**Produtibilidade de soja no estado do Rio Grande do Sul: uma análise estatística espacial de 1981 a 2010**

**Resumo**

O objetivo do presente artigo é verificar quais fatores estão associados ao aumento ou perda da produtibilidade de soja no estado do Rio Grande do Sul (RS), através de uma análise de dados estatísticos espaciais. Para viabilizar o estudo foram utilizados dados da World Clim de temperatura, precipitação, umidade do solo, altitude e evapotranspiração, do qual se analisou a produtibilidade de soja através de uma análise de autocorrelação, utilizando-se Índice de Moran global e local, além de uma Regressão Geograficamente Ponderada (GWR). Os resultados obtidos sugerem que os dados analisados apresentam uma autocorrelação espacial positiva significativa, sendo estatisticamente relevante e não podendo ser explicada pelo acaso. Como também variável resposta Valor de Produção pode estar fortemente correlacionada com as outras variáveis no modelo.

Palavras-chave: Produtibilidade de soja – Autocorrelação – Índice de Moran – Regressão Geograficamente Ponderada

**Abstract**

The objective of this article is to verify which factors are associated with the increase or loss of soybean productivity in the state of Rio Grande do Sul (RS), through an analysis of spatial statistical data. To make the study viable, data from the World Clim of temperature, precipitation, soil moisture, altitude and evapotranspiration were used, from which the soybean productivity was analyzed through an autocorrelation analysis, using global and local Moran Index, in addition to a Geographically Weighted Regression (GWR). The obtained results suggest that the analyzed data present a significant positive spatial self-correction, being statistically relevant and cannot be explained by chance. As well as the Production Value response variable, it can be strongly correlated with the other variables in the model.

Keywords: Soybean productivity – Autocorrelation – Moran Index – Geographically Weighted Regression

**Introdução**

A soja (Glycine Max) é uma planta originária da costa leste da Ásia. Diferentemente da que é cultivada hoje, ancestralmente eram plantas rasteiras que se desenvolviam principalmente ao longo do rio Yangtse, na China. Foi devido a cruzamentos naturais entre duas espécies selvagens que ela foi sendo domesticada e melhorada por cientistas da antiga China. Seus primeiros grãos surgiram no período entre 2883 e 2838 AC, chegando a soja a ser considerada um grão sagrado. E ainda para alguns autores a soja é mais antiga, citada no Livro de Odes, que foi publicado em chinês arcaico (EMBRAPA).

Na Europa, a soja surgiu no final do século XV, nos jardins botânicos da Inglaterra, França e Alemanha. E em meados do século XX as indústrias mundiais começam a se interessar pelo teor de óleo e proteína do grão. Entretanto, na Rússia, Inglaterra e Alemanha as iniciativas de implementação comercial da soja falharam, possivelmente em função das adversas condições climáticas (EMBRAPA).

Trazida dos Estados Unidos, alguns semeares começaram a chegar no Brasil no final do século XIX para realização de pesquisas como planta forrageira. Os primeiros plantios surgiram no século XX em São Paulo e Rio Grande do Sul, e o cultivo em grande escala no ano de 1950 na região Sul e Sudeste (SIQUEIRA, 2004).

No final da década de 60, foi vista como um produto comercial no país desencadeada por duas razões. Era uma opção de verão na sequência da cultura do trigo, além de que havia uma iniciativa na produção de aves e suínos, o que demandava, portanto, pelo farelo de soja (EMBRAPA).

E foi nos estados do Rio Grande do Sul, Santa Catarina e Paraná que a produção da soja se expandiu nesse período, devido a fatores climáticos favoráveis da Região Sul que promovem condições edafoclimáticas para a produção do grão (PELLENZ; ALMEIDA; FREITAS, 2019).

Em termos de produção, segundo a Food and Agriculture Organization (FAO) o Brasil foi o maior produtor mundial de soja em 2020, produzindo 122 milhões de toneladas, dentre todo o continente americano, que foi responsável por 86,2% de toda a soja produzida no mundo (ATLAS, 2022).

