Regionais, Urbanas e Ambientais (Dirur) do Ipea.

** Bolsista da Dirur/Ipea.

2 ATRIBUTOS

A seguir, serão citadas algumas das ferramentas que podem ser encontradas no *software*.

2.1 Parte visual: ferramentas de visualização e seleção de mapa

Assim como nos outros programas, no IpeaGEO precisamos ler a malha digital e sequencialmente fazer o relacionamento com a base de interesse. O IpeaGEO disponibiliza algumas malhas e bases padrões, como a de municípios do Brasil e uma base de dados que incorpora mais de mil variáveis. Entre elas, as de estimativa populacional, área da unidade territorial (km²), informações estatísticas, população e domicílios, censo agropecuário, produto interno bruto (PIB) dos municípios, informações de saúde, ensino, estatísticas do registro civil, representação política, estatísticas do cadastro central de empresas, instituições financeiras, finanças públicas, frota, pobreza e desigualdade.

A partir desse ponto, o IpeaGEO permite selecionar polígonos específicos da malha e atualizar a base para exibir apenas estes. Também pode-se selecionar linhas específicas da base e atualizar a malha de acordo. As ferramentas de *zoom* e de informação permitem analisar características específicas de cada polígono, em especial quando o número deles é muito alto, como no caso dos municípios brasileiros.

2.2 Mapa temático

Um mapa temático é simples, feito para refletir um tema particular em determinada área geográfica. Eles esboçam a distribuição espacial de eventos físicos, políticos, culturais, econômicos, sociológicos ou qualquer outro aspecto da região (cidade, município e setor censitário).

O mapa temático é dividido em classes, em que cada uma é representada por uma cor. Assim sendo, podem-se observar padrões espaciais por meio delas. Duas questões surgem imediatamente antes da construção do mapa: o número de classes que possuirá e o tamanho das respectivas.

2.2.1 O número ótimo de classes

É fácil perceber que se forem criadas muitas classes, pouco consegue se observar no mapa. Este ficaria diluído e os padrões espaciais não seriam observados. Analogamente, se o mapa for criado com poucas classes, quase nenhuma informação é obtida, já que ele se tornaria denso e estes padrões se perderiam entre as classes.

Existem muitas respostas à questão do número ótimo de classes, uma vez que a definição de boas classes ou melhor número de classes muda dependendo do analista.

2.2.2 Método de criação das classes

Escolhido o número de classes, o próximo passo é alocar os dados em cada uma delas. Dependendo do método, a amplitude das classes pode ser diferente.

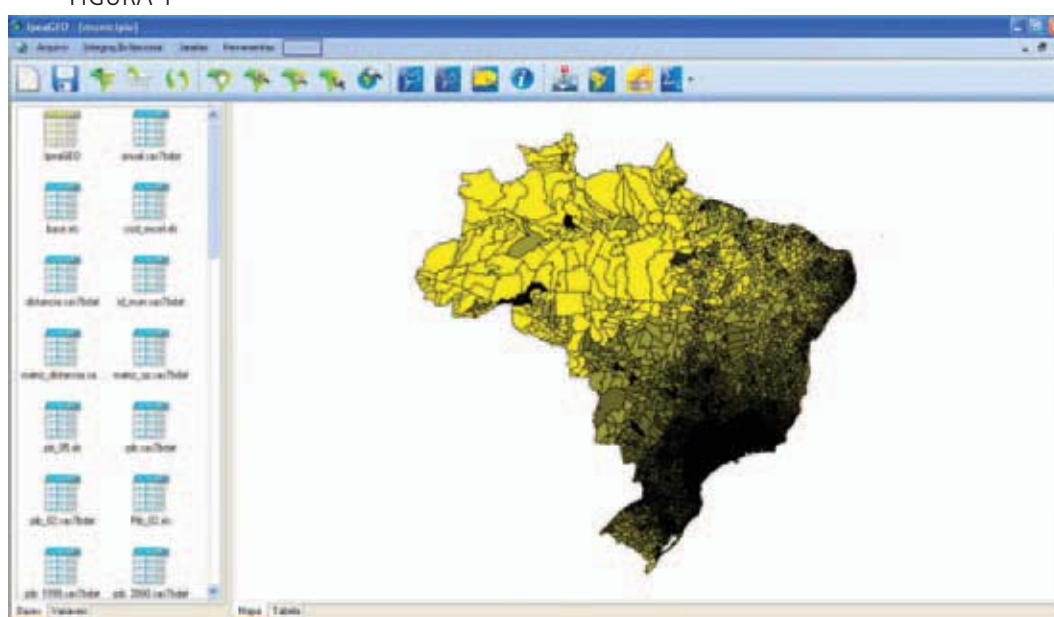
O IpeaGEO possui os seguintes métodos de criação de classes:

1. Quantil: intervalos são selecionados de tal maneira que o número de observações em cada intervalo é igual.
2. Equal: os valores são divididos em k classes com intervalos iguais.
3. Jenks: também conhecido como *natural breaks*, é o método mais utilizado nos pacotes de análise espacial. Sua composição tem como objetivo criar classes com menor variância interna e maior variabilidade externa.

