

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CÂMPUS MEDIANEIRA

RODOLFO SEFFRIN

**ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS APLICADA A
PRODUTIVIDADE DE MILHO NO ESTADO DO PARANÁ**

DISSERTAÇÃO

MEDIANEIRA
2017

RODOLFO SEFFRIN

**ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS APLICADA A
PRODUTIVIDADE DE MILHO NO ESTADO DO PARANÁ**

Dissertação de mestrado apresentado ao programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio – PPGTCA da Universidade Tecnológica Federal do Paraná – UTFPR – Campus Medianeira, como requisito parcial para a obtenção do título de “Mestre em Tecnologias Computacionais para o Agronegócio” - Área de concentração: Interdisciplinar.

Orientador: Prof. Dr. Everton Coimbra de Araújo
Coorientador: Prof. Dr. Claudio Leones Bazzi

MEDIANEIRA
2017

Dados Internacionais de Catalogação na Publicação

S453a SEFFRIN, Rodolfo

Análise exploratória de dados espaciais aplicada a produtividade de milho no estado do Paraná. / Rodolfo Seffrin – 2017.

96 f. : il. ; 30 cm.

Orientador: Everton Coimbra de Araujo

Coorientador: Claudio Leones Bazzi

Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócios. Medianeira, 2017.

Inclui bibliografias.

1. Análise espacial – Estatística. 2. Correlação (Estatística). 3. Computação – Dissertações. I. Araujo, Everton Coimbra de, orient. II. Bazzi, Claudio Leones, coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócios. III. Título.

CDD: 004

Biblioteca Câmpus Medianeira
Marci Lucia Nicodem Fischborn CRB: 9-1219



TERMO DE APROVAÇÃO

ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS APLICADA A PRODUTIVIDADE DE MILHO NO ESTADO DO PARANÁ

Por

RODOLFO SEFFRIN

Essa dissertação foi apresentada às dez horas, do dia vinte de abril de dois mil e dezessete, como requisito parcial para a obtenção do título de Mestre em Tecnologias Computacionais para o Agronegócio, Linha de Pesquisa Tecnologias Computacionais Aplicadas à Produção Agrícola, no Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio - PPGTCA, da Universidade Tecnológica Federal do Paraná. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho _____.

Prof. Dr. Everton Coimbra de Araújo (Orientador – PPGTCA)

Prof. Dr. Jose Airton Azevedo Dos Santos (Coorientador – PPGTCA)

Prof. Dr. Paulo Lopes de Menezes (Membro Interno – PPGTCA)

Profa. Dra. Denise Maria Grzegozewski (Membro Externo – UFGD – Dourados)

A via original com as assinaturas encontra-se na secretaria do programa.

AGRADECIMENTOS

Agradeço com um carinho especial meu orientador, Prof. Dr. Everton Coimbra de Araújo pela sua dedicação e pela orientação deste trabalho e, por meio dele, eu me reporto a toda a comunidade da Universidade Tecnológica Federal do Paraná (UTFPR) pelo apoio no desenvolvimento do trabalho.

A todos os colegas de trabalho gostaria de externar minha satisfação em poder conviver com eles durante essa etapa da minha vida.

Ao Prof. Dr. Claudio Leones Bazzi pela oportunidade e apoio na elaboração deste trabalho.

Agradeço ao SIMEPAR, IAPAR, INMET, IBGE, IPARDES pelo apoio recebido na execução deste projeto.

Gostaria de deixar registrado também, o meu reconhecimento à minha família, pois acredito que sem o apoio deles seria muito difícil vencer esse desafio.

RESUMO

SEFFRIN, Rodolfo. Análise exploratória de dados espaciais aplicada a produtividade de milho no estado do Paraná. Trabalho de Mestrado do Curso de Pós-graduação em Tecnologias Computacionais para o Agronegócio - PPGTCA. Universidade Tecnológica Federal do Paraná. Medianeira, 2017.

A cultura do milho é uma das atividades agrícolas mais importantes para a economia no Brasil e a utilização de modelos estatísticos pode auxiliar a tomada de decisão neste setor produtivo. O presente estudo visou identificar áreas com correlação e autocorrelação espacial para a produtividade de milho e suas variáveis preditoras (temperatura média, precipitação pluvial, radiação solar, potencialidade agrícola do solo e altitude), e também, verificar o modelo de regressão espacial mais adequado para a explicação da cultura. O estudo foi realizado utilizando dados de municípios do estado do Paraná referente a safras de verão dos anos agrícolas de 2011/2012, 2012/2013 e 2013/2014. Os softwares utilizados para a análise estatística e geração dos mapas temáticos foram o ArcMap 9.3 e GeoDa 1.6.7. A identificação da dependência espacial entre as variáveis foi realizada por meio do índice de Moran Global (Univariado e Bivariado) e o índice local de associação espacial (LISA), concluindo-se que para todos os anos e critérios de vizinhança utilizados, houve autocorrelação espacial significativa ao nível de 1% para todas as variáveis. Verificou-se ainda que a temperatura média, precipitação e altitude, estão correlacionadas significativamente ($P\text{-value} < 5\%$) com a produtividade do milho em todos os anos e critérios estudados. As variáveis: radiação solar e potencialidade agrícola do solo não apresentaram correlação significativa para alguns dos anos (2012/2013) e matrizes de vizinhança (contingência *queen* e vizinho mais próximo). Para determinar o modelo de regressão mais apropriado para a estimativa da produtividade de milho, foi adotado o diagnóstico estatístico do modelo de regressão OLS - *Ordinary Least Square*, que verifica se é necessário aplicar algum modelo de regressão espacial para explicação dos dados. Para todos os anos agrícolas foi recomendado a utilização do modelo de regressão espacial SAR – *Spatial Lag Model*, sendo que apenas para o ano agrícola 2013/2014 pode ser recomendado o modelo *Spatial Error Model* (CAR). A regressão espacial (SAR e CAR) adotada para a estimativa da produtividade de milho em diferentes anos, obteve melhores resultados quando comparada com os resultados da regressão que não incorpora a autocorrelação espacial dos dados (OLS). O coeficiente de determinação R^2 , os critérios de informação bayesiano (BIC) e o máximo valor do logaritmo da função verossimilhança (*Log-likelihood*), apresentou melhora significativa na estimação da produtividade do milho quando utilizado SAR e CAR.

PALAVRAS-CHAVE: Autocorrelação Espacial; Índice de Moran; Regressão Espacial, Modelo Espacial Auto Regressivo, Modelo do Erro Espacial

RESUMO EM LINGUA ESTRANGEIRA

SEFFRIN, Rodolfo. Exploratory analysis of spatial data applied to corn yield in the state of Paraná. Trabalho de Mestrado do Curso de Pós-graduação em Tecnologias Computacionais para o Agronegócio - PPGTCA. Universidade Tecnológica Federal do Paraná. Medianeira, 2017.

Corn cultivation is one of the most important agricultural activities for the Brazilian economy and the use of statistical models can help decision making in this productive sector. The present study aimed to identify areas with correlation and spatial autocorrelation for maize productivity and its predictive variables (mean temperature, rainfall, solar radiation, soil potential and altitude), and to verify the spatial regression model most appropriate for The explanation of culture. The study was carried out using data from municipalities in the state of Paraná referring to summer crops of the 2011/2012, 2012/2013 and 2013/2014 agricultural years. The software used for the statistical analysis and generation of thematic maps were ArcMap 9.3 and GeoDa 1.6.7. The identification of spatial dependence among variables was performed using the Moran Global Index (Univariate and Bivariate) and the local spatial association index (LISA). It was concluded that for all the years and neighborhood criteria used, there was spatial autocorrelation Significant at the 1% level for all variables. It was also verified that the average temperature, precipitation and altitude, are correlated significantly (P-value <5%) with corn yield in all years and criteria studied. The variables: solar radiation and soil agricultural potential did not present significant correlation for some of the years (2012/2013) and neighborhood matrices (queen contingency and nearest neighbor). To determine the most appropriate regression model for estimating maize productivity, the OLS - Ordinary Least Square regression model was used to verify if it is necessary to apply some spatial regression model to explain the data. For all agricultural years it was recommended to use the spatial regression model Spatial Lag Model (SLM), and only for the agricultural year 2013/2014 can the Spatial Error Model (SEM) be recommended. The spatial regression (SLM and SEM) adopted for estimating maize productivity in different years yielded better results when compared with regression results that did not incorporate spatial data autocorrelation (OLS). The coefficient of determination R^2 , the Bayesian information criteria (BIC) and the maximum log-likelihood value, showed a significant improvement in corn productivity estimation when using SLM and SEM.

KEY WORDS: Spatial Autocorrelation; Moran's Index; Spatial Regression, Spatial Lag Model, Spatial Error Model

LISTA DE QUADROS

Quadro 1- Softwares utilizados para o desenvolvimento do trabalho.....	66
--	----

LISTA DE FIGURAS

Figura 1 - Potencialidade dos Solos Agrícolas Brasileiro	21
Figura 2 - Matriz de proximidade espacial de primeira ordem, normalizada pelas linhas. (Câmara et al., 2005)	36
Figura 3 - Representação dos tipos de contiguidade entre áreas. (a) Contiguidade Queen (rainha), (b) Contiguidade Rook (torre) e (c) Contiguidade Bishop (bispo) (Araújo, 2012).....	36
Figura 4 - Matriz de Diagramas de Dispersão de Moran apresentado por Anselin et al. (2004).....	39
Figura 5 - Diagrama de espalhamento de Moran	41
Figura 6 - (Box Map) Mapa de Inclusão/Exclusão Social de São Paulo.....	43
Figura 7 - (LISA Map) Mapa de Inclusão/Exclusão Social de São Paulo	44
Figura 8 - (Moran Map) Mapa de Inclusão/Exclusão Social de São Paulo.....	45
Figura 9 - Determinação de áreas de influência pelo método de polígono de Thiessen para 12 locais.....	47
Figura 10 - Diagrama dos processos a serem realizados no trabalho	53
Figura 11 - Mapa do estado do Paraná - Brasil.....	54
Figura 12 - Mapa dos municípios do estado do Paraná com suas respectivas estações meteorológicas	57
Figura 13 - Polígonos de Thiessen calculados para as estações meteorológicas dos municípios do PR	59
Figura 14 - Histograma de conectividade para o critério contingência queen	60
Figura 15 - Histograma de conectividade para o critério da distância entre o centroide	61
Figura 16 - Decisão do Processo de Regressão Espacial (Anselin, 2005; Song, 2014)	64
Figura 17 - Mapa de Espalhamento de Moran para a produtividade de Milho	69
Figura 18 - Mapa de Espalhamento de Moran para as variáveis climáticas (Radiação Solar, Temperatura Média, Precipitação Pluvial)	70
Figura 19 - Mapa de Espalhamento de Moran para as variáveis altitude e potencialidade agrícola do solo	71
Figura 20 - Mapa de agrupamento para a produtividade de milho (Moran Map).....	72
Figura 21 - Mapa de agrupamento para os dados climáticos (Moran Map)	74

Figura 22 - Mapa de agrupamento para a altitude e potencialidade agrícola do solo (Moran Map).....	75
Figura 23 - Mapa da correlação espacial das variáveis climáticas com a produtividade de Milho (Moran MAP)	78
Figura 24 - Mapa da correlação espacial das variáveis altitude e potencialidade agrícola do solo com a produtividade de milho (Moran MAP)	79
Figura 25 - Mapa dos resíduos padronizados da regressão espacial gerada pelos modelos Spatial Error e Spatial Lag, considerando o método do desvio-padrão.	82

LISTA DE TABELAS

Tabela 1 - Índice de Moran global aplicado para as variáveis do estudo	68
Tabela 2 - Índice de Moran Global Bivariado aplicado para a produtividade de milho com as demais variáveis do estudo	76
Tabela 3 - Diagnóstico OLS para verificar o modelo adequado para aplicar a regressão conforme diagrama da Figura 1	80
Tabela 4 - Modelo de Regressão Espacial aplicado para a predição da produtividade de milho.....	81

SUMÁRIO

1 INTRODUÇÃO	14
1.1 OBJETIVO GERAL	15
1.2 OBJETIVO ESPECIFICOS	16
1.3 JUSTIFICATIVA	16
2 REFERENCIAL TEÓRICO	18
2.1 A CULTURA DO MILHO	18
2.2 CLASSIFICAÇÃO DO SOLO	19
2.3 ALTITUDE	21
2.4 DADOS AGROMETEOROLÓGICOS	22
2.4.1 Temperatura do Ar	23
2.4.2 Precipitação Pluvial	24
2.4.3 Radiação Solar	25
2.5 GEOPROCESSAMENTO	25
2.5.1 Sistema de Informações Geográficas (SIG)	26
2.6 ESTATÍSTICA ESPACIAL	28
2.6.1 Análise Espacial	28
2.6.2 Análise Exploratória de Dados Espaciais (AEDE)	29
2.6.3 Dependência Espacial	31
2.6.4 Estatística Espacial de Área	32
2.6.5 Análise de Variáveis Espaciais de Áreas	33
2.7 ANÁLISE INFERENCIAL	34
2.7.1 Autocorrelação Espacial	34
2.7.2 Matriz de Proximidade Espacial	35
2.7.3 Autocorrelação Espacial Global Univariada	36
2.7.4 Autocorrelação Espacial Global Multivariada	38
2.7.5 Autocorrelação Espacial local univariada	39
2.7.6 Autocorrelação Espacial Local Multivariada	40
2.7.7 Diagrama de Dispersão de Moran	41
2.7.8 Representação Gráfica: <i>Box Map</i> , <i>LISA Map</i> e <i>Moran Map</i>	42
2.8 ÁREAS DE INFLUÊNCIA	45
2.9 REGRESSÃO ESPACIAL	48
2.9.1 Modelo de Regressão SAR – <i>Spatial LAG Model</i>	49
2.9.2 Modelo de Regressão CAR – <i>Spatial Error Model</i>	49
2.10 SISTEMA GERENCIADORES DE BANCO DE DADOS ESPACIAIS	50
2.10.1 Banco de Dados PostgreSQL	51
2.10.2 Extensão Espacial Postgis	51

3 MATERIAIS E MÉTODOS	53
3.1 OBTENÇÃO DO DADOS	54
3.2 ORGANIZAÇÃO DOS DADOS	55
3.3 POLÍGONOS DE THIESSEN	58
3.4 MATRIZ DE PROXIMIDADE ESPACIAL	59
3.5 ANÁLISE DE DADOS DE ÁREA.....	61
3.6 MODELOS DE REGRESSÃO ESPACIAL	63
3.7 SOFTWARES UTILIZADOS.....	65
4. RESULTADOS E DISCUSSÕES.....	67
5. CONCLUSÃO	84

1 INTRODUÇÃO

Compreender fenômenos resultantes da distribuição espacial de dados no espaço é um grande mistério para diversas áreas do conhecimento. Vários estudos estão sendo realizados para se conseguir adquirir algum padrão que esteja associado à distribuição espacial, seja na saúde, em ambiente, em geologia ou na agricultura, entre outras. Portanto, essas pesquisas estão se tornando cada vez mais comuns em consequência da disponibilidade dos Sistemas de Informação Geográfica (SIG), que em alguns casos são encontrados com baixo custo e possuem uma interface amigável. Assim, esses sistemas fazem com que seja possível a visualização espacial de variáveis, como: população de indivíduos, índices de qualidade de vida e vendas de empresas em uma determinada região, sendo ilustradas por mapas, gráficos e tabelas. Com isso, é necessário se ter um banco de dados com informações de uma base geográfica (como um mapa de municípios), e um SIG para descobrir e possibilitar a visualização de algum padrão espacial nos dados analisados (DRUCK *et al.*, 2004).