O país ainda ocupa uma importante posição no mercado internacional, sendo considerado um dos maiores produtores de commodities, devido a privilégios advindos de fatores edafoclimáticos, ambientais, culturais e do fato de utilizar tecnologias de alto padrão na produtividade de suas culturas, promovendo assim maiores ganhos (PELLENZ; ALMEIDA; FREITAS, 2019).

Entre os estados da federação que mais produz soja, o terceiro maior do país é o Rio Grande do Sul, ficando atrás apenas dos estados de Mato Grosso e Paraná. Nas regiões norte-noroeste e centro encontram-se as áreas mais produtivas do Rio Grande do Sul (ATLAS,2022).

**A Soja no Brasil e Rio Grande do Sul**

Segundo dados econômicos divulgados pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), a produção de soja no Brasil foi de 123.829,5 milhões de toneladas na safra de 2021/22, em uma área plantada de 40.921,9 milhões de hectares, sendo uma produtividade de 3,026 kg/ha. Em relação à produção de soja no mundo no mesmo período, que foi de 355,588 milhões de toneladas, o Brasil é o maior produtor mundial, produzindo mais de 34% do total (EMBRAPA).

Desde 1970 o Brasil investe amplamente em pesquisa, e associado ao empreendedorismo dos agricultores vem resultando no sucesso da sojicultura brasileira. Nas últimas cinco décadas esta foi a que mais cresceu no país, segundo levantamento da Embrapa Soja. O Brasil consegue produzir uma quantidade maior em uma área menor, e com bastante eficiência, devido ao incremento baseado em ciência e tecnologia (EMBRAPA), aplicação de técnicas de manejo avançadas (FREITAS, 2011).

Dentre os estados da federação o Rio Grande do Sul ocupou o terceiro lugar na safra de 2021/22 como um dos maiores produtores do país, 9.727,7 milhões de toneladas, em uma área plantada de 6.358,0 milhões de hectares, assim sendo uma produtividade de 1.530 kg/ha (EMBRAPA).

Até a década de 1980 os cultivares eram importados dos Estados Unidos e se adaptavam somente à região Sul do país, pois o plantio necessitava de um clima temperado e subtropical, onde as latitudes eram próximas de 30° (EMBRAPA).

Dentro desse contexto, uma produção de soja está diretamente associada a vários fatores ambientais. Neste estudo serão abordados o valor da produção, a temperatura, a umidade do solo, a altitude e a evapotranspiração.

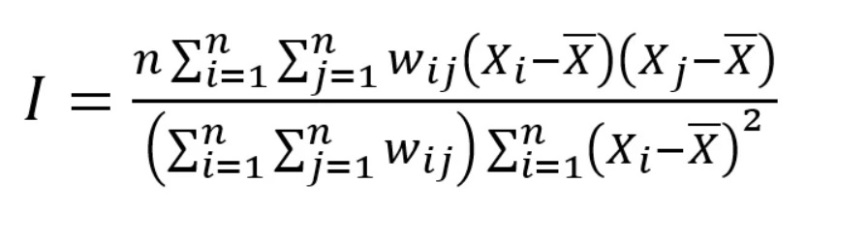
Nesse sentido, entender mais sobre a relação entre os fatores que influenciam na produtibilidade de soja pode conduzir discussões sobre melhores formas de dar continuidade ao aumento da produção e quais as regiões mais propícias para isso. O objetivo deste artigo é realizar uma análise espacial de quais fatores aumentam ou diminuem a produtibilidade de soja e como eles se relacionam entre si no estado do Rio Grande do Sul.

**Metodologia**

Para analisar a autocorrelação entre as variáveis foram usadas duas medidas estatísticas, índice de Moran global e local, e Regressão Geograficamente Ponderada (GWR) após definidas as variáveis.