4. Geométrico: este é um sistema de classificação em que as quebras de classe se baseiam em intervalos de classe que têm uma série geométrica.
5. Desvio-padrão: a média e o desvio-padrão dos valores são calculados e as classes são criadas com base na distância do atributo em relação à distância da média da normal padronizada com seu desvio-padrão.
6. Valores únicos: cada valor é tratado separadamente, por exemplo, um mapa com cores distintas para cada polígono.
7. Classificação manual: o usuário especifica os limites entre as classes. Após a criação do mapa temático, o IpeaGEO gera um relatório que permite que quem o utilize exporte o resultado obtido.

Na figura 1, podemos visualizar um mapa gerado pelo programa para a malha de Brasil e municípios. Foi feito um deste para a variável mercado de produção, utilizando-se o método de Jenks com cinco classes.

FIGURA 1



Fonte: IpeaGEO.

Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

2.3 Modelos de regressão

O IpeaGEO dispõe dos modelos tradicionais de regressão não espacial. Em uma interface *point-and-click* de fácil utilização, o usuário pode escolher a variável dependente, as variáveis independentes e o modelo de regressão. Os métodos disponíveis são: Mínimos Quadrados Ordinários, Mínimos Quadrados em dois Estágios e Método de Momentos Generalizados.

Após a regressão, o usuário tem acesso ao vetor de resíduos, além de uma tela com resultados dos coeficientes e dos testes estatísticos.

2.4 Modelos de regressão espacial

Um dos grandes diferenciais do IpeaGEO é a possibilidade de rodar modelos de regressão que incorporam a dimensão espacial. O usuário disporá, automaticamente, da possibilidade de acrescentar, como variáveis explicativas do modelo, matrizes com variáveis ponderadas pela vizinhança. Dessa maneira, o modelo poderá incorporar a influência da proximidade espacial como variável explicativa.

Os modelos de regressão espacial disponíveis são o Spatial Auto-Regressive Model (SAR), Spatial Error Model (SEM) e o Método de Momentos Generalizados Espacial.

Após a regressão, é possível visualizar os resíduos, a matriz de variâncias e as covariâncias e testes estatísticos relevantes.

2.5 Calculadora

Função de considerável utilidade nos estudos econômicos, a calculadora do IpeaGEO representa a possibilidade de criar variáveis durante o processo. O usuário dispõe de uma série de opções de operações com variáveis para a criação de variáveis derivadas; por exemplo, a opção da transformação de Box-Cox para a normalização de uma variável. A criação de variáveis também conta com a opção de filtros e condições, de forma que as variáveis derivadas podem ser criadas de forma mais precisa.

2.6 Ferramentas de georeferenciamento

O IpeaGEO permite que o usuário exporte arquivos com informações espaciais descritivas da sua malha em utilização. Podem ser gerados arquivos com coordenadas dos polígonos, de centróides, *bounding box* e matrizes de distâncias e de vizinhança (de ordem 1 a *k*, normalizadas ou não, construídas considerando-se as definições de vizinhança *Queen* ou *Rook*).

2.7 Indicadores de dependência espacial

O IpeaGEO conta com uma série de indicadores que auxiliam na detecção de dependências espaciais relativas às variáveis nos polígonos. Os índices de dependência global (Moran, Geary, Getis Ord, Tango e Rogerson) e local (Lisa, Getis Ord e Escore) são fornecidos em conjunto com seus respectivos níveis de significância, calculados por meio de simulações de Monte Carlo.

2.8 Indicadores de segregação

O usuário também dispõe da opção de gerar indicadores de segregação. Estes índices são utilizados como indicadores da relação existente entre as populações das diversas áreas. Eles têm também como utilidade a possibilidade de mensurar a interação intrazonal entre grupos raciais.

Esses índices medem a proporção da população de um grupo que deveria se mudar para outras áreas para que a composição populacional em cada uma das unidades de área seja idêntica à da cidade como um todo.

Como opções de cálculo, o IpeaGEO oferece indicadores tipo *Evenness*, *Exposition* e *Clustering* para apenas um grupo, entre dois grupos e multigrupos.

A seguir, são listados os indicadores que podem ser calculados no *software*. Indicadores de segregação para apenas um grupo: *Segregation Index*, *Segregation Index Adjusted for Contiguity*, indicador de entropia, coeficiente de Gini, Índice de Atkinson, *Isolation Index*, *Correlation Ratio*, *Absolute Clustering*, *Mean Proximity between Members of Group X*, *Mean Proximity between Members of One Group (exp)*, e *Distance Decay Isolation Index*.

Indicadores de segregação entre dois grupos: *Index of Dissimilarity between XY*, *Dissimilarity Adjusted for Contiguity between XY*, *Interaction Index*, *Mean Proximity between Members of groups X and Y*, *Mean Proximity between Members of One Group X and Group Y*, *Spatial Proximity Index*, *Relative Clustering Index* e *Distance Decay Interaction Index*.

Indicadores de segregação multigrupos: *Dissimilarity*, coeficiente de Gini, *Information Theory*, *Normalized Exposure*, *Relative Diversity* e *Squared Coefficient of Variation*.