Neste contexto, a estatística espacial situa-se em um cenário amplo, a análise espacial de dados. O objetivo é determinar a existência de algum padrão espacial, levando em consideração a localização geográfica do fenômeno em estudo. Assim, conceitos como autocorrelação espacial e dependência espacial são de suma importância. A dependência espacial pode ser obtida quando as ocorrências naturais ou sociais apresentam dependência em relação a distância (Druck *et al.*, 2004). Assim sendo, o objetivo é analisar como a dependência espacial varia em relação aos valores das amostras dos seus vizinhos (FOOK, 2005).

A técnica empregada a estatística espacial é diferente das demais técnicas que são utilizadas para análise descritiva, por levar em consideração a localização das informações no processo de coleta, descrição ou análise dos dados. É usado o termo autocorrelação espacial para poder diferenciar a correlação da estatística convencional. Por exemplo, a correlação é adquirida em meio a duas variáveis diferentes, sem qualquer referência a sua posição no espaço, já no caso de autocorrelação, utiliza-se no cálculo os valores de uma mesma variável em duas posições diferentes (ROCHA, 2004).

A análise de dados espaciais de área está vinculada a métodos que são usados quando a localização está ligada a áreas delimitadas por polígonos, um exemplo é quando se lida com eventos agregados por cidades, bairros ou setores

censitários, em que não se sabe a localização correta dos eventos no espaço, apenas de um único valor por área (DRUCK et al., 2004).

Portanto, o intuito do trabalho foi a aplicação da estatística espacial de área com o objetivo de investigar a correlação e autocorrelação espacial entre a produtividade de milho, classificação do solo, altitude e das variáveis agrometeorológicas: precipitação pluvial, temperatura média do ar e radiação solar global média, usando técnicas como o índice I de Moran Global e Local univariado e bivariado e os testes de significância. Para a predição da produtividade de milho e suas variáveis explicativas (classificação do solo, altitude, precipitação pluvial, temperatura média do ar e radiação solar global média) foi aplicado os modelos de regressão espacial SAR (*Spatial LAG Model*) e CAR (*Spatial Error Model*), a utilização desses modelos foram recomendadas pelo diagnóstico OLS (*Ordinary Least Square*) de Anselin (2005). O estudo foi baseado nos municípios do estado do Paraná, Brasil, com informações da safra dos anos de 2011/2012, 2012/2013 e 2013/2014.

Devido a estes dados, expostos anteriormente, utilizou-se técnicas de estatística espacial de área para verificar a existência da correlação espacial entre as variáveis: produtividade de milho ($t\ ha^{-1}$), classificação do solo, altitude (m) e as variáveis agrometeorológicas: precipitação pluvial [Prec] (mm), temperatura média do ar [TMed] ($^{\circ}C$), radiação solar global média [Rs] (KJ/m^2). Além da aplicação do diagnóstico OLS (*Ordinary Least Square*) para descobrir o modelo de regressão recomendado para a explicar a produtividade de milho.

1.1 OBJETIVO GERAL

Aplicar a estatística espacial de áreas para os municípios do estado do Paraná com o intuito de descobrir a correlação e autocorrelação espacial entre a produtividade de milho ($t\ ha^{-1}$), classificação do solo, altitude (m) e as variáveis agrometeorológicas: precipitação pluvial [Prec] (mm), temperatura média do ar [TMed] ($^{\circ}C$), radiação solar global média [Rs] (KJ/m^2), aplicando técnicas como o índice I de Moran Global e Local univariado e bivariado e os testes de significância. Verificar o melhor modelo de regressão (em meio ao diagnóstico OLS - ordinary least square) para a predição da produtividade de milho, sendo utilizado como variáveis explicativas: classificação do solo, altitude (m) e as variáveis agrometeorológicas: precipitação pluvial [Prec] (mm), temperatura média do ar [TMed] ($^{\circ}C$), radiação solar global média [Rs] (KJ/m^2)

1.2 OBJETIVO ESPECIFICOS

Os principais objetivos específicos são:

- a) Obter as informações da produtividade de milho, variáveis agrometeorológicas, altitude e classificação do solo no estado do Paraná;
- b) Aplicar o método dos polígonos de Thiessen para simular os valores das variáveis agrometeorológicas dos municípios que não realizam medições diárias dessas informações;
- c) Calcular o Índice de Moran global e local uni e bivariado para a produtividade de milho e variáveis agrometeorológicas com o objetivo de verificar a correlação e autocorrelação espacial;
- d) Aplicar o diagnóstico OLS (*Ordinary Least Square*) para descobrir qual o modelo de regressão mais adequado para explicação da produtividade de milho, podendo ser recomendado modelos que incorporem a autocorrelação espacial, como o SAR - *Spatial LAG Model* e CAR - *Spatial Error Model*, ou o modelo OLS que não utiliza a autocorrelação espacial na sua predição;

1.3 JUSTIFICATIVA

O presente trabalho tem o objetivo de demonstrar as técnicas de Análise Estatística Espacial de Áreas para a obtenção de respostas não alcançadas em meio a métodos convencionais de análise. Para uma resposta satisfatória serão utilizados dados oficiais pertencentes à produtividade de milho e fatores influentes, com o objetivo de oferecer respostas na definição de parâmetros e modelos, os quais podem ser utilizados para determinar políticas públicas de prevenção ou incentivos agrícolas.

Em relação aos aspectos de caráter inovador este trabalho pode contribuir de várias maneiras, como:

- Definir metodologias e procedimentos de aplicação de conceitos de estatística espacial de áreas para dados georreferenciados, referentes à produção agrícola e seus principais limitantes;
- Definição de modelos para que possam ser simuladas previsões de safra, em função de dados climatológicos e da potencialidade agrícola do solo;

- Determinar a correlação espacial entre a potencialidade agrícola do solo, produtividade média de milho e dados climáticos, visando dar suporte a aplicação de políticas públicas.

Portanto, com o término da pesquisa, será possível verificar a existência de correlação e autocorrelação espacial entre as variáveis: produtividade de milho, altitude, potencialidade agrícola do solo, temperatura média, radiação solar e precipitação. E em meio a isso, definir um padrão espacial para simular previsões de safras e subsidiar a definição de políticas públicas.

2 REFERENCIAL TEÓRICO

Neste capítulo é abordado o referencial teórico do presente trabalho.

2.1 A CULTURA DO MILHO

O Milho (*Zeamays L.*) é uma espécie que pertence à família Gramineae/Poaceae, com origem no teosinto, *Zeamays*, subespécie mexicana (*Zeamaysssp. mexicana* (Schradler) Iltis, há mais de 8000 anos. É cultivado em vários países (Estados Unidos, China, Brasil, França, Indonésia e África do Sul, dentre outros). Possui uma grande adaptabilidade, representado por vários genótipos, o seu cultivo é realizado desde a linha do Equador até o limite das terras temperadas, e também ao nível do mar até altitudes superiores a 3600 metros, portanto, encontra-se em diversos climas como: tropicais, subtropicais e temperados. A finalidade desta planta é para a utilização na alimentação humana e animal, por possuir elevadas qualidades nutricionais, sendo composto por quase todos os aminoácidos conhecidos (BARROS; CALADO, 2014).

A escolha da cultura do milho, como variável a ser analisada espacialmente, se deve pela sua importância nos cenários nacional e internacional. O Brasil é o terceiro maior produtor mundial de milho, ficando atrás apenas de Estados Unidos e China, os três juntos representam aproximadamente 66% da produção global. Nos anos de 2014 e 2015 o Brasil teve uma produção de 84,8 milhões de toneladas. O estado com maior produção de milho é o Mato Grosso com 20,77 milhões de toneladas ou 23% da produção total, em seguida vem o Paraná com 16 milhões de toneladas, correspondendo a 20% da produção no país. Estes dois estados representam 43,3% de toda a produção nacional deste cereal. O destino das exportações brasileiras está concentrado no Irã e Vietnã, os dois juntos compram mais de 36% do que é exportado (SEAB, 2015).

No Paraná, o milho produzido é utilizado na sua maioria para atender a produção de aves e suínos, sendo que nos anos de 2014/15 a produção ultrapassou 16 milhões de toneladas, tornando o estado o segundo maior produtor brasileiro da cultura. Sendo que no ano de 2014 exportou mais de três milhões de toneladas. Até outubro de 2015 as exportações totalizaram 2,47 milhões de toneladas, volume 1,7% superior ao mesmo período do ano anterior. Em 2014 o milho teve participação de 7,5% do valor bruto de produção paranaense, totalizando 5,27 bilhões de reais. No

grupo de grãos a cultura representa 22% do total, tendo a soja com 62% no grupo SEAB (2015).

De acordo com a EMBRAPA (2011), a importância do milho está relacionada a diversas formas de utilização, que vai da alimentação animal até a indústria de alta tecnologia. Para o uso em grão na alimentação animal a sua representação varia entre 70% a 90% da produção total do país. Embora a quantidade produzida tenha pouca representatividade para a alimentação humana, é um cereal de muita importância para a população de baixa renda. Por exemplo, no Nordeste, o milho é a fonte de energia para várias pessoas que vivem no semiárido (DUARTE et al, 2010).

2.2 CLASSIFICAÇÃO DO SOLO

Segundo Andreolli (2013), o solo tem uma importância fundamental para o desenvolvimento de várias atividades humanas, algumas das principais são: construção civil (fundações, telhas e tijolos, dentre outros), tratamento de resíduos (esgoto, resíduos sólidos), produção de alimentos (agropecuária), ornamentação (i.e.: produção de espécies vegetais para paisagismo.), silvicultura (produção de madeira para móveis, residências), além de outras não comuns no dia a dia.

A palavra solo tem inúmeros significados, mas normalmente a definição mais comum afirma que o solo é o meio natural para o desenvolvimento de plantas terrestres (USDA, 2012). Para a Embrapa (2009), seguindo uma denominação técnica, define o solo como um conjunto de corpos naturais, formados por fragmentos sólidos, líquidos e gasosos, tridimensionais, dinâmicos, constituído por materiais minerais e orgânicos que ocupam a maior parte do manto superficial das extensões continentais do planeta, possuem matéria viva e também podem ser vegetados na natureza onde ocorrem e podem, eventualmente, ter sido modificados por interferências antrópicas.

A Pedologia é a ciência responsável pelo estudo dos solos. Esta palavra tem origem grega: pedo(n) = solo, terra; logos = estudo de um assunto particular, no entanto, a etimologia da palavra pedologia corresponde ao estudo dos solos (ANDREOLLI, 2013). De acordo com o IBGE (2007), essa ciência teve origem na União Soviética, em meados de 1880, por Vasily Dokuchaiev, que é considerado o pai da pedologia. Lepsch (2002) a descreve como sendo aquela que se dedica a estudar os solos, considerando sua origem, morfologia, classificação e mapas, formulando

propostas para seu melhor uso, dentro dos preceitos da sustentabilidade (proteção ambiental).

Ainda com Lepsch (1987), o solo faz parte de um conjunto complexo de fatores de produção, sendo importante pelo seu papel de fornecer às plantas o suporte físico, água e nutrientes, que auxiliam no seu desenvolvimento. Portanto, conhecer as características inerentes de cada solo, os chamados fatores edáficos, é importante para se concluir a potencialidade agrícola do solo em determinadas regiões.

O território brasileiro é caracterizado por possuir uma grande variedade de tipos de solos, correspondendo também a intensidade das diferentes manifestações de formas e relevo, clima, material de origem, vegetação e organismos associados, os quais, por sua vez, direcionam para a existência de variados processos formadores de solos (IBAMA 2002).

Essa diversidade nacional, apresentando uma vasta variabilidade de solo, clima e relevo, reflete diretamente no potencial agrícola das terras, na diversificação das paisagens e também nos aspectos ligados ao uso do solo (IBAMA 2002). Com base nessas questões, o IBGE (2002), fornece um mapa temático com a classificação, por região, da potencialidade agrícola dos solos, conforme Figura 1.

De acordo com o IBGE (2002), a classificação pode ser dada pelos seguintes critérios para a prática da atividade agrícola:

- a) Boa: Praticamente sem limitações;
- b) Boa e Regular: Deficiência de nutrientes;
- c) Regular e Boa: Riscos de inundações, deficiência de nutrientes e drenagem;
- d) Regular: Deficiência de Nutrientes e Teores elevados de Alumínio;
- e) Regular e Restrita: Deficiência de Nutrientes, teores elevados de alumínio, pequena profundidade e fortes declives;
- f) Restrita: Fortes declives, susceptibilidade a erosão, teores elevados de alumínio, deficiência de nutrientes e drenagem
- g) Restrita e Desfavorável: Teores elevados de sódio, deficiência de drenagem e riscos de inundações;
- h) Áreas Desaconselháveis: Áreas não recomendadas para a prática agrícola, por possuírem limitações muito fortes de solos e/ou topografia.

