

## Análise exploratória espacial aplicada ao cultivo de arroz no Brasil

Exploratory analysis applied to Brazilian rice crop

Letícia Marasca<sup>I</sup>, Steffani Nikoli Dapper<sup>II</sup>, Adriano Mendonça Souza<sup>III</sup>,  
 Cláudia Aline de Souza Ramser<sup>IV</sup>

### RESUMO

O Brasil ocupa papel de destaque na produção de arroz, sendo o Rio Grande do Sul o maior estado produtor. O cultivo do arroz representa um papel estratégico no país, tanto economicamente como no âmbito social. O objetivo dessa pesquisa é determinar a autocorrelação espacial da produtividade agrícola do arroz entre as microrregiões brasileiras, utilizando a Análise Exploratória dos Dados Espaciais (AEDE). Como resultados, concluiu-se que a produtividade agrícola do arroz está distribuída de maneira heterogênea no país. Os valores obtidos pela estatística I de Moran possibilitaram rejeitar a hipótese nula de aleatoriedade espacial da produtividade agrícola do arroz, permitindo a identificação de clusters de alta e baixa produtividade.

**Palavras-chave:** Produtividade agrícola do arroz. Autocorrelação espacial. Análise exploratória de dados espaciais

### ABSTRACT

Brazil stands out in world agricultural production and Rio Grande do Sul is the largest producing state. Rice cultivation plays a strategic role in the country, both economically and socially. The objective of this research is to determine the spatial autocorrelation of rice agricultural productivity among Brazilian microregions, using the Exploratory Spatial Data Analysis (AEDE). As results, it was concluded that rice agricultural productivity is heterogeneously distributed in the country. The values obtained by I Moran statistics made it possible to reject the null hypothesis of spatial randomness of rice agricultural productivity, allowing the identification of high and low yield clusters.

**Keywords:** Rice agricultural productivity. Spatial autocorrelation. Exploratory analysis of spatial data

<sup>I</sup> Universidade Federal de Santa Maria, Brasil; e-mail: leticiamarasca@yahoo.com.br;

<sup>II</sup> Universidade Federal de Santa Maria, Brasil; e-mail: steffani.dapper@fisma.com.br;

<sup>III</sup> Universidade Federal de Santa Maria, Brasil; e-mail: amsouza.sm@gmail.com;

<sup>IV</sup> Universidade Federal de Santa Maria, Brasil; e-mail: claudiaramser@hotmail.com;



## 1 INTRODUÇÃO

O cultivo do arroz desempenha um estratégico papel social e econômico. De grande valor econômico, considerado o cultivo mais importante em diversos países em desenvolvimento, como a Ásia e a Oceania, que possuem 70% do total da população em desenvolvimento do mundo e mais de 60% da população subnutrida, segundo a Agência Embrapa de Informação Tecnológica - AGEITEC (2017). A Ásia tem um papel de destaque, com 90% de toda produção mundial de arroz produzida e consumida em seu território (AGEITEC, 2017).

É um dos alimentos de grande balanceamento nutricional, fornecendo 1/5 da energia e 15% da proteína necessária para o ser humano. É considerado uma cultura multifacetada, se adaptando a diferentes solos e climas, sendo por esse motivo considerado o grão com maior potencial para combater a fome do mundo (EMBRAPA, 2017). Com 75% de sua produção por sistema de cultivo irrigado, é o alimento base de 2,4 bilhões de pessoas, com estimativas para atender o dobro dessa população nos próximos 30 anos (MDIC, 2018).

O Brasil ocupa papel de destaque, estando entre os 10 principais produtores de arroz do mundo, sendo o maior produtor mundial fora do continente Asiático (POERSCHKE; MORAIS, 2014). No Brasil, é o segundo grão em extensão de área cultivada, superado apenas pelo trigo.

O cultivo desse grão no Brasil representa 75% da produção nacional. Essa produção se dá no país principalmente pela orizicultura irrigada, e a produção é considerada como estabilizadora da safra do país, pois é um tipo de cultura que não é tão dependente do clima para se desenvolver.

Segundo Klering et al. (2013), o estado do Rio Grande do Sul é um dos maiores produtores e exportadores de diversos grãos, respondendo por mais de 60% da produção brasileira de arroz (mais de 8 milhões de toneladas), possuindo uma área plantada de mais de 1 milhão de hectares.