Com esse leque de ferramentas, o usuário que utilizar o IpeaGEO poderá fazer uma rica análise estatística e geográfica por meio de um *software* gratuito e de fácil manuseio. Além disso, o programa disponibiliza ferramentas até então inéditas, o que caracteriza outro grande diferencial em relação aos outros *softwares* presentes no mercado.

REFERÊNCIAS

ANSELIN, L. **Spatial Econometrics: methods and models**. Dordrecht: Kluwer Academic Publishers, 1988.

_____. Local indicators of spatial association (Lisa). **Geographical Analysis**, v. 27, p. 93-115, 1995.

APPARICIO, P. Les indices de ségrégation résidentielle: un outil intégré dans un système d'information géographique. **Cybergeog: european journal of geography**, n. 134, 17 p. 2000.

APPARICIO, P.; LELOUP, X.; RIVET, P. La diversité montréalaise à l'épreuve de la ségrégation: pluralisme et insertion résidentielle des immigrants. *Revue de l'intégration et de la migration internationale*. **Journal of International Migration and Integration**, v. 8, p. 63-87, 2007.

APPARICIO, P.; SÉGUIN, A. M. La division de l'espace résidentiel montréalais en fonction de la langue maternelle: apport des indices de ségrégation résidentielle. **Canadian Journal of Urban Research: Revue canadienne de recherche urbaine**, v. 11, p. 265-297, 2002.

BAILEY T.; GATRELL, A. **Interactive Spatial Data Analysis**. Harlow: Longman, 1995.

CLIFF, A. D.; ORD, J. K. **Spatial Processes: models and applications**. London: Pion, 1981.

CRESSIE, N. A. C. **Spatial Statistics**. New York: John Wiley, 1993.

DAVIS, C. C. G.; MONTEIRO, A. M. **Introdução à ciência da geoinformação**. São José dos Campos: DPI/Inpe, 2001.

DOANE, D. P. Aesthetic frequency classifications. **American Statistical Association**, v. 30, n. 4, p. 181-183, 1976.

DRUCK, S. C. G.; MONTEIRO, A. M. **Análise espacial de dados geográficos**. Planaltina: Embrapa, 2004.

DUNCAN, O. D.; CUZZORT, R. P.; DUNCAN, B. **Statistical Geography: Problems in Analyzing Areal Data**. Illinois: The Free Press of Glencoe, 1961.

DUNCAN, O. D.; DUNCAN, B. A methodological analysis of segregation indexes. **American Sociological Review**, n. 20, p. 210-217, 1955a.

_____. Residential distribution and occupational stratification. **American Journal of Sociology**, n. 60, p. 493-503, 1955b.

GEARY, R. C. The contiguity ratio and statistical mapping. **Incorporated Statistician**, n. 5, p. 115-141, 1954.

GOODMAN, L.; KRUSKAL, W. H. Measures of Association for Cross-Classifications. **Journal of the American Statistical Association**, n. 49, p. 732-764, 1954.