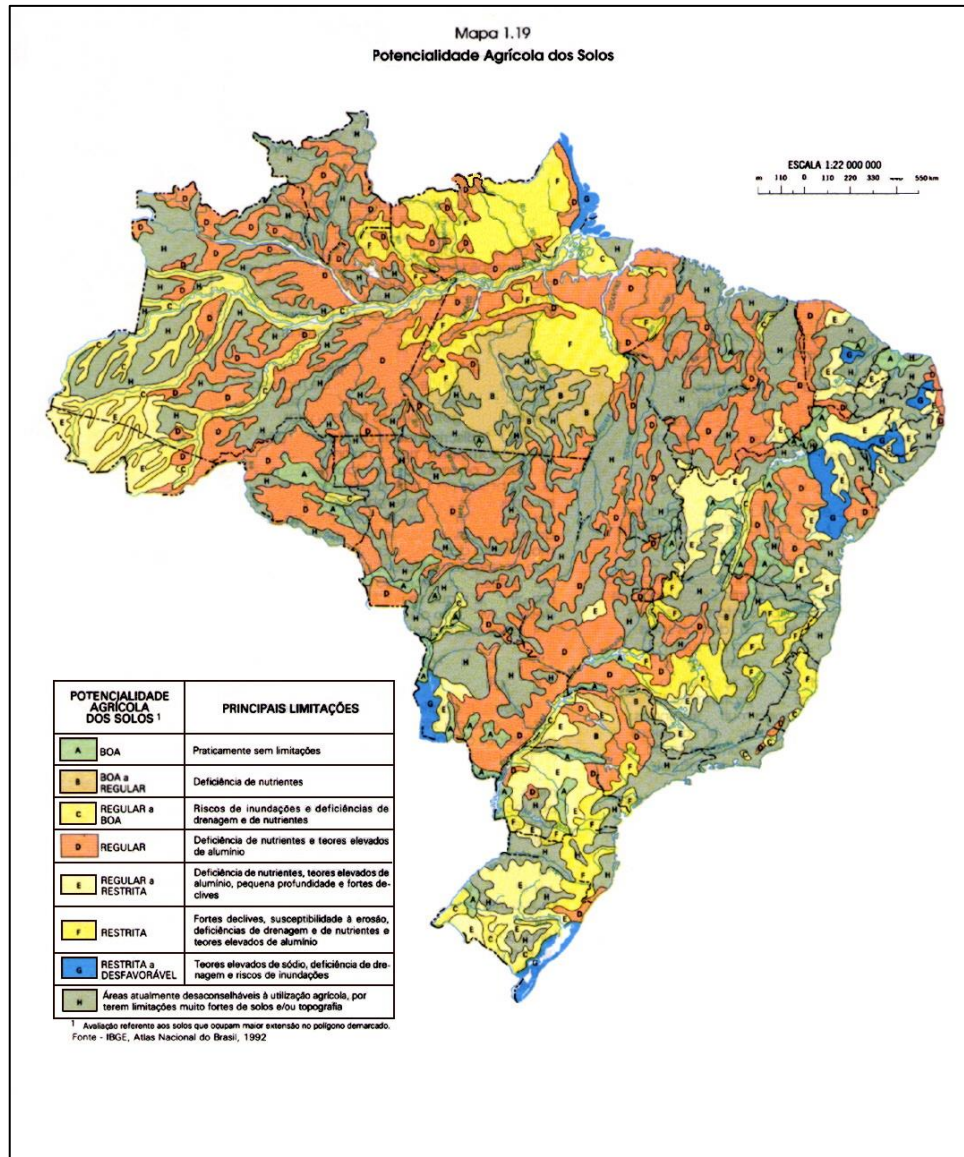


Figura 1 - Potencialidade dos Solos Agrícolas Brasileiro
FONTE: IBGE (2002)

Portanto, os municípios foram classificados de acordo com as regiões com o seu potencial produtivo representado pelo mapa da Figura 1.

2.3 ALTITUDE

De acordo com a EMBRAPA (2011), a altitude influencia diretamente na temperatura, tanto diurna como na noturna, tendo efeito também no processo fotossintético da planta, como a respiração. Para as condições brasileiras, o milho quando plantado em regiões com maiores altitudes apresenta um maior número de dias para atingir o pendoamento, aumentando o ciclo e tendo um melhor rendimento de grãos. Portanto, aumentando o enchimento dos grãos, em consequência aumentará a produtividade, em temperatura máximas menores e mais próximas da

ideal para o cultivo do milho, enquanto que no período noturno a temperatura será menor, reduzindo a taxa de respiração, o que ocasionará a redução do ponto de compensação (processo em que a fotossíntese e a respiração são idênticas), o que também auxilia no aumento da produtividade.

Uma análise realizada sobre diferentes cultivares nas regiões do Brasil, mostrou que os ensaios plantados nas áreas com altitude maior que 700 metros obtiveram um rendimento melhor (7.429 t ha⁻¹) e florescimento masculino de 65 dias, comparado com os ensaios cultivados em altitudes menores que 700 metros, que tiveram um rendimento de 6.473 t ha⁻¹ e florescimento masculino de 65 dias (EMBRAPA, 2011).

Em um estudo realizado por Durães (2002) verificou-se que a altitude influenciou no rendimento médio de milho. A análise dos dados de Ensaios Nacionais de Cultivares de Milho (33 Normal, 71 Precoce e 32 Superprecoce – 2000/2001) em diversos locais das regiões centro (cerca de 37 cultivares) e sul (7 cultivares) do Brasil, apresentaram um rendimento médio menor de grãos de milho em altitudes menos elevadas (<700 m), comparando com regiões com maiores altitudes, mostrando uma disparidade maior na região sul para todos os grupos de maturidade de milho.

O estado do Paraná possui um território extenso no qual a altitude varia muito entre os municípios, portanto, é de suma importância aplicar essa variável nas técnicas de estatística espacial de área, como pode-se verificar em estudos anteriores, pois a altitude tem uma forte influência na produtividade da cultura de milho.

2.4 DADOS AGROMETEOROLÓGICOS

Segundo Vianello e Alves (2001), os dados agrometeorológicos como: temperatura média do ar, precipitação pluvial e radiação solar, podem ser adquiridos em meio a leituras ou registros contínuos, diretamente dos aparelhos que realizam esse processo de captação dessas informações climáticas. As observações meteorológicas são efetuadas nas estações meteorológicas, que são lugares específicos e preparados para capturar essas informações.

De acordo com Sentelhas e Monteiro (2009), o conhecimento entre as relações das condições físicas do ambiente (solo e atmosfera) e as várias espécies que são cultivadas, permite a obtenção de informações precisas em torno da influência do

tempo/clima no crescimento, no desenvolvimento e também na produtividade dessas culturas. Esta influência específica das condições ambientais em determinadas culturas agrícola denomina-se agrometeorologia dos cultivos.

Para Hoogenboom (2000), as principais variáveis agrometeorológicas que influenciam no desenvolvimento, crescimento e produtividade das culturas são a chuva, temperatura do ar e a radiação solar. Foram escolhidas essas três variáveis agrometeorológicas por se tratarem dos elementos que mais influenciam na produtividade agrícola. Segundo Carmo Neto et al. (2011), a agricultura é a atividade econômica que sofre maior influência das condições climáticas, pois as variáveis climáticas interferem em todo o ciclo fenológico das plantas e influenciam a produtividade das culturas

O período de desenvolvimento do milho é limitado por três variáveis agrometeorológicas, que são a água, temperatura e radiação solar ou luminosidade. Portanto, a cultura do milho depende que os índices dos fatores climáticos, como a temperatura, precipitação pluviométrica e fotoperíodo, atinjam um nível considerado ótimo, para que o desenvolvimento em todo o ciclo de crescimento na produção se expresse ao máximo. Por isso, os principais elementos, pertinentes a este estudo, são: temperatura do ar, precipitação pluvial e radiação solar (CRUZ *et al*, 2006).

2.4.1 Temperatura do Ar

Silva (2008), afirma que, como a maioria dos gases, o ar não tem um bom desempenho na condução de calor, além de tardar muito para conseguir atingir o equilíbrio térmico com os demais corpos no qual se tem contato. A autora diz também que nas camadas de ar adjacentes ao solo é onde se encontram as maiores variações dos valores de temperatura do ar e a partir de certa altitude (correspondente à superfície de 850 hPa) é que se encontra um decréscimo um pouco mais regular. A autora ressalta ainda que a distribuição de temperatura no planeta tem interferência de vários fatores, como a latitude, a distribuição dos continentes e mares, as correntes marítimas, os ventos predominantes e pela ação das massas de ar.

A temperatura tem uma grande influência nos processos de desenvolvimento das plantas, como a respiração de manutenção, a transpiração, o repouso vegetativo, a duração das fases fenológicas das culturas, a indução ao florescimento, o conteúdo de óleo em grãos a taxa de germinação de sementes (PEREIRA *et al*, 2002, MAVI e TUPPER, 2004). Segundo Camargo (1974 e 1977), esses fatores citados

anteriormente, fazem com que a temperatura, juntamente com a chuva, seja a principal variável meteorológica nos zoneamentos agroclimáticos.

De acordo com Pereira e Cruz (2003), as etapas metabólicas do milho são influenciadas pela temperatura, o autor diz que a faixa de conforto para a cultura situa-se entre 20° C e 30° C, sendo que para a fase de emergência à floração é entre 26° C e 30° C.

2.4.2 Precipitação Pluvial

A precipitação é definida como um conjunto de partículas líquidas ou sólidas que caem das nuvens, alguns exemplos: chuva, chuvisco, granizo e neve. (INMET, 1999). No Brasil, a chuva é a precipitação com maior significância em termos de volume, é o elemento alimentador na etapa terrestre da fase hidrológica, e por consequência, é um dos fatores mais importantes nos processos de escoamento superficial direto, infiltração, evaporação, evapotranspiração, recarga de aquíferos e vazão dos rios (SILVA, 2008).

Barbosa (1983), afirma que devido à grande área foliar, o milho transpira intensamente, e por este motivo, é extremamente exigente em água. Magalhães e Durães (2006), complementam que o milho é uma planta cultivada em localidades onde a pluviosidade situa-se entre 300 mm a 5000 mm anuais, e que o consumo aproximado de água é em torno de 600 mm, durante o ciclo. Santos e Carlesso (1998), ressaltam que o milho é tolerante ao déficit hídrico na fase vegetativa, no entanto, os ciclos da emergência, florescimento e formação de grãos são os períodos que mais exigem água.

De acordo com Fancelli e Dourado Neto (2000), caso ocorra uma semana de déficit hídrico na fase de apendoamento, pode ocorrer a queda aproximada de 50% na produção de grãos, e de 25% a 32%, se o período seco acontecer após a fase de polinização.

Segundo Hoogenboom (2000), a chuva não influencia diretamente as etapas dos processos metabólicos das plantas, em consequência, age indiretamente no desenvolvimento e crescimento das culturas. No entanto, períodos com muita chuva ocasionam a redução da oxigenação dos solos, fazendo com que a atividade radicular seja reduzida, ou seja, a absorção de água e nutrientes.

2.4.3 Radiação Solar

A radiação solar é um dos principais parâmetros para a cultura do milho, sem a radiação a etapa fotossintética é inibida e a planta acaba sendo impedida de expressar o seu máximo potencial produtivo. A maior parte da matéria seca do milho, aproximadamente 90%, provém da fixação de CO² pelo processo fotossintético. O milho pertence ao grupo C4 que é altamente eficiente na utilização da luz. Por exemplo, se ocorrer uma redução entre 30% a 40% da intensidade luminosa, por um longo período, retarda a maturação dos grãos, podendo ocasionar uma queda na produção (CRUZ *et al*, 2010).

Ayoade (1986), afirma que a radiação solar é a energia que aciona o sistema agrícola, determinando as características térmicas do ambiente, especialmente as temperaturas do ar e do solo. Outro fator importante, é que a radiação solar determina a duração do dia, influenciando o fotoperiodismo (resposta dos vegetais a luminosidade).

De acordo com Taiz e Zeiger (2009), o milho, por pertencer ao grupo de plantas C4, praticamente não manifesta saturação por radiação solar, plantas C4 são vegetais cujo primeiro composto orgânico estável da fotossíntese é uma molécula de 4 carbonos que pode ser o malato ou o aspartato. Por esse motivo, segundo Fancelli e Dourado Neto (2000), incrementa rendimentos maiores, à medida que ocorre o aumento na intensidade de luz nos lugares onde se está sendo cultivada a cultura.

Para Pereira Filho e Cruz (2003), os fatores que podem influenciar na diminuição da incidência de luz está a nebulosidade e o número de plantas/ha. Fancelli e Dourado Neto (2000), consideram que a distribuição espacial das plantas também pode prejudicar na redução da incidência de luz. Sans e Guimarães (2011) afirmam que capacidade produtiva do milho safrinha está fortemente associada as limitações de radiação solar, além de temperatura e precipitação.

2.5 GEOPROCESSAMENTO

Segundo Druck *et al.* (2004), a coleta de informações utilizando a distribuição geográfica de recursos minerais, animais, propriedades e plantas, sempre exerceu uma importância nas atividades das sociedades organizadas. Portanto, no início dos estudos que envolviam informações de localizações geográficas, eram realizadas coletas apenas com documentos e mapas em papel. Em meio a isso, se tornava difícil

uma análise que pudesse utilizar diversos mapas e dados. Com o desenvolvimento da computação na segunda metade do século XX, tornou-se viável armazenar e realizar a representação dessas informações em um ambiente computacional, começando então o aparecimento do geoprocessamento.

Com os avanços tecnológicos da cartografia automatizada, dos sistemas de gerenciamento de banco de dados e também do processamento digital de imagens, junto com o desenvolvimento da computação, surgiu um conjunto distinto de ferramentas, com o objetivo de capturar automaticamente os dados relacionados com a superfície terrestre para o gerenciamento, análise e apresentação das informações geradas. A junção técnica e conceitual de todo o processo de desenvolvimento destas ferramentas resultou no desenvolvimento de uma grande variedade de métodos utilizados para o processamento de dados geográficos. Esse processamento é chamado, na literatura nacional, de Geoprocessamento. (CÂMARA; MEDEIROS, 1996) e de Geomática (OSTENSEN, 1995; CLEMENT *et al.*, 1997).

Para Silva (2003), geoprocessamento representa qualquer processamento de dados georreferenciados. O autor ressalta também que o geoprocessamento envolve técnicas e conceitos de cartografia, sensoriamento remoto, e também Sistema de Informações Geográficas (SIG).

2.5.1 Sistema de Informações Geográficas (SIG)

Dantas *et al.* (1996) divide as fases de evolução do SIG em três partes: manipulação e visualização de banco de dados (Fase 1), operações analíticas de dados não gráficos e estrutura organizacional (Fase 2) e análise espacial (Fase 3).