De acordo com Zanin (2013) o aumento da produção no país se deu pelos ganhos de produtividade, uma vez que a área utilizada para o plantio de arroz diminuiu ao longo do tempo. O Rio Grande do Sul, maior estado produtor, é considerado o principal

responsável pela industrialização desse grão, quando se fala em unidades beneficiadores e também em valores de venda da indústria.

Essa pesquisa objetiva determinar a autocorrelação espacial da produtividade agrícola do arroz entre as microrregiões brasileiras, utilizando para isso a metodologia da Análise Exploratória dos Dados Espaciais (AEDE), visando contribuir para um maior entendimento sobre os efeitos espaciais da produtividade agrícola do arroz no país.

Além da introdução, a pesquisa está estruturada da seguinte maneira: o segundo item, a metodologia, expõe o banco de dados utilizado, além do processo metodológico; o terceiro item, os resultados e discussões, apresenta os resultados obtidos na pesquisa; seguido pelo quarto item, as conclusões; e por fim, os agradecimentos e as referências.

## 2 METODOLOGIA

O banco de dados utilizado nessa pesquisa corresponde à produtividade agrícola do arroz referente as 558 microrregiões brasileiras, compreendendo os anos de 1992, 1997, 2002, 2007, 2012 e 2017, totalizando 3.348 observações anuais. O banco de dados e o *shapefile* foram obtidos no *site* do Instituto Brasileiro de Geografia e Estatística (IBGE), na pesquisa de Produção Agrícola Municipal (PAM) do Sistema IBGE de Recuperação Automática (Sidra).

Os softwares utilizados na modelagem foram o *GeoDa* para as estatísticas espaciais e o *Qgis* para manipulação dos mapas. Para realização da Análise Exploratória dos Dados Espaciais, o primeiro passo foi realizar a estimação do Modelo Clássico de Regressão Linear (MCRL), pelo método dos Mínimos Quadrados Ordinários para a variável produtividade agrícola do arroz. Após estimação do MCRL, ocorreu a análise dos resíduos do modelo de regressão, utilizou-se, para isso, a estatística I de *Moran*. Seus resultados indicaram a rejeição da hipótese nula de aleatoriedade espacial, com a confirmação da autocorrelação espacial positiva (PASCHOALINO et al., 2017).

Após a confirmação da autocorrelação espacial, definiu-se a matriz de pesos espaciais (W), selecionando a matriz que melhor captou a autocorrelação espacial presente nos dados em análise, sendo a que apresentou o maior valor para a estatística I de *Moran* (ALMEIDA; PEROBELLI; FERREIRA, 2008).

Nessa pesquisa a matriz selecionada foi a matriz binária de distância geográfica, onde regiões próximas geograficamente tem interação espacial maior do que as regiões mais distantes. Utilizou-se a matriz dos  $k$  vizinhos mais próximos (*k-nearest neighbors weight matrix*) ( $W_{ij}(k)$ ). Tem-se então a Equação (1):

$$W_{ij}(k) = \begin{cases} 1, & \text{se } d_{ij} \leq d_i(k) \\ 0, & \text{se } d_{ij} > d_i(k) \end{cases} \quad (1)$$

Onde:  $d_i(k)$  é a distância de corte para a região  $i$ , a fim de que a região  $i$  tenha  $k$  vizinhos;  $W_{ii}(k) = 0$ .

A vantagem da utilização dessa forma de convenção é o equilíbrio da conectividade da matriz, pois todas terão o mesmo número de vizinhos e também não haverá regiões sem nenhum vizinho (ALMEIDA, 2012). Após a definição da melhor matriz que captou a autocorrelação ajustada ao banco de dados, testou-se a hipótese da distribuição aleatória dos mesmos. O primeiro coeficiente de autocorrelação espacial foi proposto por Moran (1948) e está descrito na Equação (2):

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j W_{ij} Z_i Z_j}{\sum_{i=1}^n Z_i^2} \quad (2)$$

Os Indicadores Locais de Associação Espacial ou *Local Indicator of Spatial Association* - LISA foram gerados para definir o índice local de associação espacial, sendo uma decomposição do índice global de autocorrelação espacial (ANSELIN, 1995). O coeficiente LISA está expresso conforme Equação (3):

$$I_i = Z_i \sum_{j=1}^j W_{ij} Z_j \quad (3)$$

As informações dos indicadores LISA são mapeadas no mapa de significância LISA *map*, obtendo-se assim o mapa de *clusters* (TEIXEIRA; BERTELLA, 2015).