A autocorrelação é a relação entre duas ou mais variáveis de modo que uma delas possa ser explicada pelas demais. O índice de Moran global é uma medida que avalia a correlação espacial em um conjunto de dados, variando seu valor de -1 a 1. Quando o índice está próximo de zero, indica que a variável em questão não apresenta padrões espaciais de aglomeração bem definidos. Valores positivos do índice indicam a presença de autocorrelação, ou seja, altos valores da variável tendem a estar próximos de outros altos valores, sugerindo aglomeração espacial. Por outro lado, valores negativos do índice indicam que altos valores da variável tendem a estar próximos de valores baixos, o que sugere uma dispersão espacial para o atributo analisado. A interpretação correta do Índice de Moran Global é essencial para compreender as estruturas espaciais presentes nos dados e pode ser aplicada em diversos campos (ACACIO, 2020).

Fórmula do Índice de Moran (YAMAMOTO, 2020):



Onde:

n = n° de pontos dados;

X = variável aleatória contínua;

wij = elementos da matriz W.

O índice de Moran Local, também conhecido como Indicador de Associação Espacial Local (LISA), é uma ferramenta estatística que fornece valores proporcionais à estatística global, permitindo a descrição do grau de semelhança ou diferença de cada evento em relação aos eventos mais próximos. Demonstra-se que a soma total dos valores LISA para todas as áreas é proporcional ao valor obtido para o índice de Moran Global (ACACIO, 2020).

Equação do Índice Moran Local:

Ii = Índice local de Moran;

Zi = diferença entre o valor do atributo no local e a média de todos os atributos;

Zj = diferença entre o valor do atributo no local e a média de todos os atributos;

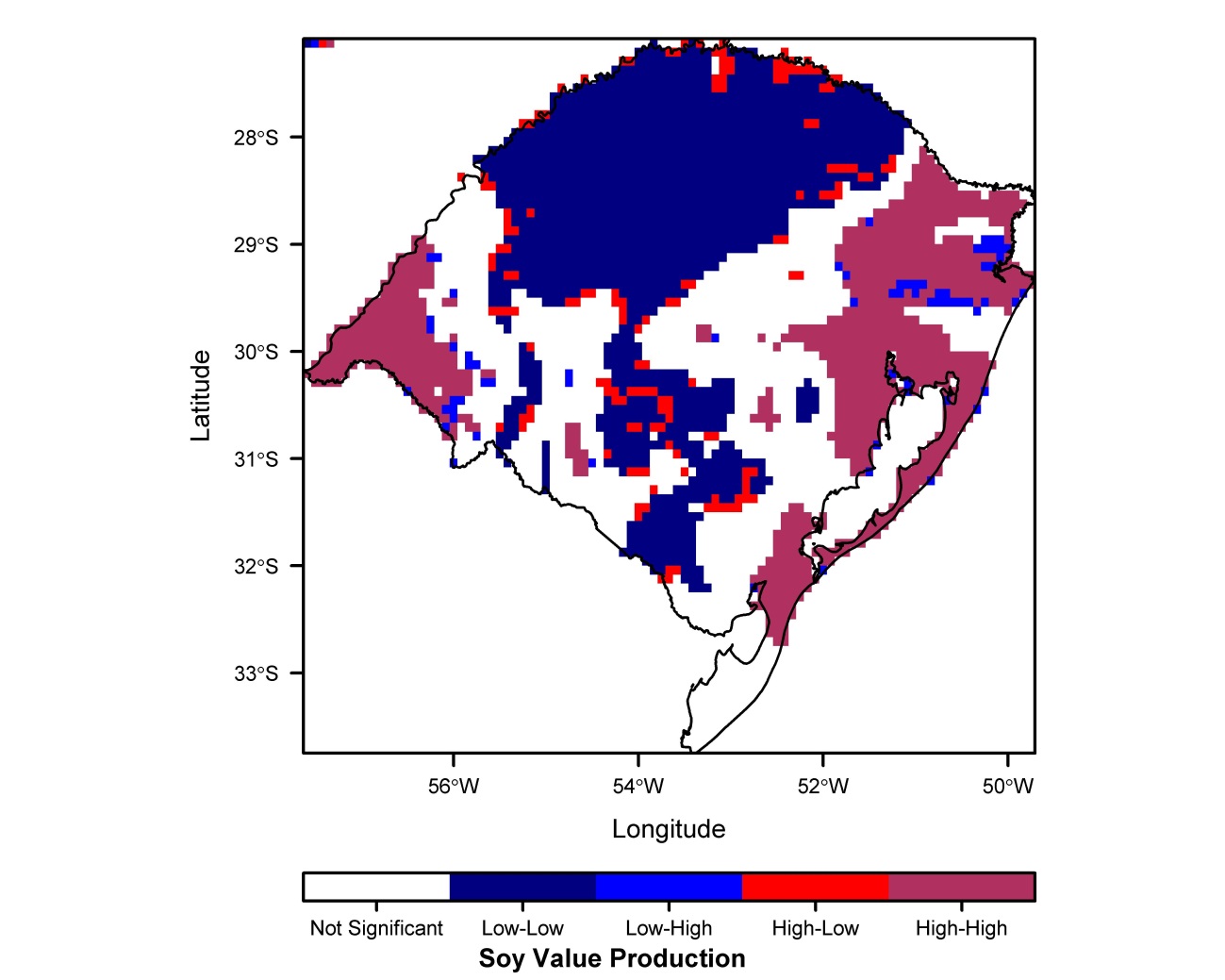
wij = pesos ou graus de conectividade atribuídos conforme a relação topológica entre as i e j.