Na primeira fase, início da década de 50, o marco foi a necessidade de se poder armazenar, organizar, processar e visualizar dados. Em meio a isso, surgiu as primeiras versões de SIG, que foram baseados na manipulação e visualização de dados (MENESES, 2003). Segundo Teixeira *et al* (1995), a segunda fase se caracterizou pelo aumento da capacidade de memória e processamento dos computadores, tornando-se possíveis novas concepções e a popularização do SIG. Dantas *et al.* (1996), complementa que nesta nova perspectiva, as operações analíticas são destacadas pelos modelos matemáticos com informações numéricas não gráficos. E o crescimento da popularização do SIG fez com que surgisse a formação de recursos humanos para trabalhar no desenvolvimento desta nova

tecnologia, além da preocupação com a elaboração de uma estrutura organizacional para o uso do SIG (DANTAS, 1996).

A terceira fase surgiu no início da década de 80 e foi marcada pela diminuição dos recursos para a pesquisa científica, ao mesmo tempo houve o avanço do setor industrial e comercial do SIG. Neste período, todo o potencial do SIG é mais explorado, realizando a combinação de atributos não geográficos com as relações topológicas dos objetos geográficos, para então realizar operações de análise espaciais utilizando dados georeferenciados (DANTAS, 1996).

De acordo com Demers (1997) e Teixeira *et al.* (1995), existem diversas definições para SIG, mas eles destacam duas em específico. A primeira é o SIG como um *software*. Já a segunda, mais abrangente, considera que o SIG é um sistema que incorpora vários elementos, inclusive o software. Teixeira *et al.* (1995), define o SIG como:

“Conjunto de programas, equipamentos, metodologias, dados e pessoas (usuários), perfeitamente integrados, de forma a tornar possível a coleta, o armazenamento, o processamento e a análise de dados georeferenciados, bem como a produção de informação derivada de sua aplicação.”

Segundo Pina (1994), os SIGs são sistemas baseados em computador utilizados para realizar a armazenagem e manipulação das informações geográficas. Permitem reunir uma enorme variedade de informações convencionais de expressão espacial, estruturando-se adequadamente, de uma maneira a otimizar o tratamento integrado de seus três componentes no processamento de suas análises e aplicações gráficas: posição, topologia e atributos. O autor ressalta que o SIG é composto por um sistema gerenciador de banco de dados georeferenciados, sendo possível realizar operações de análises espaciais complexas em meio a uma rápida formação e alternância de cenários. O autor afirma também que se caracterizam por possuírem um potente instrumento de análise, ao contemplar uma ampla variedade de alternativas nas avaliações e simulações dos cenários projetados.

Camara *et al.* (1996), destacam a função de armazenagem, recuperação e análise de dados geográficos que são realizadas em sistemas SIG. Isso é possível pela integração entre os dados espaciais (ex. dados censitários, mapas cartográficos, imagens de satélite) em um único banco de dados, manipulada pela utilização de algoritmos que são capazes de gerar diversas respostas, além destas informações poderem ser consultas e visualizadas em formato gráfico (ex. mapas temáticos, gráficos).

2.6 ESTATÍSTICA ESPACIAL

2.6.1 Análise Espacial

No Brasil, as técnicas possíveis para realizar a automatização da análise espacial são chamadas de: Geomática, Geoprocessamento, Sistema de Informação Geográfica e mais recentemente como Geotecnologias (ROSA, 2011). No entanto, cada um desses termos apresentam uma definição diferente, mas abordam uma temática semelhante. E esses conceitos são apresentados nas obras de Tomlin (1990), Maguire *et al.* (1997), Camara *et al.* (1996), Burrough; McDonnell (1998), Xavier da Silva e Zaidan (2004), Tomlinson (2005) e Demers (2009).

De acordo com Rosa (2011), para realizar uma análise espacial, é necessário obter informações por meio de observações diretas da realidade, e em meio a isso, tentar descobrir algum padrão que esteja relacionado com a distribuição geográfica. Rosa (2011) complementa que essas informações podem ser representadas de várias maneiras, como: texto, número, gráficos ou imagens, podendo ou não estar ligada a uma localização geográfica. Rosa (2011), afirma que esses dados em formato de texto ou números são chamados de alfanuméricos, no qual podem estar associados a uma localização específica ou a um momento (no tempo) específico. As coletas das informações podem ser adquiridas de fontes primárias ou secundárias. Os dados devem ter três dimensões: temporal (tempo), temática (o quê) e espacial (localização geográfica).

Rosa (2011) diz ainda que a coleta dos dados está ligada diretamente com o resultado final do trabalho, por exemplo, a qualidade dos dados pode ser determinada pela sua precisão e também exatidão. Para Rosa (2011), a precisão dos dados corresponde ao nível de detalhe das informações, e a exatidão é o grau de proximidade com o valor real.

Bailey (1994) define a análise espacial como uma ferramenta que possibilita a manipulação dos dados espaciais de várias maneiras e em meio a isso extrair um conhecimento adicional como resposta, incluindo procedimentos básicos de consultas espaciais dentro de uma determinada área de interesse, podendo gerar mapas e sumários estatísticos dessa informação, incorporando também, funções como investigação de padrões e relacionamento dos dados na região de interesse, para então, conseguir um entendimento melhor do fenômeno e a possibilidade de se fazer previsões.

Para Câmara *et al.* (2002) a análise espacial é um conjunto de procedimentos encadeados, cuja finalidade é a escolha de um modelo inferencial que considere explicitamente o relacionamento espacial presente no fenômeno. Os procedimentos iniciais da análise incluem o conjunto de métodos genéricos de análise exploratória e a visualização dos dados, em geral por meio de mapas. Essas técnicas permitem descrever a distribuição das variáveis de estudo, identificar observações atípicas, não só em relação ao tipo de distribuição, mas também em relação aos vizinhos e buscar a existência de padrões na distribuição espacial. Por meio desses procedimentos é possível estabelecer hipóteses sobre as observações, de maneira a selecionar o modelo inferencial melhor suportado pelos dados.

2.6.2 Análise Exploratória de Dados Espaciais (AEDE)

As técnicas de estimação convencionais usadas nos modelos clássicos de econometria trabalham com deduções básicas de que o erro aleatório possui média zero, com a variância igual e que não possuem correlação entre si. As variáveis independentes ou explicativas podem ser fixas em amostras que se repetem, entre outras. Portanto, é incomum, para não dizer impossível, que um conjunto de elementos perto de um determinado fenômeno em estudo se comporte de acordo com os pressupostos do modelo clássico de regressão linear (RAMANATHAN, 1998; ALMEIDA, 2004).

Para Anselin (1996) e Almeida (2004), a econometria convencional não tem elementos teóricos que consigam explicar os fenômenos que existem no espaço, como os efeitos espaciais da especificação. Segundo os autores, para poder superar esta limitação da análise econométrica convencional, a técnica tem que conseguir trabalhar com os efeitos provocados pela autocorrelação espacial, heterogeneidade espacial e também com a estrutura espacial desses tipos de dados.

Almeida (2004), afirma que a AEDE é tida como pré-requisito ao estudo da econometria espacial, subcampo da econometria que lida com as complicações ocasionadas pela interação espacial (autocorrelação espacial) e pela estrutura espacial (heterogeneidade espacial) em modelos de regressão para dados na forma de *cross-section* e painel de dados. Portanto, mostra-se um requisito importante na identificação de observações discrepantes no espaço e possíveis padrões de associação espacial.

A Análise Exploratória de Dados Espaciais (AEDE) utiliza dados georeferenciados e é frequentemente usada para verificar a existência de padrões espaciais, por exemplo, a heterogeneidade espacial e a dependência espacial, no qual pode indicar a semelhança entre as regiões vizinhas. Este método considera a distribuição e o relacionamento dos dados no espaço. A AEDE, é importante nos estudos dos processos de difusão espacial, por indicar padrões de autocorrelação espacial (dependência espacial entre os objetos geográficos) (ANSELIN, 1994; HAINING, 1997; GOODCHILD *et al.*, 2000).

De acordo com Anselin (1988), o objetivo principal da análise exploratória é a questão espacial das informações em sentido a associação ou dependência espacial e na heterogeneidade espacial, conforme já descrito. A técnica tem como intuito verificar as distribuições que ocorrem no espaço, tentar descobrir algum padrão de associação espacial ou agrupamento, e ainda, propor diferentes formas de instabilidade espacial (não estacionariedade), além de descrever situações atípicas, como *outliers*.

Muitas pesquisas foram realizadas no Brasil utilizando a análise exploratória de dados espaciais, por exemplo, estudos envolvendo economia regional e urbana de Monastério e Avila (2004), Monastério, Salvo e Damé (2008) estudaram a estrutura espacial das aglomerações industriais e do crescimento econômico e as desigualdades regionais ocorridas no Rio Grande do Sul. Alves e Silveira (2008), Perobelli *et al.* (2007) e Gonçalves (2005) em Minas Gerais, utilizaram a análise espacial para avaliar o comportamento de variáveis econômicas no estado. Em São Paulo, Neves *et al.* (2000) usaram a mesma metodologia para estudar dados socioeconômicos, e no Nordeste, Silva e Resende (2006) e Gama e Strauch (2008), realizaram um estudo que abordou o crescimento econômico das cidades e também indicadores de sustentabilidade locais, respectivamente.

Araújo (2012), utilizou as técnicas de análise exploratória na região oeste do Paraná, aplicando a estatística espacial de área na produtividade de soja e variáveis agrometeorológicas. Reis e Castro (2008), no estado do Para, conseguiram identificar relações de dependência espacial em atividades do *cluster* potencial agropecuário-florestal. Esta abordagem espacial constatou efeitos de transbordamento entre regiões vizinhas que são importantes para elaboração de aglomerados produtivos no estado.

2.6.3 Dependência Espacial

De acordo com Câmara *et al* (2002), uma definição importante para o entendimento e também análise dos fenômenos espaciais é a dependência espacial. Esta afirmação é baseada na definição de Tobler (1970), referindo-se a Primeira Lei da Geografia, em que todos os objetos que estão no espaço estão relacionados, no entanto, os que estão mais próximos são mais semelhantes, e também na afirmação de Cressie (1992), de que a dependência espacial está presente em todas as direções, mas à medida que a distância entre os objetos aumenta a autocorrelação espacial diminui. Por consequência a maior parte das ocorrências, sejam naturais ou sociais, possuem uma relação que depende da distância. Câmara *et al.* (2002), cita como exemplo para essa contextualização o fato de que se for detectada poluição em certo local de um lago é bem provável que, nas regiões próximas, as amostras também estejam poluídas.

No caso mais específico de transportes, Miller (1999) argumenta que os modelos de regressão usados para a previsão de demanda por transportes usaram para o estudo variáveis explicativas, como: uso do solo e também fatores demográficos (domicílio). Essas variáveis são dados geográficos, então terão o efeito da autocorrelação espacial. O autor complementa que essa questão não é bem reconhecida na literatura de modelagem de transportes, apesar de gerar um problema potencial.

A expressão computacional do conceito de dependência espacial é a autocorrelação espacial (CÂMARA *et al.*, 2002), termo baseado no conceito estatístico de correlação, e isto é usado para medir a dependência espacial entre duas variáveis aleatórias. A preposição “auto” informa que o cálculo da correlação será realizado com a mesma variável aleatória, calculada em lugares diferentes no espaço. A autocorrelação é positiva se a ocorrência de um certo evento tem influência para que outro semelhante ocorra em uma área próxima, o que ocasiona um espalhamento aglomerado de eventos. Caso a ocorrência de um mesmo evento impeça ou dificulte o acontecimento de outros eventos parecidos ao seu redor, diz-se que a autocorrelação é negativa, que nesta situação é uma distribuição equidistante dos eventos (TEIXEIRA, 2003).

Outro ponto importante na análise espacial é a caracterização da dependência espacial. Para Anselin (1992), a maneira como os dados estão distribuídos no espaço

é um fator indicador da interação espacial, que é mostrado nas medidas de associação ou autocorrelação espacial. Portanto, a análise exploratória indica como os valores das áreas do estudo estão no espaço, e em meio a isso estima a relevância da dependência espacial.

Anselin (1994) e Moran (1947), concordam que a dependência espacial pode ser vista em termos globais usando o Índice de Moran Global, e também para análises locais utilizando o LISA – *Local Indicators of Spatial Association*. Caso seja uma análise global, o resultado considerado da associação espacial dos dados é a região como um todo, senão, se for uma análise local, são fornecidas estatísticas que indicam a dependência do local considerando a situação média de uma vizinhança definida *a priori*.

2.6.4 Estatística Espacial de Área

De acordo com Miller (1999), o uso de métodos de análise de dados espaciais, onde a localização está associada a áreas que estão delimitadas por polígonos, acontece com bastante frequência quando se trabalha com eventos agregados por cidades, bairros, setores censitários ou zonas de tráfego, em que não se tem o valor correto do local onde os eventos ocorreram, mas sim de um valor por área. Portanto, nessas situações, procura-se saber padrões espaciais de distribuição, aglomerados e sinais de dependência espacial, no entanto, deve se ter o conhecimento de alguns problemas que podem acontecer, como a descontinuidade das fronteiras, pois, na maioria das vezes, valores que estão próximos da fronteira, em áreas contíguas, tendem a ser parecidos, embora cada região seja representada por sua média, o que pode distorcer a análises nas áreas de fronteira, apresentando um erro de valor que não seria o real. O outro problema é a de Unidade de área Modificável (*Modifiable Areal Unit Problem - MAUP*), em que a primeira evidência foi abordada por Gehlke e Biehl (1934), depois também por Robinson (1950) e que passou a ser estudado cada vez mais por pesquisadores como Cressie (1995, 1996) e Miller (1999). O que se observa é que o grau de agregação das subáreas, e também a escala do mapa, pode proporcionar erros na etapa de análise.

Segundo Assunção (2001), dados de área englobam elementos que juntam o mapa geográfico a um banco de dados. Este mapa geográfico é separado em áreas, onde cada uma dessas regiões possui uma ou mais variáveis aleatórias que correspondem ao valor de toda a área e não para uma local em específico. Por

exemplo, nesta situação não se tem a informação do local exato da ocorrência, mas sim o valor agregado por área. No presente trabalho foi trabalhado com análise espacial de área por não se saber o local exato da ocorrência de cada evento, que neste caso é a produtividade de milho e as variáveis agrometeorológicas.