Todas as etapas descritas anteriormente compreendem a metodologia da AEDE, parte exploratória da econometria espaciais e permitiram identificar os padrões espacial da produtividade agrícola do arroz entre as microrregiões brasileiras.

### 3 Resultados e discussões

A autocorrelação espacial global para a variável produtividade agrícola do arroz nos anos de 1992, 1997, 2002, 2007, 2012 e 2017 foi calculada e o valor esperado para a estatística I de *Moran* é -0.00179. Com base nesse valor são analisadas suas estatísticas calculadas em todos os anos em análise (Tabela 1):

Tabela 1 – Valor esperado e calculado do Índice I de Moran\*

Variável	Ano	I de <i>Moran</i> Esperado	I de <i>Moran</i> Calculado	Matriz
Arroz	1992	-0.00179533213644	0.809247	k (3)
	1997	-0.00179533213644	0.752341	k (5)
	2002	-0.00179533213644	0.739656	k (3)
	2007	-0.00179533213644	0.754401	k (2)
	2012	-0.00179533213644	0.766171	k (4)
	2017	-0.00179533213644	0.707435	k (5)

Fonte: Organização do autor (2018).

\*H0: aleatoriedade espacial; H1: autocorrelação espacial

Com a análise da Tabela 1, observa-se que, para todos os anos selecionados nessa pesquisa, o I de *Moran* calculado é maior que o seu valor esperado, o que confirma a presença de autocorrelação espacial, com a rejeição da hipótese nula de aleatoriedade dos dados. A presença de autocorrelação espacial positiva é confirmada pelos coeficientes positivos e indica similaridade entre a variável produtividade agrícola do arroz e sua localização espacial.

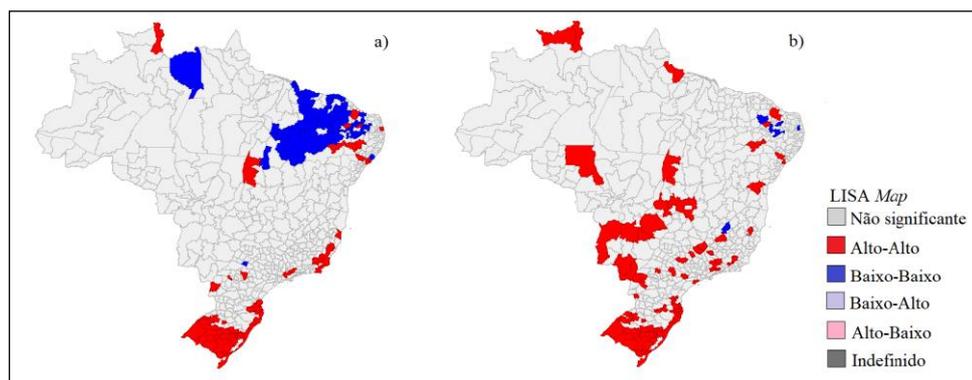
Assim confirma-se a formação de *clusters* de alta e baixa produtividade agrícola do arroz entre as microrregiões brasileiras. Isso significa que microrregiões com alta produtividade agrícola do arroz estão agrupadas em uma região do país, com regiões vizinhas com valores também altos e, de maneira análoga, microrregiões com baixa produtividade agrícola de arroz estão agrupadas em outra região, rodeadas por outras microrregiões com valores também baixos.

Com a análise da matriz de ponderação espacial (Tabela 1), observa-se que todas as selecionadas seguiram a convenção dos *k* vizinhos mais próximos (*k-nearest neighbors weight matrix*) e chegaram ao mesmo resultado em relação a sua significância e ao sinal, o que sugere bons resultados na sua escolha.

A produtividade agrícola do arroz apresentou autocorrelação espacial em todos os períodos em análise. Diversos fatores políticos, econômicos e climáticos influenciaram tanto positivamente quanto negativamente a variável em análise e esses serão analisados nessa pesquisa.

A Figura 1 apresenta os LISA maps referentes à produtividade agrícola do arroz para os dois primeiros anos analisados.

Figura 1 - LISA map das microrregiões brasileiras para a produtividade agrícola do arroz nos anos de 1992 e 1997



Fonte: Organização do autor (2018).