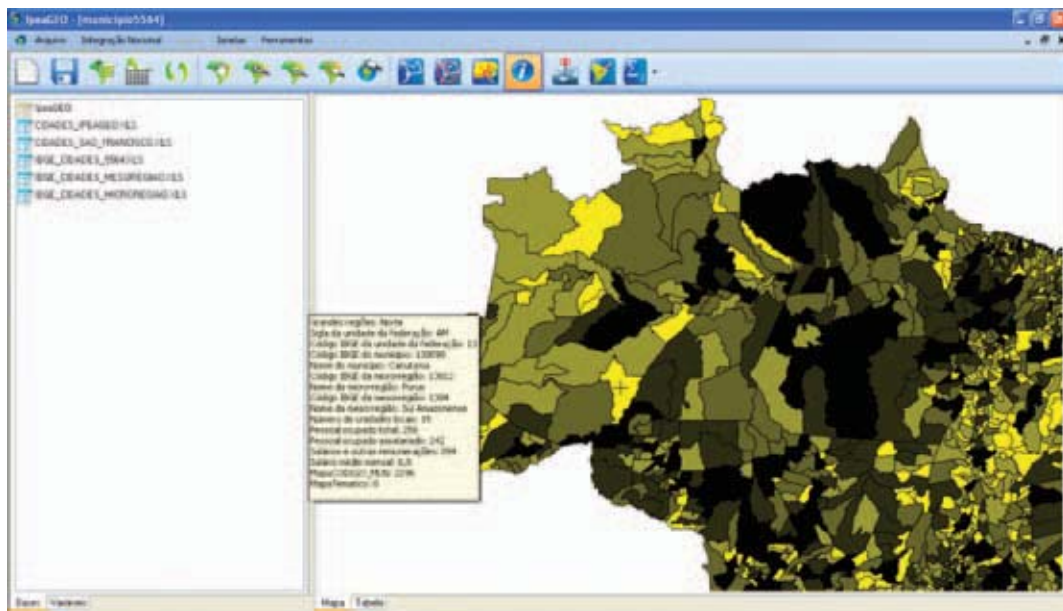
- HIRAI, Y. Some Remarks on Class Interval of Histograms. 1989. Disponível em: < ousar.lib.okayama-u.ac.jp/file/9696/.../082_0113_0117.pdf>. Acesso em: out. 2010.
- HOOVER, E. M. Interstate redistribution of population: 1850-1940. **Journal of Economic History**, n. 1, p. 199-205, 1941.
- LASZLO, M. **Computational Geometry and Computer Graphics in C++**. Upper Saddle River: Prentice Hall, 1996.
- LIBERSON, S. Measuring population diversity. **American Sociological Review**, n. 34, p. 850-862, 1969.
- MASSEY, D. S.; DENTON N, A. The dimensions of residential segregation. **Social Forces**, n. 67, p. 281-315, 1988.
- MASSEY, D. S.; DENTON, N. A.; PHUA V. C. The dimensions of residential segregation revisited. **Sociological Methods & Research**, n. 25, p. 172-206, 1996.
- MATEOS, P.; WEBBER, R.; LONGLEY, P. How segregated are name origins? A new method of measuring ethnic residential segregation. *In: THE GIS RESEARCH UK (GISRUK) ANNUAL CONFERENCE*, p. 285-291, 2006. **Anais**. 2006.
- MORAN, P. A. P. Notes on Continuous Stochastic Phenomena. **Biometrika**, v. 37, p. 17-33, 1950.
- MORI, T. An optimal length of class interval for histogram. **Japanese Journal of Applied Statistics**, v. 4, n. 1, p. 17-24, 1974.
- RIPLEY, B. D. **Spatial Statistics**. New York: John Wiley, 1981.
- ROGERSON, P. The detection of clusters using a spatial version of the chi-squared goodness-of-fit test. **Geographical Analysis**, v. 31, p. 130-147, 1999.
- ROGERSON, P.; YAMADA, I. **Statistical Detection and Surveillance of Geographic Clusters**. Boca Raton, Florida: Chapman & Hall/CRC Press, 2009.
- STURGES, H. A. The choice of a class interval. **Journal of the American Statistical Association**, v. 21, p. 65-66, 1926.
- RAY, B. K. Plural geographies in canadian cities: interpreting immigrant residential spaces in Toronto and Montreal. **Canadian Journal of Regional Science**, n. 22, p. 65-86, 1999.
- REARDON, S. F. Measures of racial diversity and segregation in multigroup and hierarchical structured populations. *In: ANNUAL MEETING OF THE EASTERN SOCIOLOGICAL SOCIETY*, Philadelphia, PA, 1998.
- REARDON, S. F.; FIREBAUGH, G. Measures of multigroup segregation. **Sociological Methodology**, v. 32, p. 33-67, 2002.
- ROMERO, M. A. B. *et al.* Indicadores de sustentabilidade dos espaços públicos urbanos: aspectos metodológicos e atributos das estruturas urbanas. *In: SEMINÁRIO A QUESTÃO AMBIENTAL URBANA: EXPERIÊNCIAS E PERSPECTIVAS*, 1., 2004, Brasília, DF. **Anais**. Brasília, DF: UnB, 2004.
- ROSS, N. A. *et al.* Dimensions and dynamics of residential segregation by income in urban Canada: 1991-1996. **Canadian Geographer**, n. 48, p. 433-445, 2004.

- SAKODA, J. N. A generalized Index of dissimilarity. **Demography**, n. 18, p. 245-250, 1981.
- SEIDEL, R. A simple and fast incremental randomized algorithm for computing trapezoidal decompositions and for triangulating polygons. **Computational Geometry: theory and applications**, n. 1, p. 51-64, 1995.
- SUZIGAN, W. *et al.* Coeficientes de Gini Locacionais (GL): aplicação à indústria de calçados do estado de São Paulo. **Nova Economia**, v. 13, n. 2, p. 39-60, jul./dez. 2003.
- TANGO, T. A class of tests for detecting general and focused clustering of rare diseases. **Statistics in Medicine**, n. 7, p. 649-660, 1955.
- THEIL, H. **Statistical Decomposition Analysis**. Amsterdam, North-Holland, 1972.
- WALLER *et al.* A chronic disease surveillance and testing of clustering of disease and exposure application to leukemia incidence and TCE-contaminated dumpsites in upstate New York. **Environmetrics**, n. 3, p. 281-300, 1992.
- WHITE, M. J. Segregation and diversity measures in population distribution. **Population Index**, n. 52, p. 198-221, 1986.
- WONG, D. W. S. Implementing spatial segregation measures in GIS. **Computers, Environment and Urban Systems**, v. 27, p. 53-70, 2003.
- _____. Geostatistics as measures of spatial segregation. **Urban Geography**, n. 20, p. 635-647, 1999.
- _____. Measuring multiethnic segregation. **Urban Geography**, v. 19, p. 77-87, 1998.
- _____. Enhancing segregation studies using GIS. **Computers, Environment and Urban Systems**, n. 20, p. 99-109, 1996.