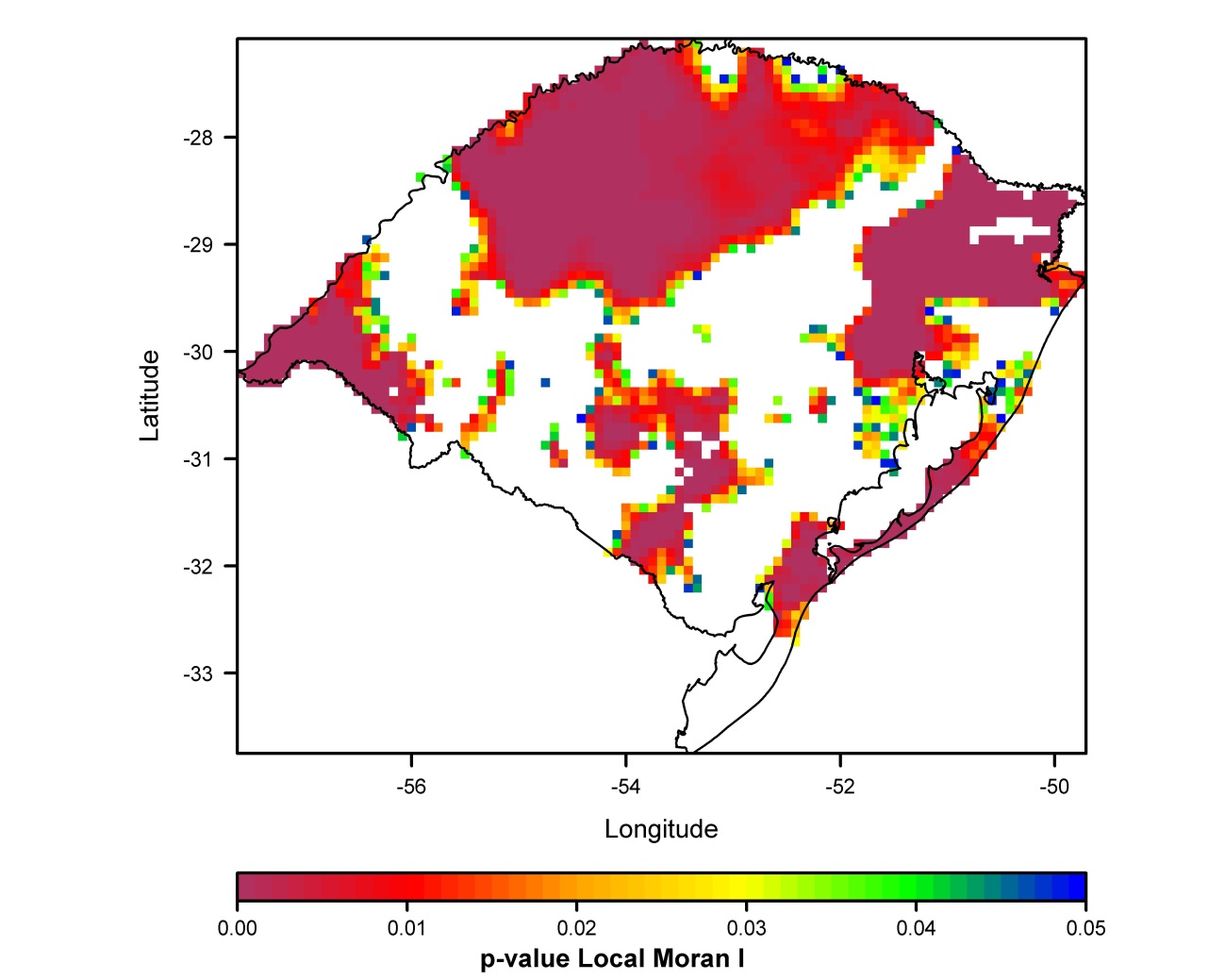
Uma outra maneira de se estudar o Índice de Moran é através do Diagrama de Dispersão de Moran. O diagrama de dispersão de Moran possui quatro quadrantes distintos, cada um representando diferentes padrões espaciais. O primeiro quadrante, chamado de baixo-alto (BA), mostra um grupo de localidades em que uma área com baixo valor da variável de interesse está cercada por localidades com alto valor dessa variável. O segundo quadrante, alto-alto (AA), representa um grupo de localidades que exibem altos valores da variável de interesse e estão cercadas por outras localidades com valores altos também. O terceiro quadrante, baixo-baixo (BB), refere-se a um grupo de associação espacial com baixos valores da variável de interesse, cercado por grupos também de baixos valores. Por fim, o quarto quadrante, chamado de alto-baixo (AB), mostra um agrupamento onde o valor da variável de interesse é alto, mas está circundado por localidades com valores baixos (PELLENZ; ALMEIDA; FREITAS, 2019).