O mapa a) (Figura 1) corresponde à produtividade agrícola do arroz das microrregiões brasileiras para o ano de 1992. Nele, pode-se observar que sessenta e uma microrregiões compõem três clusters na cor vermelho intenso (microrregiões caracterizadas como alto-alto) e cinquenta e sete microrregiões compõem um cluster azul intenso (microrregiões caracterizadas como baixo-baixo), além de quatrocentas e quarenta microrregiões que não apresentaram o índice I de Moran como significativo.

O mapa b) (Figura 1) também corresponde à produtividade agrícola do arroz, mas para o ano de 1997. Nele observa-se oitenta e duas microrregiões que formam cinco clusters caracterizados como alto-alto, em vermelho intenso e quarenta e duas microrregiões compõem os três clusters caracterizado como baixo-baixo (azul-intenso), além de quatrocentas e trinta e quatro microrregiões sem significância.

O período que compreende o ano de 1992 foi de grande aceleração inflacionária no país, com taxa de câmbio valorizada no período, prejudicando a agricultura do país como um todo, reduzindo as receitas. Nesse período foi instaurado o regime de flutuação suja, onde o mercado estabelecia a taxa de câmbio. Isso causou uma

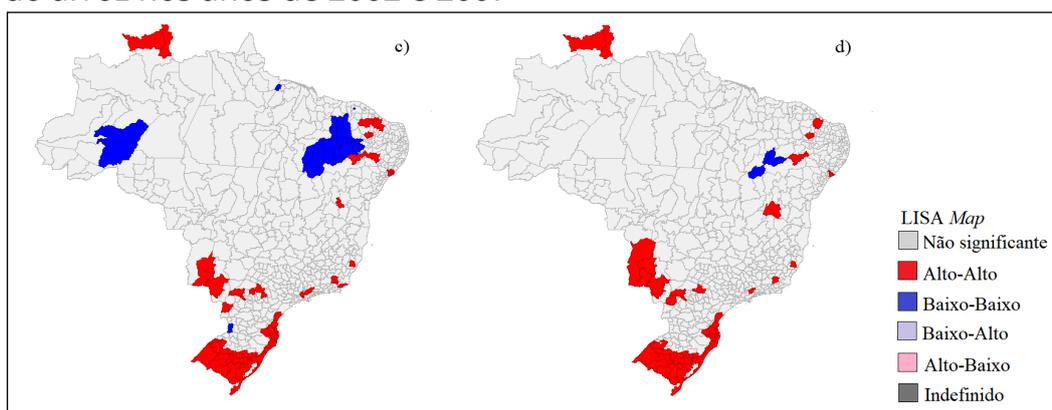
desvalorização no período de 1991 e 1992 e uma valorização no ano de 1993, o que equilibrou o mercado.

No início dos anos 90, também correspondente ao ano de 92 em análise, ocorreu a criação de financiamentos privados, entre eles o Certificado do Depósito Agrícola (CDA) e a Célula do Produtor Rural, além da possibilidade de renegociação das dívidas rurais (BACHA; DANELON; FILHO, 2006). No ano de 1992 ocorreu também o impeachment do Presidente Fernando Collor de Mello, aumentando significativamente a instabilidade econômica e política no país.

O próximo ano analisado, de 1997, foi marcado por um aumento nos subsídios oferecidos, depois de um longo período de cortes. Foi nesse ano, de 1997, que foi criado o sistema de minibandas cambiais, ocorrendo uma valorização cambial no país, favorecendo a compra de insumos e equipamentos. Foi também nesse ano, 1997, que ocorreu uma grande seca, sentida inclusive em anos posteriores.

Apesar de todos os problemas citados, o setor apresentou crescimento no período, devido ao aumento da produtividade da terra, da mão-de-obra e do uso de novas tecnologias. A Figura 2 apresenta os LISA maps da produtividade agrícola do arroz referentes aos anos de 2002 e 2017.

Figura 2 - LISA map das microrregiões brasileiras para a produtividade agrícola do arroz nos anos de 2002 e 2007



Fonte: Organização do autor (2018).