ANEXOS

FIGURA 1A

Ferramentas visuais do IpeaGEO – o zoom e a informação sobre o polígono

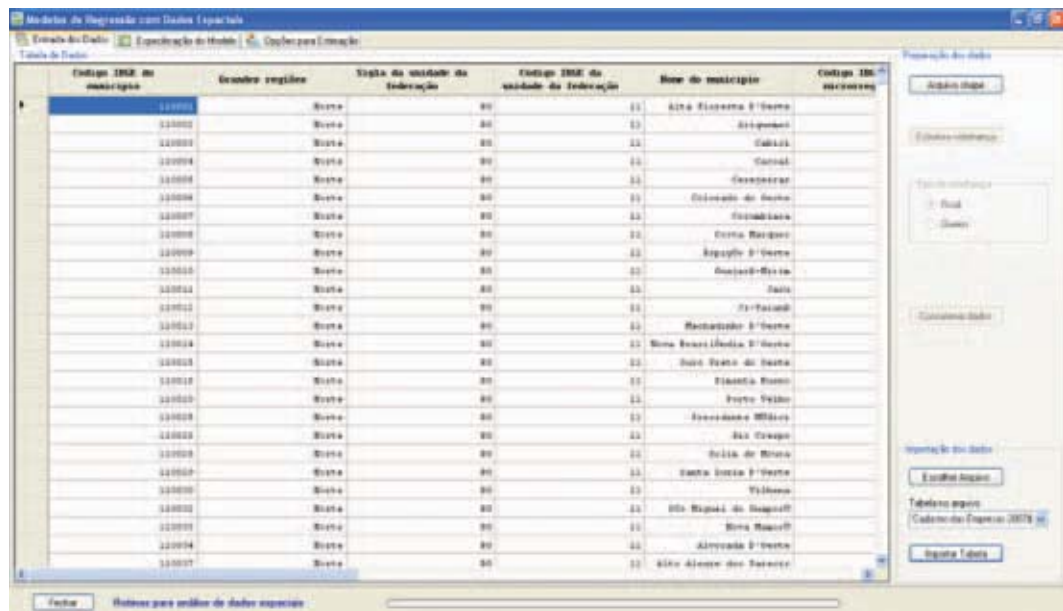


Fonte: IpeaGEO.

Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 2

Formulário presente na parte de ferramentas



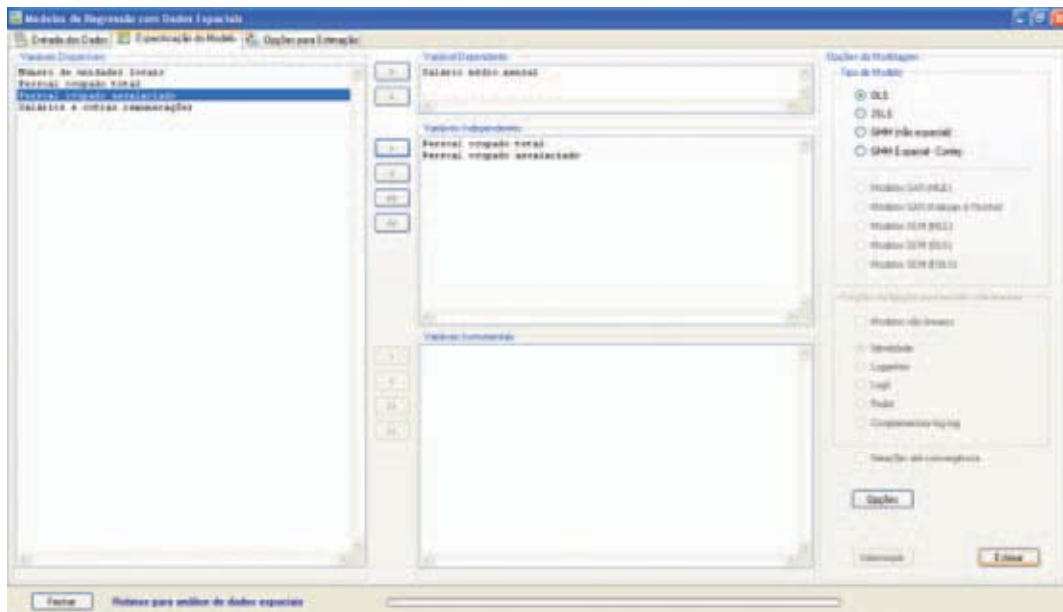
Fonte: IpeaGEO.

Obs.: Para se fazer uma regressão, é utilizado um formulário a parte do IpeaGEO em que não é obrigatório ter um arquivo *shape*.

A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 3

Estimativa dos parâmetros de regressão utilizando OLS, 2OLS, GMM não espacial e GMM espacial com o programa