Após a autocorrelação espacial, que é uma pré-condição para a aplicação da Regressão Geograficamente Ponderada (GWR) (ZHAO; YANG; ZHOU, 2010), a metodologia do modelo GWR tem por finalidade abordar a não estacionalidade espacial presente em determinada região geográfica, onde um modelo de regressão espacial global não seria suficiente para explicar adequadamente as relações entre algumas variáveis. Essa abordagem consiste em ajustar um modelo de regressão individual para cada ponto do conjunto de dados, levando em conta a distância das demais observações em relação a esse ponto específico. A ponderação das observações é realizada com base na proximidade, ou seja, quanto mais próximo um ponto está do ponto em estudo, maior será sua influência nos parâmetros estimados da regressão (ZILLI; DROUBI; HOCHHEIM, 2022).

**Resultados e discussõe**s

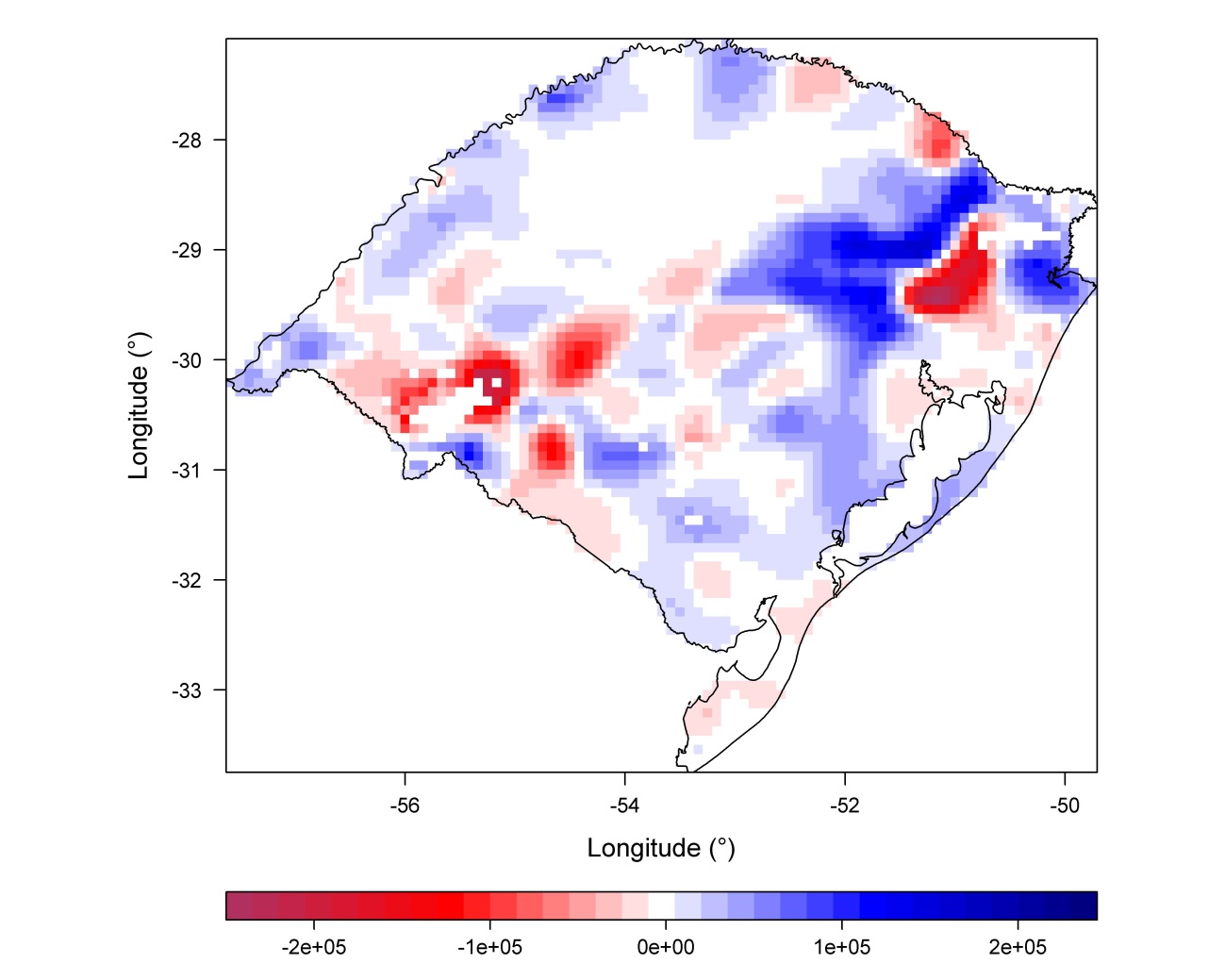
O software R foi usado para análise de dados estatísticos e geração de mapas. Ordinariamente, a partir do Teste de Moran I sob randomização, infere-se como resultado que a estatística de Moran I padronizada é igual a 141,14, indicando uma forte autocorrelação espacial. O valor do p-valor (<2.2 e -16) indica que essa autocorrelação é estatisticamente significativa. O resultado também indica que a hipótese alternativa é maior, sugerindo uma tendência de aglomeração espacial positiva nos dados. Nas estimativas amostrais, o valor da estatística de Moran I observada é de 0.6158210971, indicando a autocorrelação presente nos dados. O valor da expectativa teórica sob a hipótese nula de aleatoriedade espacial é -0.0002734482, o que significa que, sob a hipótese nula, espera-se que não haja autocorrelação espacial. A variância esperada sob a hipótese nula é de 0.0000209460, indicando a variabilidade esperada para a estatística de Moran I sob a hipótese nula.