O intuito da análise espacial de área é descobrir a existência de um padrão espacial nos valores que estão sendo estudados. A maneira mais comum de apresentar essas informações por área é em meio a mapas coloridos com o padrão espacial do fenômeno. Por exemplo, quando este padrão espacial é analisado, é necessário verificar se o padrão apresenta algum agrupamento ou se é aleatório, ou também se a distribuição está associada a causas mensuráveis (KREMPI, 2004).

Portanto, para as análises de dados de área é necessário calcular a associação espacial das variáveis em estudo, e isso pode ser realizado por índices que calculam esta autocorrelação (Índice de Moran), gráfico de espalhamento (*Box Map*) e mapas (Distribuição espacial). Esses índices identificam os agrupamentos de objetos cujo os elementos possuem valores altos e baixos, área de transição e também situações atípicas (KREMPI, 2004).

2.6.5 Análise de Variáveis Espaciais de Áreas

Segundo Carvalho (1997), uma das técnicas mais usadas para estudar fenômenos relacionados a eventos de área é a análise da autocorrelação espacial. Essa técnica faz com que seja identificada a estrutura de correlação espacial que melhor exemplifica o padrão de distribuição dos dados. A ideia é estimar a dependência espacial entre as regiões, evidenciando como os valores estão se comportando no espaço.

De acordo com Câmara *et al.* (2002), os indicadores globais de autocorrelação espacial, como o Índice de Moran, fornecem um valor único para a caracterização da associação espacial de toda a área em estudo. Portanto, quando está se trabalhando com muitas áreas, é provável que possa acontecer diferentes situações de associação espacial e que encontre locais onde a dependência espacial é ainda maior.

Portanto, as “análises locais” ou “modelagens locais”, tentam encontrar a existência de diferenças espaciais ao invés de assumir que estas não existem. Estas análises separam as estatísticas globais de acordo com a sua composição local, focando especificamente em padrões locais e não em globais (FOTHERINGHAM *et al.*, 2000).

Entre as técnicas univariadas utilizadas nas análises locais existem as abordagens gráficas e também as focadas para o desenvolvimento formal de estatísticas univariadas locais. Entre as abordagens gráficas, objetiva-se em encontrar exceções regionais para as tendências gerais na distribuição dos dados e nas relações entre os elementos. Para isso, utiliza-se histogramas, gráficos de dispersão e gráficos em três dimensões (MELO; HEPP, 2008).

2.7 ANÁLISE INFERENCIAL

Esta etapa tem o objetivo de construir modelos que melhor representem a distribuição espacial dos dados de área, com o intuito de gerar resultados aplicáveis à produtividade de milho, classificação do solo e fatores agrometeorológicos. As técnicas descritas podem ser aplicadas a estatística espacial de área.

2.7.1 Autocorrelação Espacial

Krempi (2004) disserta que, no contexto da estatística clássica, a definição de correlação é em respeito ao relacionamento de duas variáveis. Caso ocorra o aumento da dependência à medida que as áreas estão mais próximas, tem-se uma correlação positiva. Se acontecer o inverso, é uma correlação negativa. Ainda segundo o autor, se não constatar relação quantificável, pode-se concluir que as variáveis do estudo são não-correlatas, ou independente, nesta situação é uma distribuição aleatória. Portanto, a informação que é buscada no cálculo da autocorrelação espacial é de quanto o valor de uma variável em uma região é semelhante com a do vizinho mais próximo, e ao mesmo tempo, o quão diferente do vizinho mais distante.

Portanto a função de autocorrelação espacial é calcular o nível de dependência para os valores situados na área do estudo, ou seja, estimar quanto o valor de uma variável em uma determinada região é dependente dos valores dessa mesma variável nas áreas próximas (DRUCK *et al*, 2004; BRASIL/MS, 2007).

Enquanto o objetivo da correlação é comparar valores de duas variáveis aleatórias distintas, o da autocorrelação é a correlação realizada com a mesma variável aleatória calculada em locais diferentes no espaço (BRASIL/MS, 2007).

2.7.2 Matriz de Proximidade Espacial

Uma técnica importante para a análise espacial de área é a construção de uma matriz de vizinhança, que também pode ser chamada de matriz de distância, matriz de conectividade ou matriz de proximidade. Esta matriz de vizinhança aponta a relação espacial de cada área do estudo, podendo ser ligada pela lista de vizinhos de cada polígono, pela distância entre eles, e também pela conectividade ponderada pelo comprimento da fronteira comum (ARAÚJO *et al.*, 2013).

Segundo Anselin (1992), qualquer variável pode ser espacialmente defasada para qualquer unidade regional. A maneira mais comum para quantificar a estrutura de dependência espacial é utilizando a matriz de peso espacial.

Por exemplo, para calcular um conjunto de n áreas $\{A_1, \dots, A_n\}$, pode-se construir a matriz $W^{(1)}(n \times n)$, em que cada um dos elementos W_{ij} é uma medida de proximidade entre A_i e A_j . Então, esta medida é calculada seguindo os critérios:

- $w_{ij} = 1$, quando o centroide de A_i está a uma determinada distância de A_j ; senão $w_{ij} = 0$;
- $w_{ij} = 1$, quando A_i compartilha um lado comum com A_j , senão $w_{ij} = 0$;
- $w_{ij} = l_{ij}/l_i$, onde l_{ij} é o comprimento da fronteira entre A_i e A_j e l_i é o perímetro de A_i ;

Como a matriz de vizinhança é usada para cálculos de indicadores quando está na fase de análise exploratória, é importante realizar a normalização das suas linhas, fazendo com que a soma dos pesos das linhas seja igual a 1. Isto simplifica os cálculos que são feitos pelos índices de autocorrelação espacial (ARAÚJO *et al.*, 2013). A Figura 2 mostra um exemplo de matriz de vizinhança onde os valores foram normalizados:

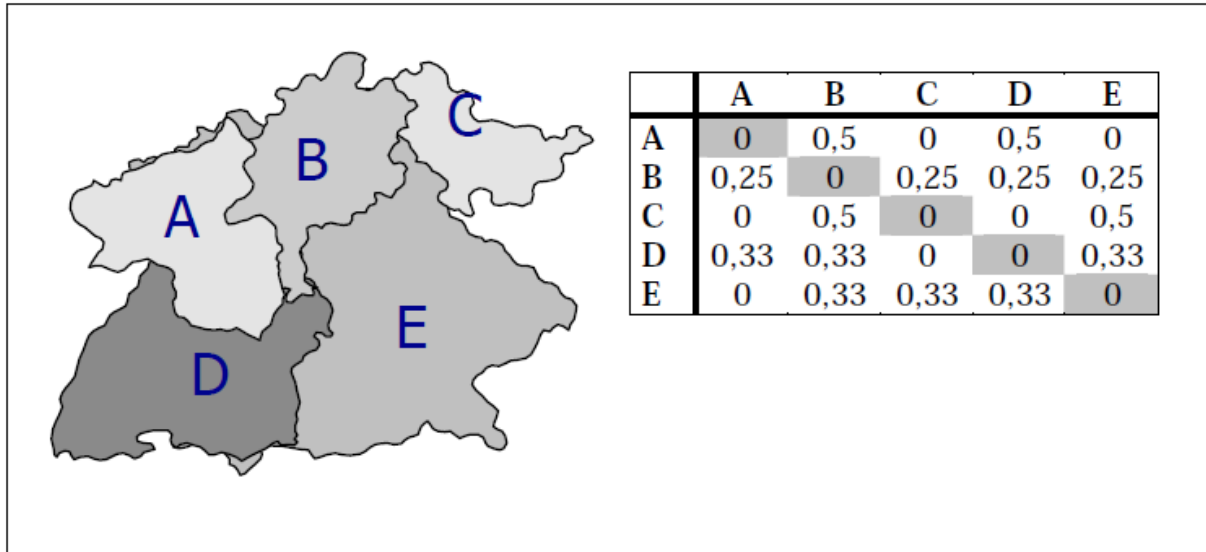


Figura 2 - Matriz de proximidade espacial de primeira ordem, normalizada pelas linhas. (Câmara *et al.*, 2005)

É de suma importância convencionar as formas de vizinhança nas matrizes de vizinhança que consideram a contiguidade. Sendo assim, as regras seguem os movimentos de peças que existem no jogo de xadrez, tais como a rainha (*Queen*), a torre (*Rook*) e o bispo (*Bishop*) (CÂMARA *et al.*, 2002). A Figura 3 ilustra um exemplo dessas regras para identificar os vizinhos na área J.

A	B	C	D
E	F	G	H
I	J	K	L
M	N	O	P

(a)

A	B	C	D
E	F	G	H
I	J	K	L
M	N	O	P

(b)

A	B	C	D
E	F	G	H
I	J	K	L
M	N	O	P

(c)

Figura 3 - Representação dos tipos de contiguidade entre áreas. (a) Contiguidade *Queen* (rainha), (b) Contiguidade *Rook* (torre) e (c) Contiguidade *Bishop* (bispo) (Araújo, 2012)

A Figura 3a mostra a regra *Queen*, onde todas as áreas que têm intersecção não-nula com a área J serão vizinhas de J. Já na Figura 3b detalha-se a regra *Rook*, considera como vizinho apenas as áreas com lado em comum da área J. E por último, a Figura 3c que demonstra a regra *Bishop*, que usa como critério de vizinhança as áreas que se localizam nas diagonais da região J.

2.7.3 Autocorrelação Espacial Global Univariada

De acordo com Krempi (2004), a autocorrelação espacial pode ter valor positivo ou negativo. Caso ocorra um nível elevado de autocorrelação espacial positiva, então os valores analisados em uma determinada localidade tenderão a ser parecidos aos

valores em regiões vizinhas. Se acontecer o contrário, ou seja, autocorrelação negativa, os valores baixos estão próximos de valores altos, o que significa que a média ponderada é maior para os vizinhos do que o valor analisado no próprio local, por exemplo, a presença de um fenômeno dificulta o acontecimento dele em locais próximos.

Segundo Anselin (1995), o índice global de Moran (I), é uma das formas para calcular a autocorrelação espacial. Esta estatística varia entre -1 e 1, fornecendo então uma medida global de associação linear (espacial) entre os valores z_t no tempo t e a média ponderada dos valores da vizinhança, ou Lags espaciais wz_t . Valores perto de zero indicam que não existe autocorrelação espacial significativa, e valores próximos de um, indica que existe dependência espacial entre a variável nas áreas vizinhas. Este índice é uma das maneiras para medir a similaridade entre as regiões e pode ser descrita conforme equação 1:

$$I_t = \left(\frac{N}{S_0} \right) \left(\frac{z_t' w z_t}{z_t' z_t} \right) \quad \text{Eq. (1)}$$

Onde:

- $t = \{ 1, 2 \dots n \}$;
- z_t = Vetor de n observações para o ano t na forma de desvio em relação à média;
- w = Matriz de peso espacial onde os elementos w_{ii} são os elementos que na diagonal principal são iguais a zero, enquanto os elementos w_{ij} indicam a forma como a região i está especialmente conectada com a região j ;
- S_0 = É um escalar igual à soma de todos os elementos de w ;

Sendo o valor esperado:

$$E(I) = - \frac{1}{n-1} \quad \text{Eq. (2)}$$

Quando a matriz de pesos espaciais é normalizada na linha, ou seja, quando a soma dos elementos de cada linha for igual a 1 a Equação (1) poderá ser reescrita como se segue:

$$I_t = \left(\frac{z_t' W z_t}{z_t' z_t} \right) \quad \text{Eq. (3)}$$

Após calcular a autocorrelação espacial das variáveis observadas é necessário verificar a significância do resultado. Para verificar se a correlação obtida é significativa, duas abordagens são possíveis, a primeira é o teste de pseudo-significância, que neste caso, são calculadas várias permutações dos valores das variáveis da área do estudo, em que cada uma dessas permutações gera um novo arranjo espacial, pois os valores são distribuídos novamente entre as regiões. Como apenas um desses arranjos correspondem a ocasião observada, é criada uma situação empírica de I . Se o valor de I for o extremo da distribuição simulada, trata-se de um evento com significância. A segunda abordagem é o teste de distribuição aproximada, que utiliza um número suficiente de sub-regiões, e pressupõe que as variáveis aleatórias associadas a cada local do atributo são independentes e normalmente distribuídas, então, assume-se que o índice I possui distribuição aproximadamente normal (KREMPI, 2004).

2.7.4 Autocorrelação Espacial Global Multivariada

A estatística I de Moran pode ser usada para calcular o grau de correlação entre duas variáveis diferentes. A ideia principal é descobrir se os valores de uma determinada variável em uma área possuem alguma relação com outra variável em áreas vizinhas (ALMEIDA, 2007a).

O coeficiente do I de Moran multivariado, com a matriz W normalizada na linha, é explicado com a seguinte equação:

$$I_{k,l} = \left(\frac{z_k^l W z_l}{n} \right) \quad \text{Eq. (4)}$$

Com n representando o número de observações e W a matriz de pesos espaciais. A soma dos quadrados no denominador é constante e igual a n independente de Z^k ou de Z^l (RIGOTTI e VASCONCELLOS, 2005). O diagrama de dispersão de Moran multivariado segue o mesmo padrão de análise do diagrama de dispersão de Moran univariado, só que no primeiro o eixo das ordenadas é representado por w_x , que é a variável x defasada para os vizinhos confrontada com a variável y no eixo das abscissas (ALMEIDA, 2007a).

Segundo Anselin *et al.* 2004, quando se utiliza uma análise multivariada, com mais de duas variáveis em estudo, usa-se a Matriz de Diagramas de Dispersão de Moran (*Moran's Scatterplot Matrix*). A representação desta matriz pode ser vista na Figura 4, onde os eixos inferiores são as variáveis espacialmente defasadas (com os *lags* espaciais aplicados as variáveis normalizadas). O objetivo desta ferramenta é analisar o padrão espacial de cada variável com ela mesma, e também a defasagem espacial com as outras variáveis.