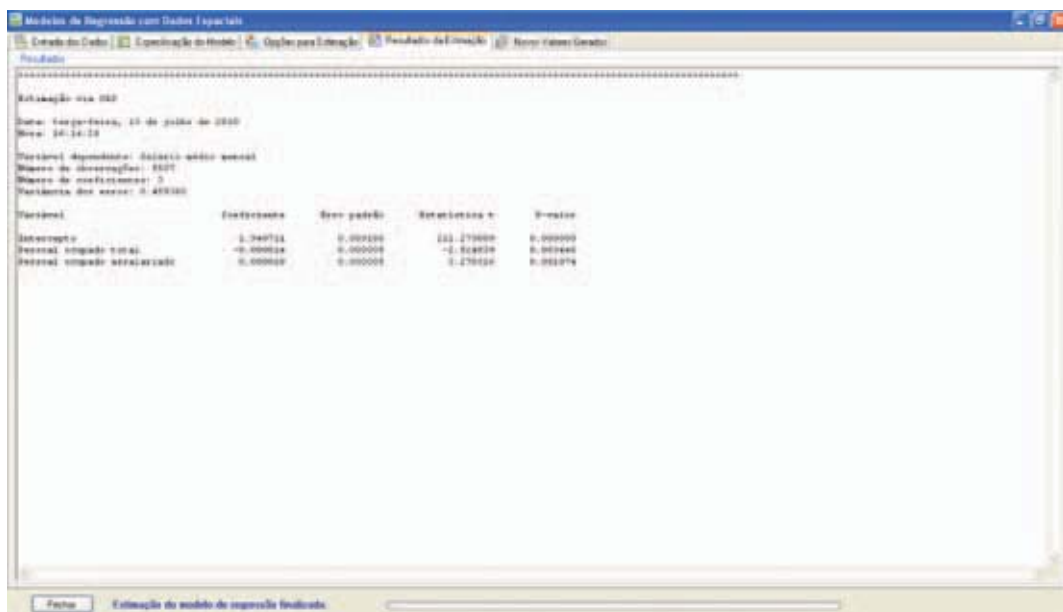


Fonte: IpeaGEO.

Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 4

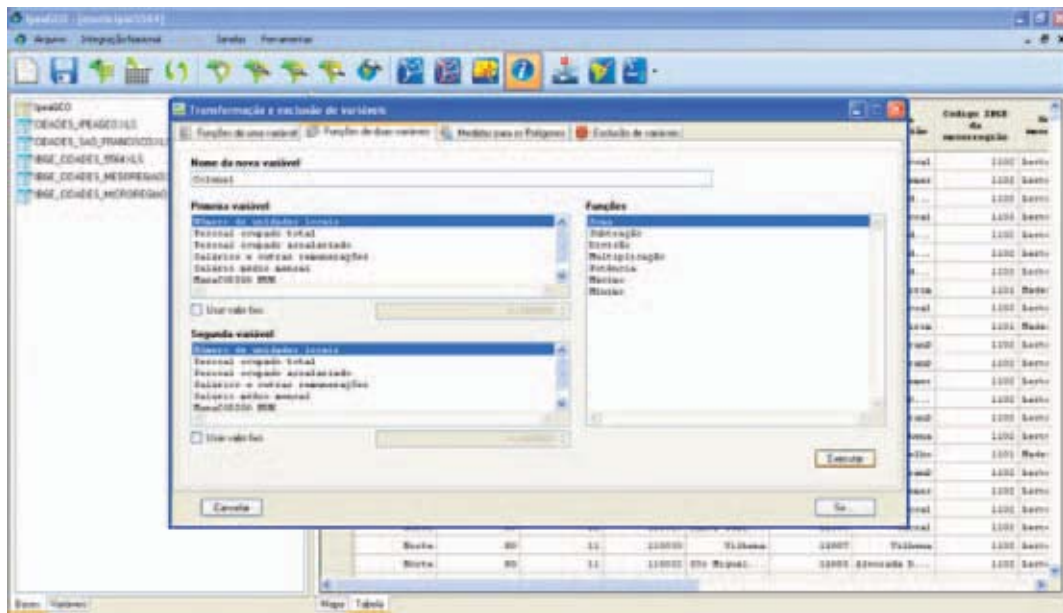
Resultado da regressão



Fonte: IpeaGEO.

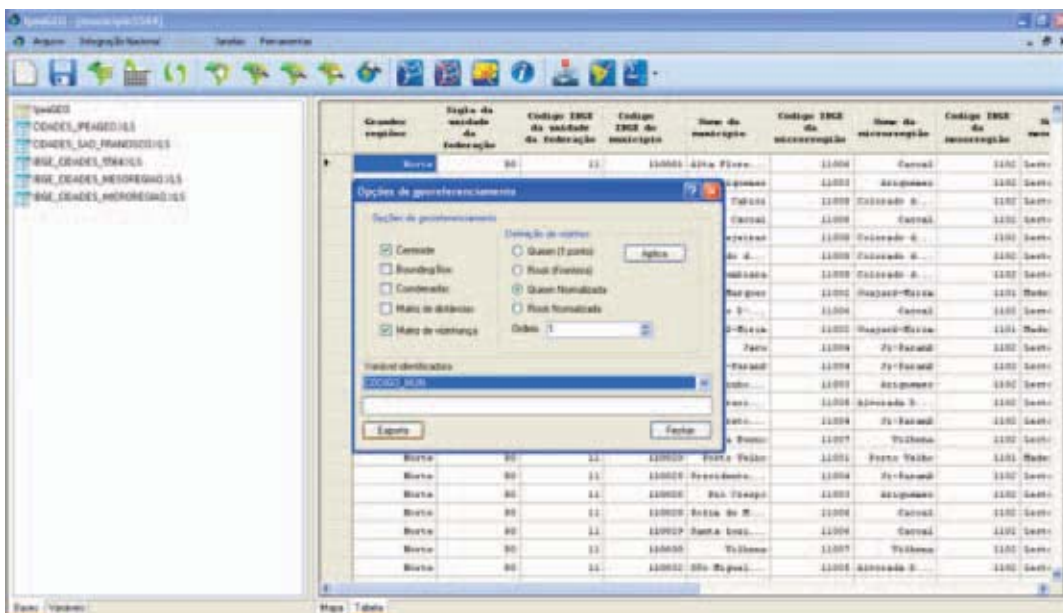
Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 5
Função de calculadora