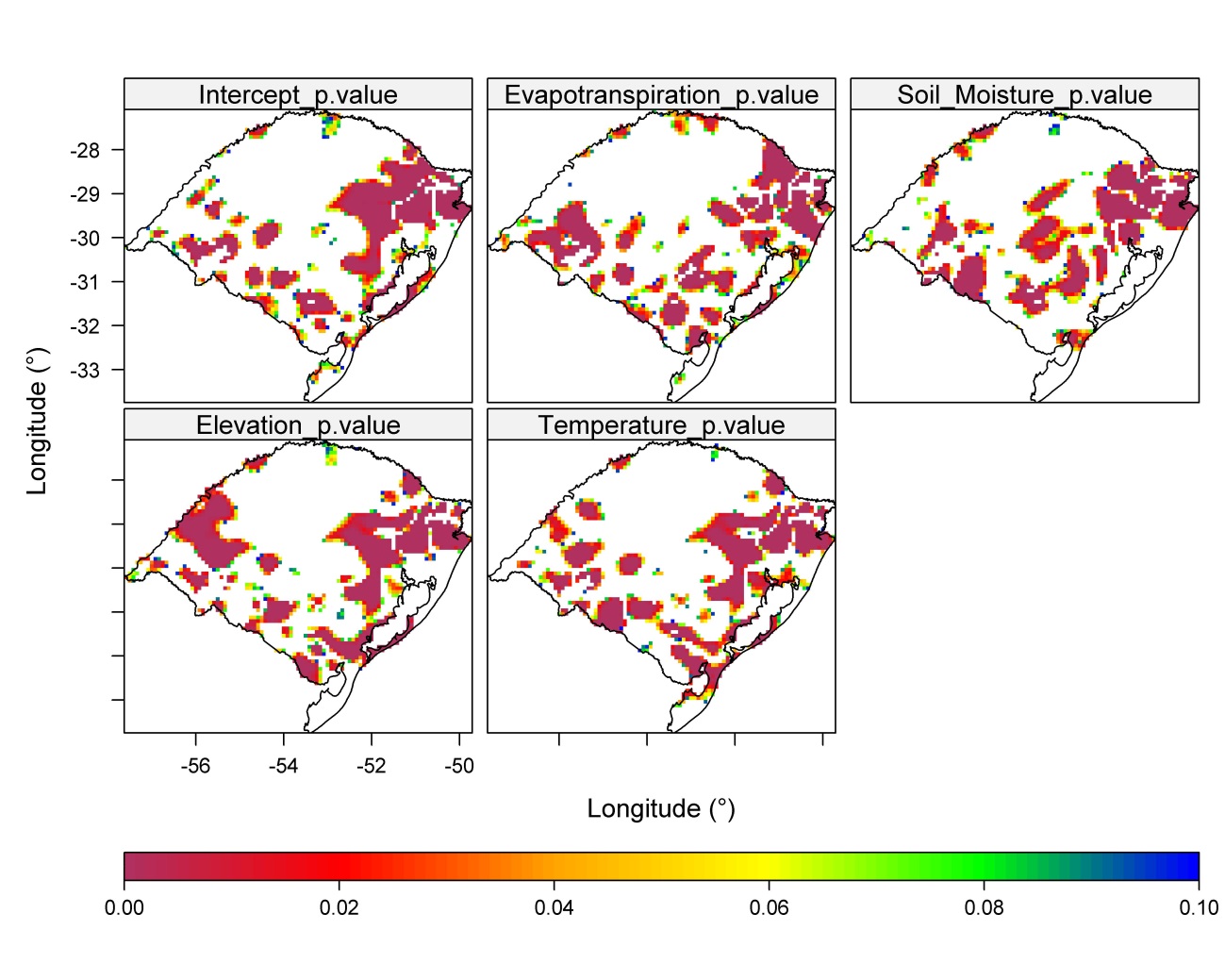




Na aplicação da Regressão Geograficamente Ponderada os dados foram organizados em uma matriz para a análise. A matriz possui inicialmente 9 colunas e cada uma preenchida com os valores das variáveis valor da produção, temperatura, umidade do solo, altitude e evapotranspiração. Após manipulação aplica-se a Regressão Linear Múltipla e o teste Durbin-Watson, utilizado para verificar a presença de autocorrelação nos resíduos de um modelo de regressão.