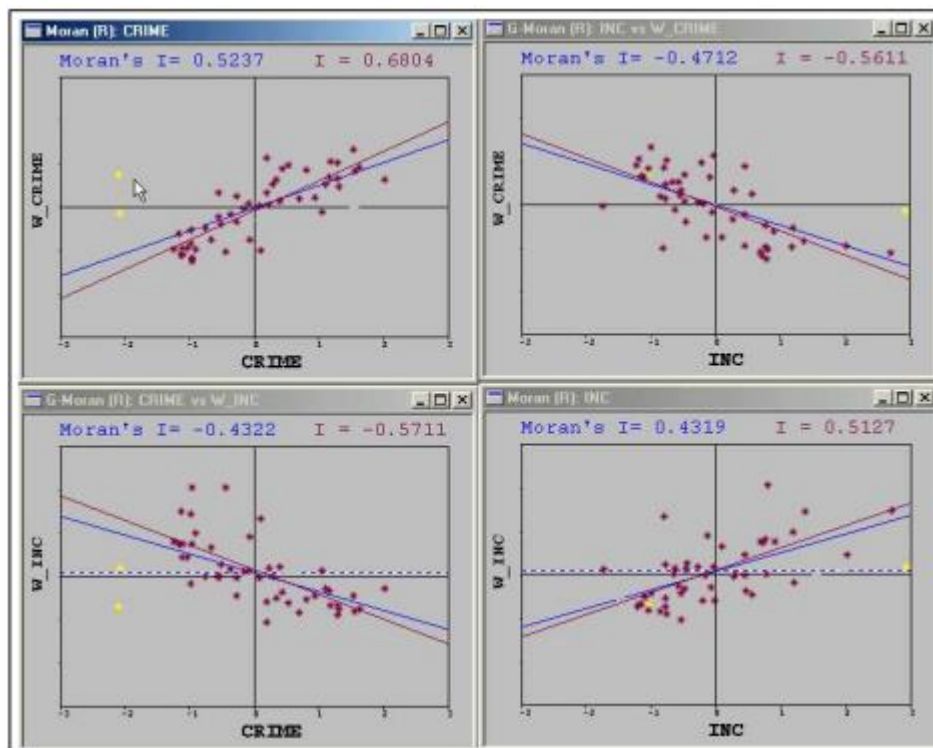


Figura 4 - Matriz de Diagramas de Dispersão de Moran apresentado por Anselin *et al.* (2004)

2.7.5 Autocorrelação Espacial local univariada

Os valores globais podem ocultar os padrões locais de autocorrelação espacial, sendo necessário analisar com maior detalhamento a dependência espacial existente. Por este motivo que existem os indicadores locais que permitem um maior detalhamento para cada região, tornando possível a identificação de diferentes padrões espaciais, como agrupamentos com valores parecidos (*clusters*) ou divergentes (*outliers*). Ao estudar os padrões locais para autocorrelação, é possível ter-se informações sobre a presença de heterogeneidade espacial. Podendo então, descobrir associações espaciais do tipo alto-alto (AA) ou baixo-baixo (BB), e também do tipo alto-baixo (AB) ou baixo-alto (BA) (ALBUQUERQUE, 2009).

Para uma melhor contextualização, as associações do tipo AA, representam o agrupamento dos municípios com uma alta produtividade de milho, já as associações BB representam os municípios com uma baixa produtividade de milho, e os agrupamentos AB e BA são os municípios com correlação negativa (*outliers*), formando grupos locais com valores divergentes (ALBUQUERQUE, 2009).

Segundo Celebioglu e Dall'erba (2009), a estatística *LISA*, ou índice *I* de Moran local, pode ser explicado pela Equação (5):

$$I_i = \frac{x_i - \mu}{\sigma_0^2} \sum_{j=1}^n w_{ij} (x_j - \mu), \quad i = 1, \dots, n \quad \text{Eq. (5)}$$

Em que, σ_0^2 a variância populacional da variável em estudo dos n polígonos (cidades); x_i é a observação de uma variável de interesse no polígono i para $i = 1, \dots, n$, μ é a média dos n polígonos (populações).

2.7.6 Autocorrelação Espacial Local Multivariada

Da mesma maneira que pode-se calcular a estatística de autocorrelação espacial global multivariada, também é possível se medir um coeficiente de autocorrelação espacial local multivariado (ALMEIDA, 2007a).

Segundo Anselin *et al.* (2004), o *I* de Moran local multivariado “dá uma indicação do grau de associação linear (positiva ou negativa) entre o valor para uma variável em uma dada locação i e a média de outra variável nas localizações vizinhas”. O coeficiente *I* de Moran local multivariado é explicado pela equação 6.

$$I_{k,l} = \left(\frac{Z'_K W_{Z,L}}{N} \right) \quad \text{Eq. (6)}$$

O indicador local de associação espacial (*LISA*) fornece o grau de autocorrelação espacial, estatisticamente significativo, em cada unidade regional (HADDAD e PIMENTEL, 2004). Como destaca Almeida (2007a), podem-se mapear os valores do *I* de Moran local multivariado estatisticamente significativos, criando um mapa de significância bivariado do *I* de Moran Local.

De acordo com Anselin *et al.* (2004), a interpretação dessa estatística (*LISA M*), representa uma indicação do grau de associação linear (positiva ou negativa) entre o

valor de um determinado elemento em uma certa localização (cidade por exemplo) i , e a média de outra variável nas localidades próximas.

2.7.7 Diagrama de Dispersão de Moran

O Diagrama de dispersão de Moran compara os valores normalizados do atributo em uma determinada região com a média normalizada das áreas vizinhas, gerando assim, um gráfico bidimensional de Z (Valores normalizados) por WZ (média dos vizinhos). Este diagrama é uma maneira de visualizar a autocorrelação espacial e poder apresentar os diferentes regimes espaciais referentes ao estudo. A Figura 5 ilustra os quatro quadrantes, Baixo-Alto, Alto-Alto, Baixo-Baixo e Alto-Baixo que representam os quatro padrões de associação local espacial entre as áreas e suas regiões vizinhas (ALMEIDA, 2007b).

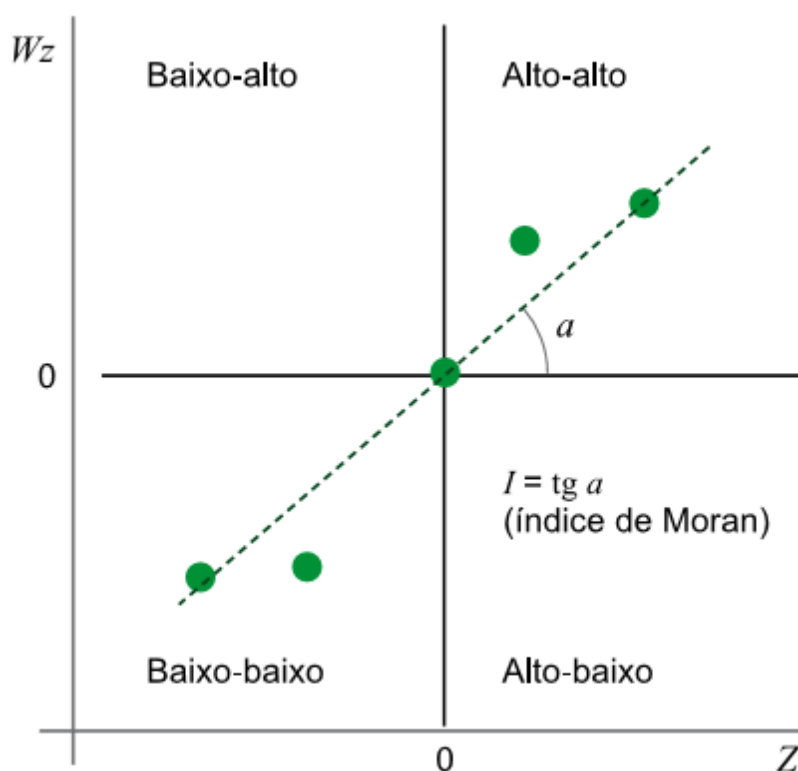


Figura 5 - Diagrama de espalhamento de Moran
 Fonte: Druck *et al.* 2004

Segundo Almeida (2006), O coeficiente I de Moran será a inclinação da curva de regressão de WZ contra Z e indicará o grau de ajustamento. O agrupamento alto-alto (AA), representa a área que possui alto valor para a variável, ou seja, valores acima da média, assim como seus vizinhos. O agrupamento baixo-baixo, expressa os

locais com valores baixos comparados com os atributos analisados, acompanhados também com seus vizinhos que possuem valores baixos. O agrupamento baixo-alto, é formado por atributos com valores baixos na região estudada, cercada por vizinhos com valores altos. O último agrupamento Alto-Baixo é constituído por áreas com valores altos, cercadas por regiões vizinhas que possuem baixos valores (ALMEIDA, 2007b).

As áreas do estudo que formarem um *cluster*, ou seja, tiverem valores parecidos, ocorrem nos quadrantes Alto-Alto e Baixo-Baixo, apresentando então, autocorrelação positiva. As regiões que são representadas pelos quadrantes Alto-Baixo e Baixo-Alto, apresentam autocorrelação negativa, então os valores dos atributos são divergentes nos seus vizinhos (ALMEIDA, 2007b).

Portanto, os quadrantes Alto-Baixo e Baixo-Alto apresentam os *outliers* espaciais do conjunto de dados, enquanto observações situadas nas associações Alto-Alto e Baixo-Baixo representam pontos de alavancagem (ARAÚJO, 2012).

2.7.8 Representação Gráfica: *Box Map*, *LISA Map* e *Moran Map*

De acordo com Neves *et al.* (2001), estas três formas gráficas são baseadas nos valores calculados pelos indicadores locais e também do gráfico de espalhamento de Moran.

Para Serrano e Valcarce (2000), uma das maneiras de se descobrir os *outliers* espaciais, para dados de área, é por meio do *Box Map*. Este mapa é um complemento do Diagrama de Espalhamento de Moran, em que os atributos de cada um dos quadrantes do gráfico, são representados por uma cor específica com seus respectivos polígonos. Como exemplo, a Figura 6 que indica a distribuição do percentual de idosos dos bairros de São Paulo.

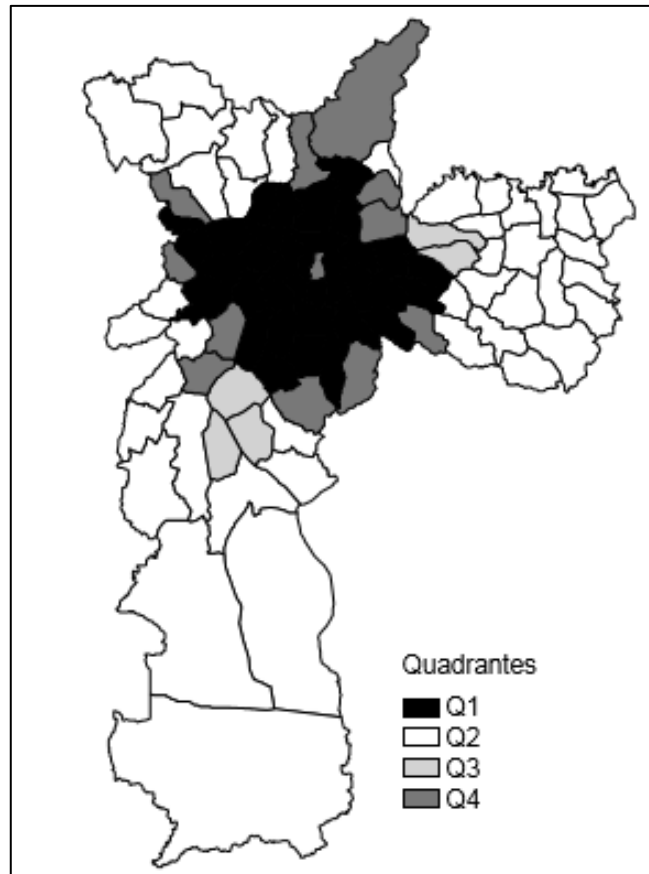


Figura 6 - (Box Map) Mapa de Inclusão/Exclusão Social de São Paulo
Fonte: <http://www.dpi.inpe.br/geopro/exclusao/mapas.html> (2000)

O *LISA Map* é utilizado para a análise dos Indicadores Locais de Associação Espacial, este mapa representa as áreas que possuem correlação local significativamente divergente das outras regiões do estudo. Estas áreas têm uma dinâmica especial própria e precisam ser analisadas mais detalhadamente. Na criação deste mapa os índices são classificados conforme Figura 7, não significantes (cor preta), com significância de 95% (cor branca), 99% (cor cinza) e 99,9 (cor cinza escuro) (KREMPI, 2004).

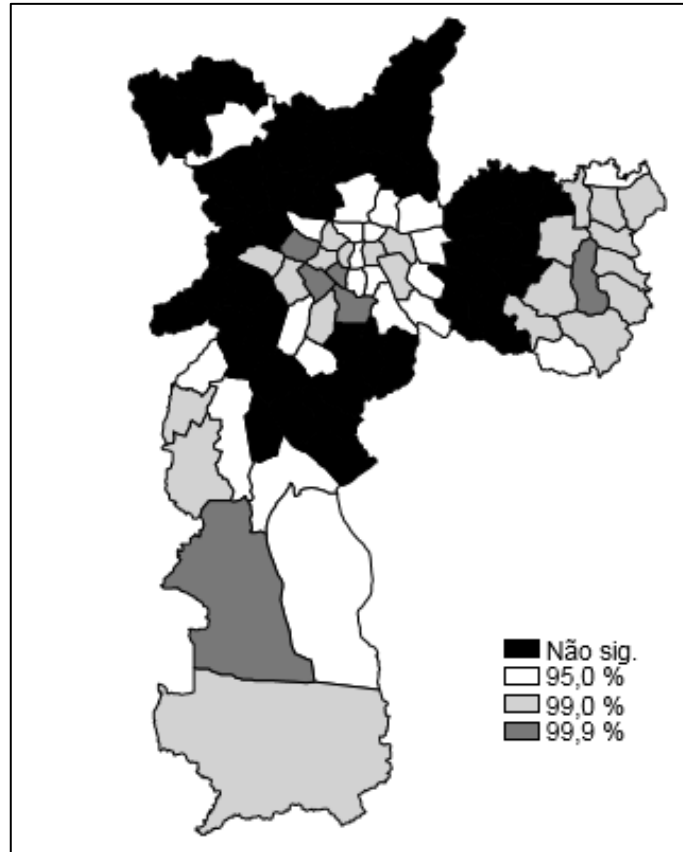


Figura 7 - (LISA Map) Mapa de Inclusão/Exclusão Social de São Paulo
Fonte: <http://www.dpi.inpe.br/geopro/exclusao/mapas.html> (2000)

Por fim, outra maneira de visualizar as informações espaciais de área é por meio do *Moran Map*, conforme Figura 8. Este mapa é parecido com o *LISA Map*, apresentam somente os valores locais que possuem significância ($p > 0,05$), porém são classificados em quatro grupos, conforme sua localização no quadrante do gráfico de espalhamento de Moran, além dos não significantes (que estão com valores zero na legenda) (KREMPI, 2004).