No mapa a) (Figura **Erro! Fonte de referência não encontrada.**) representa a produtividade agrícola do arroz para o ano de 2002. Nele observa-se que sessenta microrregiões formam seis *clusters* caracterizados como alto-alto, em vermelho intenso

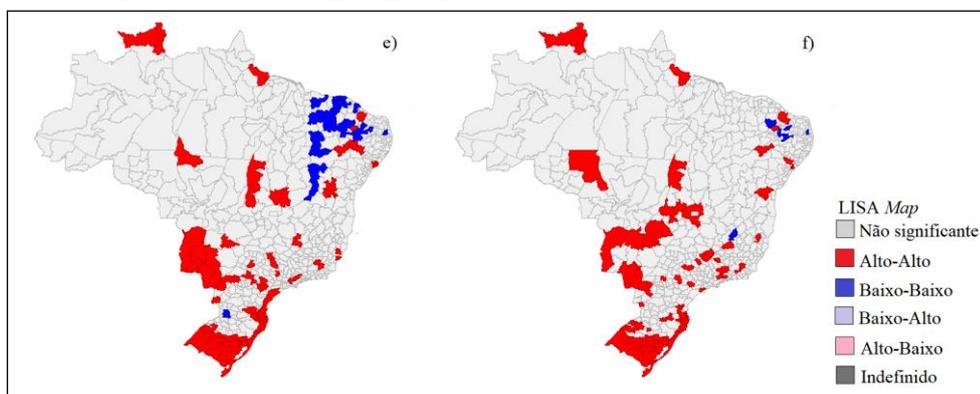
e dezenove microrregiões que compõem um *cluster* caracterizado como baixo-baixo m azul intenso, além de quatrocentas e setenta e nove microrregiões não significantes.

O mapa b) (Figura **Erro! Fonte de referência não encontrada.**) corresponde à produtividade agrícola do arroz para o ano de 2007. Nele, cinquenta e uma microrregiões formam três *clusters* alto-alto, em vermelho intenso, e duas microrregiões estão classificadas como baixo-baixo (azul intenso), além de quinhentas e cinco microrregiões não significativas.

No ano de 2002, ocorreu outra grande desvalorização cambial, que estimulou as exportações no país. A partir desse período que foram criados programas de financiamento de máquinas agrícolas. No ano de 2007 a China se tornou o maior parceiro comercial do Brasil, sendo esses países parceiros desde 1949, com a fundação da República da China (MARCOS et al., 2011).

A Figura Figura 2 apresenta os LISA *maps* da produtividade agrícola do arroz caracterizando cada microrregião para os últimos anos em análise (2012 e 2017).

Figura 2 - LISA *map* das microrregiões brasileiras para a produtividade agrícola do arroz nos anos de 2012 e 2017



Fonte: Organização do autor (2018).

No ano de 2012 (mapa a), Figura 3) a produtividade agrícola do arroz se distribui em setenta e oito microrregiões identificadas em vermelho intenso, compondo dez clusters alto-alto, trinta e quatro microrregiões em azul intenso, caracterizadas como baixo-baixo, compondo três clusters, além de quatrocentos e quarenta e seis microrregiões não significativas.

O mapa b) (Figura 3) corresponde à produtividade agrícola do arroz para o último ano em análise nessa pesquisa, de 2017. Nele observa-se que setenta e oito

microrregiões compõem oito clusters caracterizados como alto-alto, em vermelho intenso no mapa e sete microrregiões se caracterizam como baixo-baixo, em azul intenso, além de quatrocentas e setenta e três microrregiões que não apresentaram valores significativos.

O ano de 2012 foi um ano recorde na venda de equipamentos agrícolas. Foi nesse ano que ocorreu uma grande quebra de safra, devido ao fenômeno climático “Al Niña”, com grande estiagem, considerada a mais intensa em diversos estados, e se estendeu até o último ano em análise, de 2017. Esse ano foi bastante instável politicamente, pois no ano anterior ocorria o impeachment da presidente Dilma Rousseff.

Nas Figura 1, 2 e 3 observa-se os LISA maps caracterizando a produtividade agrícola do arroz para cada microrregião em análise. Esses mapas de agrupamentos locais corroboram os resultados da Tabela 1, pois neles pode-se observar a formação de clusters de alta e baixa produtividade, sendo esses clusters uma característica da presença de autocorrelação espacial positiva.

Analisando-se o primeiro e o último ano para fins de comparação (Figura 1 mapa a e Figura 3 mapa b), observa-se um aumento na produtividade agrícola do arroz entre as microrregiões brasileiras.

#### **4 CONCLUSÕES**

A produtividade agrícola do arroz está distribuída de maneira heterogênea no país, com a dependência espacial presente tanto de maneira global como local, confirmando a importância da utilização da metodologia da Análise Exploratória de Dados Espaciais.