O valor calculado para o teste Durbin-Watson é 0.30455, indicando que a estatística do teste está próxima de zero. O valor-p é menor que 2.2e-16, o que significa que a autocorrelação nos resíduos é estatisticamente significativa (p < 0.05), rejeitando-se a hipótese nula de não autocorrelação. Neste caso, a variável resposta é o "Valor da Produção" e as variáveis explicativas são "Evapotranspiração", "Umidade do Solo", "Elevação" e "Temperatura". A hipótese alternativa sugere que a autocorrelação verdadeira nos resíduos é maior que zero. Ou seja, existe autocorrelação positiva nos resíduos, o que indica dependência serial entre eles.





**Referências bibliográficas**

ACACIO, Mariana Alves (2020). Análise espacial dos indicadores sociais e econômicos do estado do Tocantins. *Monografia Graduação – Universidade Federal do Tocantins – Curso de Ciências Econômicas, 2020.*

ATLAS (2022) – Soja - Atlas Socioeconômico do Rio Grande do Sul. Acesso em 13 de julho de 2023. <https://atlassocioeconomico.rs.gov.br/soja>.

EMBRAPA. História – Portal Embrapa. Acesso em 12 de julho de 2023, <https://www.embrapa.br/soja/cultivos/soja1/historia#:~:text=A%20soja%20que%20hoje%20cultivamos,do%20rio%20Yangtse%2C%20na%20China>.

EMBRAPA. Embrapa Soja. Soja em números (safra 2021/22). Acesso em 12 de julho de 2023, <https://www.embrapa.br/soja/cultivos/soja1/dados-economicos>

EMBRAPA (2023). Pesquisa, Desenvolvimento e Inovação. Brasil lidera e é referência no desenvolvimento de tecnologias sustentáveis para produção de soja. Acesso em 12 de julho de 2023, <https://www.embrapa.br/busca-de-noticias/-/noticia/81613580/brasil-e-referencia-no-desenvolvimento-de-tecnologias-sustentaveis-para-producao-de-soja>

FREITAS, Márcio de Campos Martins (2011). A cultura da soja no Brasil: o crescimento da produção brasileira e o surgimento de uma nova fronteira agrícola. ENCICLOPÉDIA BIOSFERA, Centro Científico Conhecer - Goiânia, vol.7, N.12; 2011 Pág.1

PELLENZ, Jéssica de Lima; ALMEIDA, Mariza; FREITAS, Claílton Ataídes (2019). Distribuição espacial do valor da produção da soja no Rio Grande do Sul: distintos retratos de 2000 a 2010. *Geosul, Florianópolis, v.34, n. 71. Dossiê Agronegócios do Brasil, p 86-110. Abril, 2019.*

SIQUEIRA, Tagore Villarim (2004). O ciclo da soja: desempenho da cultura da soja entre 1961 e 2003. *Biblioteca Digital BNDES.*

ZHAO, Na; YANG, Yonghui; ZHOU, Xinyao (2010). Application of geographically weighted regression in estimating the effect of climate and site conditions on vegetation distribution in Haihe Catchment, China. *Springer Science+Business Media B.V. 2010.*

ZILLI, Carlos Augusto; DROUBI, Luiz Fernando Palin; HOCHHEIM, Norberto (2022). Avaliação em massa de apartamentos com uso de regressão geograficamente ponderada. *Revista Valorem, v. 1, n. 1 (2022)*.

YAMAMOTO, JORGE KAZUO (2020). Como calcular o índice Moran – Autocorrelação espacial. Site Geo Krigagem. Acesso em 13 de julho de 2023, <https://geokrigagem.com.br/como-calcular-o-indice-moran-autocorrelacao-espacial/>