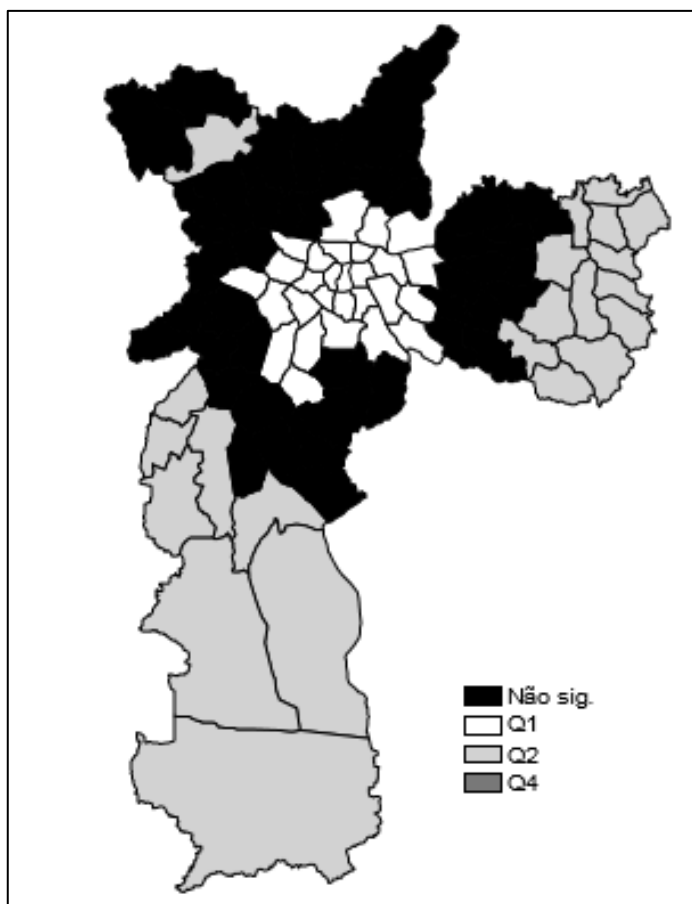


Figura 8 - (Moran Map) Mapa de Inclusão/Exclusão Social de São Paulo
 Fonte: <http://www.dpi.inpe.br/geopro/exclusao/mapas.html> (2000)

Segundo Anselin (1993), realizando uma análise gráfica de autocorrelação espacial, é possível detectar os pontos discrepantes globais e locais, também chamado de *Outliers*. Os *Outliers*, são conhecidos como observações que não possuem um valor semelhante que a maioria das informações em conjunto de dados, na estatística espacial, eles podem ser classificados de duas maneiras, *outliers* globais e locais. Os globais são os que se distanciam muito do restante das outras observações, esse distanciamento pode ser perante valores altos como também para baixos. A identificação desses pontos discrepantes globais pode ser feita por meio do mapa *Box map*, caracterizada como uma ferramenta para detectar *outliers* globais superiores.

2.8 ÁREAS DE INFLUÊNCIA

A precipitação média em uma área é considerada como a altura de água uniforme interceptada por essa mesma região em um determinado período de tempo.

Por isso na literatura, são definidas três medidas para a precipitação média, que são a nível anual, mensal ou diária (CAMURÇA, 2011).

A precipitação se altera temporal e espacialmente e o conhecimento dessa distribuição e variação da precipitação é muito importante para os estudos hidrológicos. Para se estimar a precipitação média de uma determinada superfície, é preciso usar as observações que foram captadas pelos postos dentro dessa região e também nas suas vizinhanças. Portanto, existem três métodos para estimar a chuva média: método da média aritmética, método de Thiessen e método de Isoietas (PEDRAZZI, 1999).

Para Lima (2000), estes três métodos citados anteriormente para as análises de precipitação possuem as seguintes características: a média aritmética é um método simples e eficiente, desde que os pluviômetros sejam em número suficiente para que forneçam dados coerentes e possam compor a média da precipitação da bacia em questão. O polígono de Thiessen é mais preciso que a média aritmética e não necessita da distribuição uniforme dos pluviômetros, mas, ainda, muito trabalhoso e por último o método das isoietas que permite boa precisão na visualização da precipitação na área, sendo bastante utilizado nos estudos de distribuição espacial da chuva.

O método dos polígonos de Thiessen é indicado quando não há distribuição uniforme dos postos pluviométricos dentro da bacia hidrográfica. Esse método dá bons resultados quando o terreno não é muito acidentado. Consiste em dar pesos aos totais precipitados medidos em cada posto pluviométrico, sendo estes pesos proporcionais à área de influência de cada posto. São considerados os postos inseridos na bacia, bem como postos localizados na região de entorno e que exercem influência na bacia (PINTO *et al.*, 1976).

A equação matemática para o cálculo da precipitação média \bar{P} pelo método de Thiessen, para uma determinada área, é explicada pela equação 7:

$$\bar{P} = \frac{\sum_{i=1}^n A_i P_i}{A} \quad \text{Eq. 7}$$

Onde a variável P_i representa o *i*ésimo posto como dados de precipitação em milímetros, A_i é a área de influência de cada posto em Km² e A é a área total.

Unwin e Unwin (1998) ressaltam que a área de influência para cada estação meteorológica pode ser fornecida pela utilização do método dos polígonos de

Thiessen. Inicialmente, é necessário traçar as linhas que ligam os postos pluviométricos mais próximos. O próximo passo, é determinar o ponto médio para cada uma dessas linhas de ligação, e a partir deste ponto, é traçada uma linha perpendicular. A intercepção dessas linhas médias entre si e com os limites da bacia vão definir a área de influência para cada um dos postos.

Observa-se na Figura 9 que existem 12 polígonos, chamados de polígonos de *Voronoi*, cada um representando um local. Os 7 polígonos externos se estendem infinitamente no plano, e por isso são representados como figuras abertas. Cada aresta do diagrama constitui um lugar onde os pontos são equidistantes a dois locais. Os vértices dos polígonos estão ligados a três ou mais arestas, e, portanto, são pontos de equidistância entre três ou mais locais (SILVEIRA 2006).

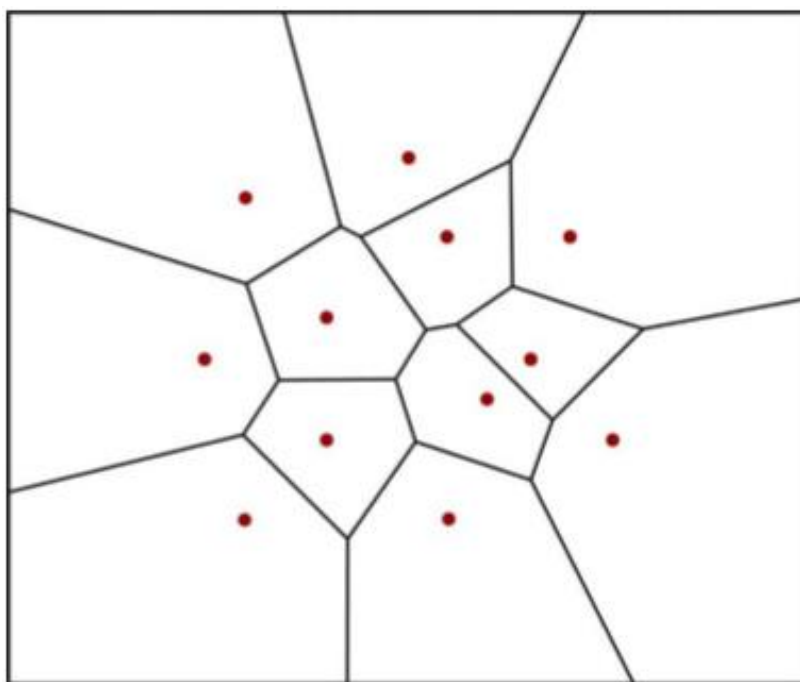


Figura 9 - Determinação de áreas de influência pelo método de polígono de Thiessen para 12 locais
Fonte: SILVEIRA, 2006

Para Silveira (2006), a construção dos diagramas de *voronoi*, podem ser realizadas da seguinte maneira:

- Conectar cada ponto amostral ao seu vizinho mais próximo, por meio de segmentos de reta;
- Construir bissetrizes, formando nos segmentos de retas que conectam os pontos;
- Unir todas as bissetrizes nas retas que conectam os pontos; e

- Unir retas bissetrizes, formando o polígono que delimitam a área de influência de um ponto amostral;

Com a realização deste procedimento, cada estação meteorológica, com sua área calculada, possui seu valor específico da variável medida, que neste caso seria a precipitação. Em relação aos municípios, eles podem representar um ou mais polígonos e o valor da variável deve ser atribuído para cada um deles, levando em consideração a área no qual a estação meteorológica influencia no município.

2.9 REGRESSÃO ESPACIAL

De acordo com Câmara *et al.* (2005), um modelo de regressão baseia-se no relacionamento entre duas ou mais variáveis de maneira que uma delas possa ser explicada ou ter seu valor predito por meio de outras variáveis. Para dados espaciais, caso seja verificado autocorrelação espacial, o modelo criado deve ser incorporado na estrutura espacial, pelo fato da dependência entre as observações afetar diretamente a capacidade de explicação do modelo.

A aplicação da regressão deve ser utilizada caso a correlação entre as variáveis seja significativa. Enquanto a correlação calcula o grau de relacionamento entre duas variáveis ou mais, a regressão fornece uma equação que explica o comportamento matemático entre elas. O intuito da aplicação de uma análise de regressão, de uma maneira geral, é mensurar a relação entre uma determinada variável de resposta dependente (Y) e um conjunto de variáveis explicativas independentes (X_1, X_2, \dots), conforme a Equação 8.

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_\rho X_\rho + \varepsilon \quad \text{Eq. 8}$$

Em que β 's são os coeficientes de regressão; ρ : é o número de covariáveis ou variáveis explicativas; e ε é o erro não observado ou resíduo.

Na regressão clássica, Bailey e Gatrell (1995), contextualizam que o método dos mínimos quadrados ordinários provê um modelo no qual o valor gerado é a soma dos quadrados dos erros. Além disso, para efetuar inferências, é necessário que os erros tenham uma distribuição normal, média zero e variância (σ^2) constante e independente entre si. Portanto, matematicamente, tem-se:

$$\varepsilon \sim N(0, \sigma^2) \text{ e } Cov(\varepsilon_i, \varepsilon_j) = 0, \text{ para } i \neq j, \quad \text{Eq. 9}$$

Isto significa que a variável de Y também deve ter uma distribuição normal, no qual pode ser explicada pela equação: $Y = X\beta + \varepsilon$, com $E(\varepsilon) = 0$ e $E(\varepsilon\varepsilon^T) = C$.

No entanto, quando aplicado a procedimentos baseados pelos mínimos quadrados esbarra-se nos erros serem independentes e também a variância desses erros ser constante. Outra situação, é que a variável Y pode não ter uma distribuição normal, sendo necessária uma transformação para a normalidade, o que nem sempre acontece.

Portanto, para introduzir explicitamente efeitos espaciais em modelos de regressão, existem diferentes maneiras, uma delas é conhecida como modelo com efeitos espaciais globais, estes modelos tem o objetivo de capturar a estrutura de correlação espacial em apenas um parâmetro e adicioná-lo no modelo de regressão clássico. Existem duas alternativas, uma delas é pelo modelo SAR – *Spatial LAG Model*, e a outra é pelo modelo CAR – *Spatial Error Model* (BAILEY e GATRELL, 1995).

2.9.1 Modelo de Regressão SAR – *Spatial LAG Model*

Câmara *et al.* (2005) explica o modelo SAR como a dependência espacial em meio a adição ao modelo de regressão de um novo termo na forma de uma relação espacial para a variável dependente. Portanto, pode ser explicado pela equação 10.

$$Y = \rho WY + X\beta + \varepsilon \quad \text{Eq. 10}$$

Em que W é matriz de proximidade espacial, e o produto WY é a dependência espacial em Y e ρ é o coeficiente espacial autoregressivo. A hipótese nula para a não existência da autocorrelação espacial pode ser dada como $\rho = 0$. O objetivo deste modelo é introduzir a autocorrelação espacial como componente dele. Portanto, este modelo pode ser explicado conforme equação 11.

$$y_i = \left(\sum_j w_{ij} y_j \right) + \sum_{i=1} x_j \beta_i + \varepsilon_i \quad \text{Eq. 11}$$

2.9.2 Modelo de Regressão CAR – *Spatial Error Model*

De acordo com Câmara *et al.* (2005), o modelo de regressão espacial CAR considera que os efeitos espaciais são um ruído, ou perturbação, por consequência

este erro precisa ser removido. Com isso, os efeitos da autocorrelação espacial são associados ao termo de erro ε e o modelo pode ser explicado pela equação 12.

$$Y = X\beta + \varepsilon, \quad \varepsilon = \lambda W\varepsilon + \xi \quad \text{Eq. 12}$$

Onde $W\varepsilon$ é o componente do erro com efeitos espaciais, λ é o coeficiente autoregressivo e ξ é o componente de erro com variância constante e não correlacionada. A hipótese nula para não existência de autocorrelação espacial é que $\lambda = 0$, ou seja, o termo do erro não é espacialmente correlacionado.

2.10 SISTEMA GERENCIADORES DE BANCO DE DADOS ESPACIAIS

Os Sistemas Gerenciadores de Banco de Dados (SGBD) controlam grandes volumes de dados, e isso é realizado de maneira persistente, organizada e inter-relacionada. Os SIGs, conhecidos como Sistemas de Informação Geográfica, são sistemas de informação que foram desenvolvidos com o objetivo de armazenar, processar e visualizar dados geográficos, e na maioria dos casos, tem como componente central um SGBD (QUEIROZ e FERREIRA, 2005).