Os valores obtidos pela estatística I de Moran possibilitam rejeitar a hipótese nula de aleatoriedade espacial da produtividade agrícola do arroz, o que sugere a formação de clusters alto-alto e clusters baixo-baixo, indicando que microrregiões com alta produtividade são vizinhas de outras microrregiões com produtividade também alta para a variável arroz, e de maneira análoga, microrregiões que apresentam baixa

produtividade se encontram em uma região do mapa com vizinhos apresentando produtividade agrícola do arroz também abaixo da média.

A observação dos mapas de significância LISA maps permite a identificação de clusters de alta e baixa produtividade. Esses agrupamentos podem ser explicados pelo clima favorável a determinadas culturas em algumas regiões do país, pelos incentivos governamentais, em abundância em algumas regiões, pela convergência da produtividade do arroz, que acaba aumentando a produção em algumas áreas que antes não recebiam cultivo, terras com características específicas, favoráveis a determinadas culturas.

Os resultados obtidos possibilitam tomar decisões acerca de investimentos governamentais, levando em consideração os padrões espaciais identificados, que possibilitou identificar áreas de maior e menor produtividade.

Para estudos futuros, sugere-se a utilização da Análise Exploratória de Dados Espaciais em outras culturas, para fins de comparação de resultados.

## **AGRADECIMENTOS**

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e também do Laboratório de Análise e Modelagem Estatística - LAME, da Universidade Federal de Santa Maria - UFSM.

## REFERÊNCIAS

AGÊNCIA EMBRAPA DE INFORMAÇÃO TECNOLÓGICA - AGEITEC [Internet]. Disponível em: <<http://www.agencia.cnptia.embrapa.br/>>.

ALMEIDA, E. S. de; PEROBELLI, F. S.; FERREIRA, P. G. C. Existe convergência espacial da produtividade agrícola no Brasil? **Revista de Economia e Sociologia Rural**. 2008;46(1):31-52.

ALMEIDA, E. **Econometria espacial aplicada**. Campinas: Alínea, 2012.

ANSELIN, L. Local Indicators of Spatial Association—LISA. **Geographical Analysis**. 1995;27(2):93-115,.

BACHA, C. J. C.; DANELON, L.; FILHO, E. D. B. Evolução da taxa de juros real do crédito rural no Brasil - período de 1985 a 2003. **Teoria e evidência econômica**. 2006;14(26):43-69.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA [Internet]. Disponível em: <<https://www.embrapa.br/>>.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE [Internet]. Disponível em: <<https://www.ibge.gov.br/>>.

KLERING, E. V.; FONTANA, D. C.; ALVES, R.; ROCHA, J.; BERLATO, M. A. Estimativa de área cultivada com arroz irrigado para o estado do rio grande do sul a partir de imagens Modis. **Ciência e Natura**. 2013;35(2):126-135.

MARCOS, C.; SÍLVIA, M.; GALVÃO DE MIRANDA, H.; RUMENOS, M.; BACCHI, P. Determinantes do comércio Brasil-China de Commodities e produtos industriais: uma aplicação VECM. **Economia Aplicada**. 2011;15(2):311-335.

MINISTÉRIO DA INDÚSTRIA, COMÉRCIO EXTERIOR E SERVIÇOS - MDIC [Internet]. Disponível em: <<http://www.mdic.gov.br/>>.

MORAN, P. A. The interpretation of statistical maps. **Journal of the Royal Statistical Society**. 1948;10(2):243-251.

PASCHOALINO, P. A. A. T.; MICHELLON, E.; BERNARDELLI, L. V.; GOBI, J. R.; SANTOS, L. G. de O. Análise espacial da produtividade do feijão nos municípios paranaenses entre 2006 e 2014. **Economia & Região**. 2017;5(1):107-123.

POERSCHKE, R. P.; MORAIS, I. A. C. de. Determinantes da demanda brasileira por importação de arroz: uma abordagem não linear. **Revista de Economia e Sociologia Rural**. 2014;52(1):177-194.

TEIXEIRA, R. F. A. P.; BERTELLA, M. A. Distribuição espaço-temporal da produtividade média do café em Minas Gerais: 1997-2006. **Análise Econômica**. 2015;33(63):275-299.

ZANIN, V. Panorama geral da orizicultura brasileira. **Indicadores Econômicos FEE**. 2013;41(2):51-66.