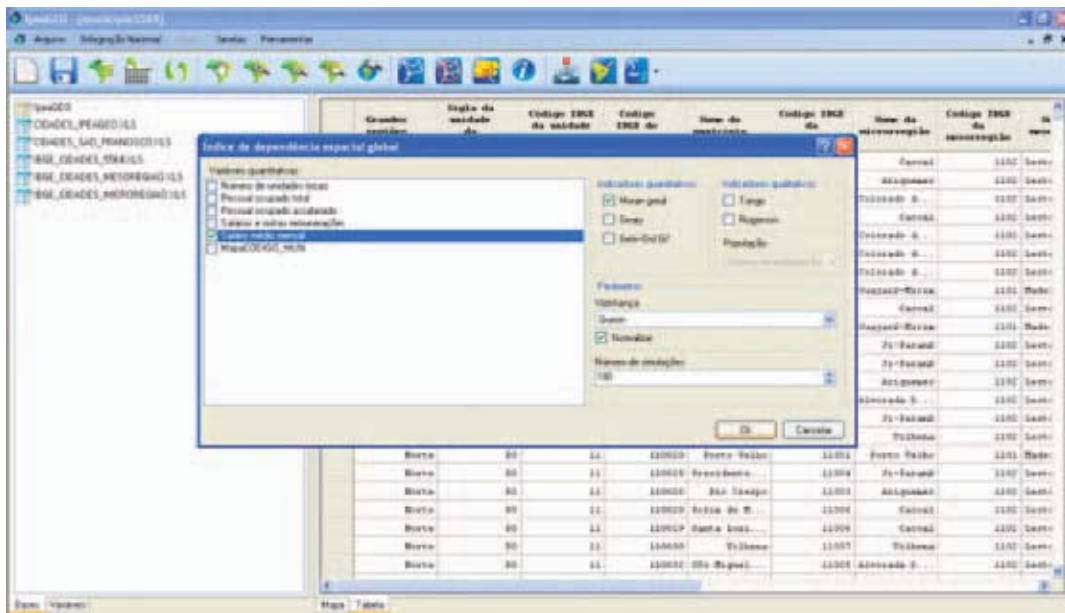
Fonte: IpeaGEO.
Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 6
Ferramenta de georreferenciamento



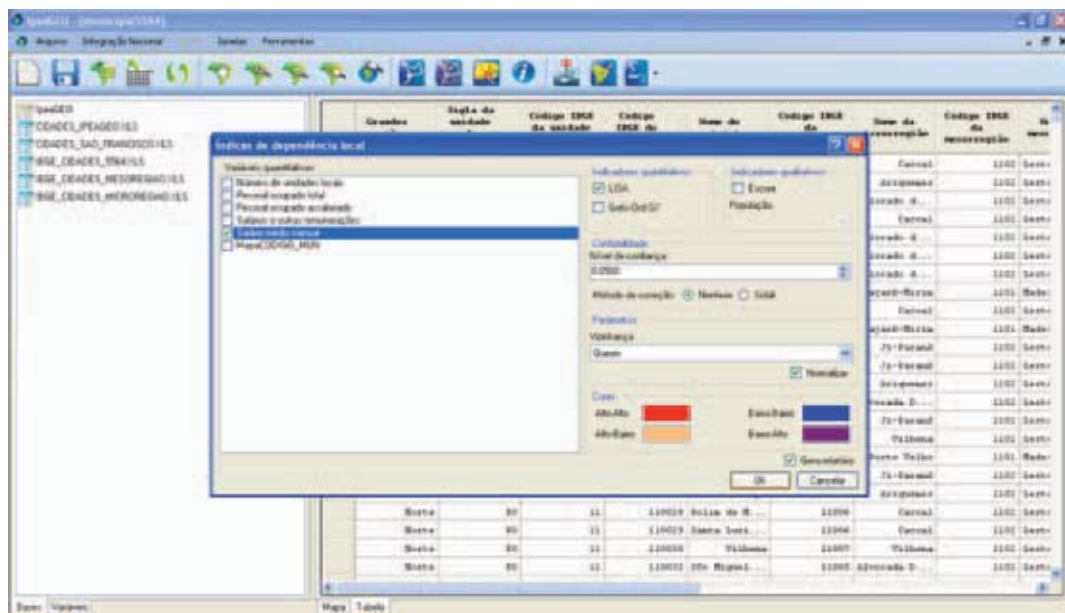
Fonte: IpeaGEO.
Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 7
Geração dos índices de dependência global



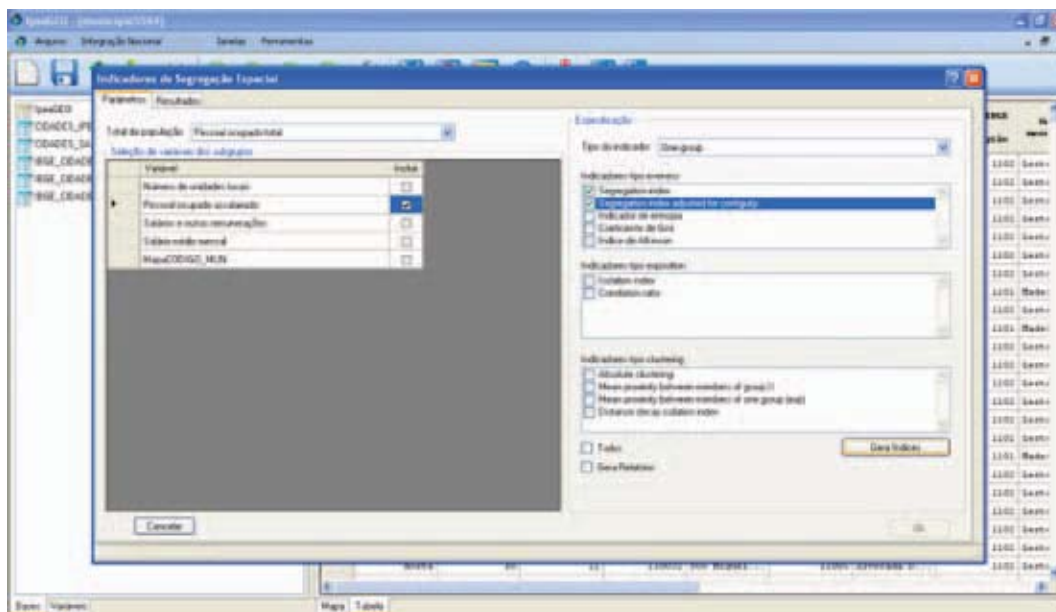
Fonte: IpeaGEO.
Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 8
Geração dos índices de dependência local



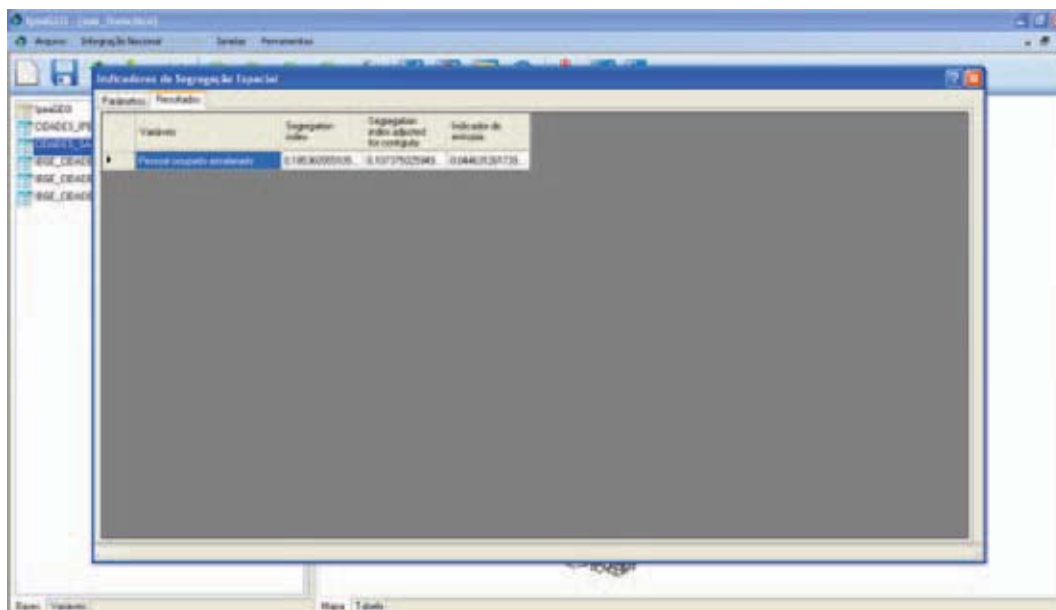
Fonte: IpeaGEO.
Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 9
Índices de segregação



Fonte: IpeaGEO.
Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

FIGURA 10
Resultados gerados



Fonte: IpeaGEO.
Obs.: A figura está reproduzida conforme o original fornecido pelos autores, cujas características não permitiram melhor ajuste para fins de impressão (nota do Editorial).

Ipea – Instituto de Pesquisa Econômica Aplicada

Editorial

Coordenação

Cláudio Passos de Oliveira

Njobs Comunicação

Supervisão

Cida Taboza

Fábio Oki

Thayse Lamera

Revisão

Ângela de Oliveira

Cristiana de Sousa da Silva

Lizandra Deusdará Felipe

Regina Marta de Aguiar

Editoração

Anderson Reis

Daniela Rodrigues

Danilo Tavares

Marília Assis

Patrícia Dantas

Capa

Eduardo Grisoni

Livraria

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo

70076-900 – Brasília – DF

Tel.: (61) 3315 5336

Correio eletrônico: livraria@ipea.gov.br

I S S N 2177- 1847



9 772177 184005

Ministério da
Integração Nacional