No início, os SGBDs tinham o gerenciamento apenas do componente convencional do dado geográfico, enquanto que o componente de localização espacial, era persistido em arquivos externos, sendo gerenciado pelo SIG. A partir dos anos 2000, os SGBDs começaram a disponibilizar as chamadas extensões espaciais, que é um conjunto de tipos, métodos de acesso e mecanismos de indexação espaciais, sendo possível realizar o gerenciamento das informações geográficas pelo próprio SGBD (QUEIROZ e FERREIRA, 2005).

Em consequência, esse crescimento tecnológico tornou possível a evolução das arquiteturas para o desenvolvimento dos SIGs, ou de qualquer sistema que faça o uso de mecanismos espaciais, para diversos ambientes como o monousuário, ou colaborativo e também distribuído no ambiente da internet. Portanto, nessa arquitetura, a maior parte das funcionalidades espaciais são desenvolvidas em meio a consultas ao SGBD, sendo responsável por realizar o processamento das informações espaciais (SHEKKAR e CHAWLA, 2003).

Os SGBDs, utilizam como linguagem de programação a Structured Query Language (SQL), que foi criada para realizar o gerenciamento dos dados em SGBD's relacionais, e as extensões espaciais incluem em sua SQL um vocabulário específico para manipular os dados espaciais (SHEKKAR e CHAWLA, 2003). Portanto, para

realizar as manipulação e organização dos dados espaciais no presente trabalho, foi utilizado o SGBD PostgreSQL com a extensão espacial Postgis.

2.10.1 Banco de Dados PostgreSQL

O PostgreSQL é um sistema gerenciador de banco de dados objeto relacional, gratuito e de código fonte aberto, seu início ocorreu a partir do projeto Postgres, que começou em 1986, na Universidade da Califórnia em Berkeley, sob comando do professor Michael Stonebraker. No ano de 1995, quando o suporte a linguagem SQL foi incorporado no SGBD, o código fonte foi disponibilizado na Web (<http://www.postgresql.org>). A partir de então, um grupo de desenvolvedores passou a manter e melhorar o código fonte sob o nome de PostgreSQL (QUEIROZ e FERREIRA, 2005).

A escolha por esse SGBD para o desenvolvimento do trabalho é por possuir um mecanismo de extensão para dados espaciais (Postgis) com muitos recursos para se aplicar na área de Banco de dados geográficos, além de apresentar uma documentação com todas as informações necessárias referente a dados espaciais. O PostgreSQL também é flexível para se trabalhar com alto volume de dados, sendo extremamente necessário quando se está utilizando dados geográficos.

2.10.2 Extensão Espacial Postgis

PostGIS é uma extensão espacial do PostgreSQL. Essa extensão foi desenvolvida pela empresa canadense *Refraction Research Inc.*, (POSTGIS, 2016).

Refractions é uma empresa de SIG e consultoria em bando de dados, que está situada na cidade de Victoria, no Canadá, especializada nas áreas de integração de dados e desenvolvimento customizado de sistemas. PostGIS é um projeto sob comando da fundação OSGeo, no qual está sempre sendo melhorado e financiado por vários desenvolvedores da FOSS4G, além de corporações por todo o mundo que utilizam e se beneficiam dos seus recursos (POSTGIS, 2016).

A extensão espacial PostGIS disponibiliza várias funções prontas para o desenvolvimento de consultas espaciais. Algumas dessas consultas, por exemplo, envolvem o cálculo de área de cada polígono, percentual de ocupação de um determinado polígono em outro (utilizado para classificar a potencialidade agrícola do solo de cada polígono), distância entre as estações meteorológicas em meio a sua

localização geográfica, criar e calcular os polígonos de Thiessen para simular os dados climáticos de cada polígono (Município).

3 MATERIAIS E MÉTODOS

Este capítulo descreve a metodologia aplicada para atingir os objetivos que foram apresentados nesta dissertação. Em seguida são discutidos os detalhes dos procedimentos utilizados para a obtenção dos dados, construção da base de dados (arquivo Shape do estado do Paraná) e aplicação das técnicas de estatística espacial de área e regressão espacial.

O trabalho é composto pelas etapas apresentadas de maneira estruturada na Figura 10.

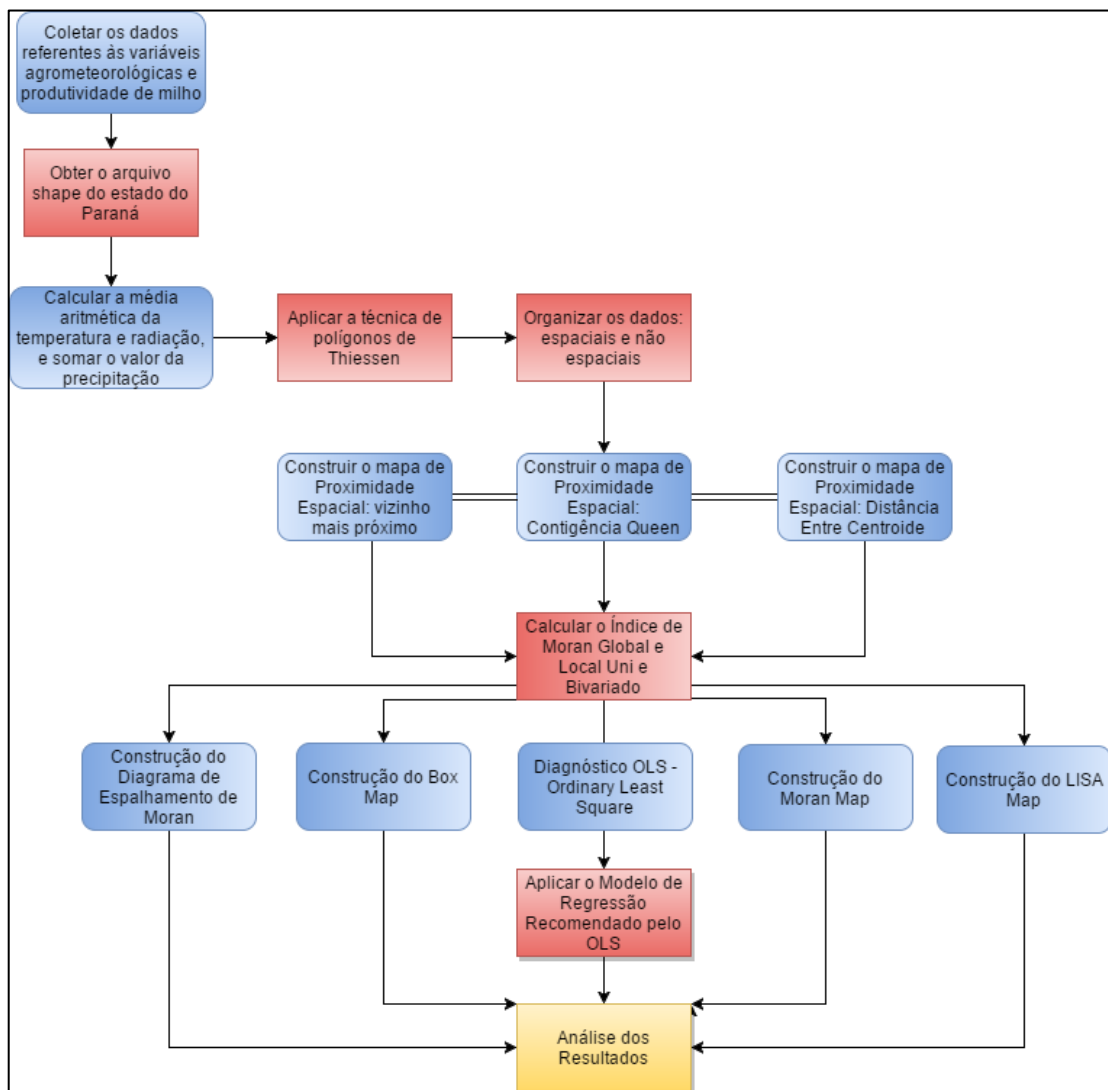


Figura 10 - Diagrama dos processos a serem realizados no trabalho

Nos próximos tópicos serão abordadas as técnicas utilizadas nos processos representados na Figura 10.

3.1 OBTENÇÃO DO DADOS

A área de estudo do trabalho abrange o estado do Paraná (Figura 11), fazendo uso de dados referentes a produtividade média do milho da primeira safra (safra de verão) dos municípios do estado, considerando atributos referentes à altitude (m), potencialidade agrícola do solo, precipitação pluvial (mm), temperatura média (°C) e radiação solar (KJ/m²). Os anos safra considerados no estudo foram 2011/2012, 2012/2013 e 2013/2014, e o período analisado para a obtenção dos dados agrometeorológicos são do início do mês de setembro até o final do mês de maio, seguindo o calendário de plantio e colheita da primeira safra do milho do estado do Paraná, de acordo com a CONAB (2015).

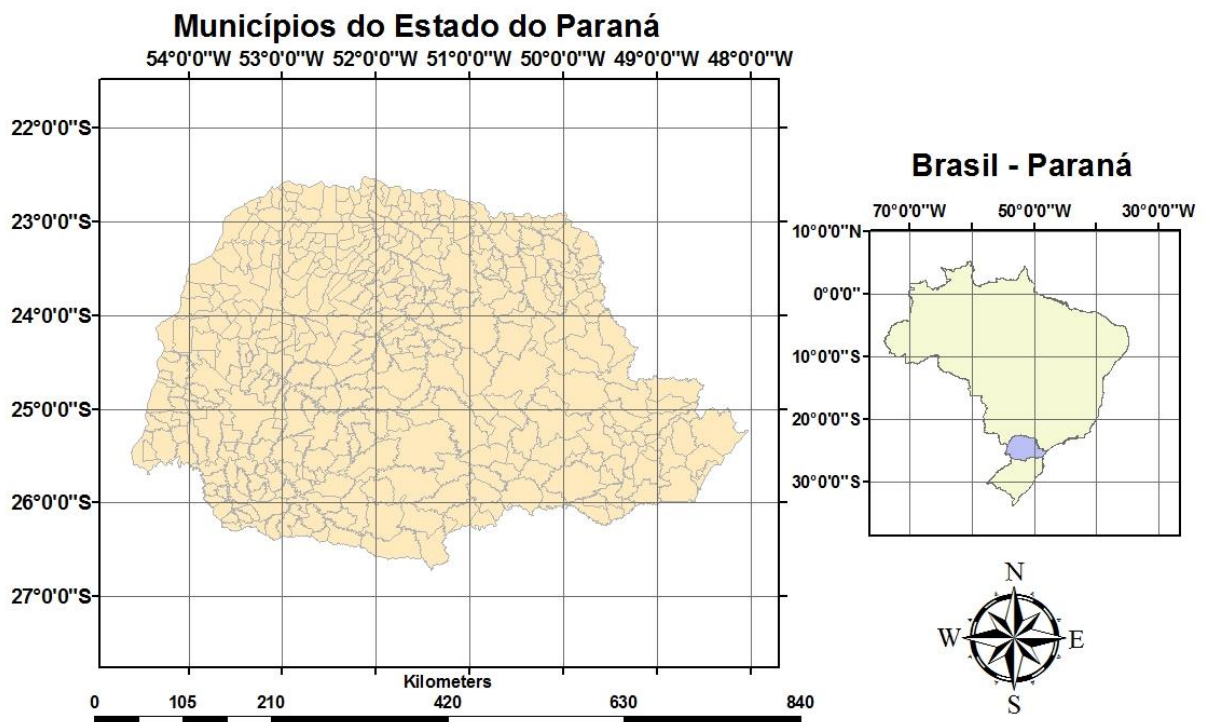


Figura 11 - Mapa do estado do Paraná - Brasil

Os dados referentes a produtividade média de milho da primeira safra e potencialidade agrícola do solo foram obtidos junto ao Instituto Brasileiro de Geografia e Estatística (IBGE, 2016) e os dados de altitude média dos municípios junto ao IPARDES (2016). As variáveis climáticas referentes à precipitação pluvial (mm) e temperatura média (°C) foram obtidas junto as empresas SIMEPAR (2016), INMET (2016) e IAPAR (2016), e os dados referentes a radiação solar (KJ/m²), pelas empresas INMET (2016) e IAPAR (2016). Os dados de potencialidade agrícola do solo e altitude foram obtidos para os 399 municípios do estado do Paraná, apesar de

que, em 35 destes municípios não houve registro de produção de milho no período, sendo estes municípios excluídos do estudo. As variáveis climáticas (precipitação pluvial, temperatura média e radiação solar) foram adquiridas para os municípios onde são disponibilizadas estações meteorológicas.

Alguns municípios não realizam medições diárias de precipitação, radiação solar e temperatura média, ou seja, não possuem estações meteorológicas, nesses casos foi aplicada a técnica de polígonos de Thiessen para simular o valor desses atributos.

Os dados da potencialidade agrícola do solo por região foram classificados com a seguinte divisão: 8 – Boa, 7 - Boa a Regular, 6 - Regular a Boa, 5 – Regular, 4 - Regular a Restrita, 3 – Restrita, 2 - Restrita a Desfavorável e 1 – Desaconselhável.

3.2 ORGANIZAÇÃO DOS DADOS

No domínio de Sistemas de Informação Geográfica (SIG), a maneira como é gerenciado o Banco de Dados Geográfico (BDG) são aplicáveis de acordo com os diferentes tipos e natureza dos dados que precisam ser gerenciados e armazenados. No contexto do presente estudo, existem diferenças entre os dados que são espaciais ou não espaciais. Por exemplo, dado espacial se refere a elementos geográficos que possuem alguma informação referente a sua localização no espaço, portanto, os municípios do estado do Paraná são classificados como elementos espaciais, em meio a isso, foi utilizado um arquivo *shape* adquirido no IBGE (2016) contendo os municípios do estado, onde cada cidade é representada por um polígono (uma área) que contém seus respectivos atributos.

Já os dados não espaciais são os atributos vinculados em cada um dos polígonos utilizados no estudo (municípios do estado do Paraná), são armazenados em um arquivo *.DBF* que é o banco de dados do arquivo *shape*, eles são as informações que caracterizam cada área geográfica, e podem ser atribuídos também para um determinado ponto ou linha no BDG. Portanto, os dados da produtividade média de milho ($t\ ha^{-1}$) da primeira safra, potencialidade agrícola do solo, altitude (m), Temperatura média [TMed] ($^{\circ}C$), Radiação Solar [Rs] (KJ/m^2) e Precipitação Pluvial [Prec] (mm) são classificados como não espacial, esses dados são os que caracterizam cada um dos municípios do estado e serão utilizados na análise espacial de área para a descoberta dos padrões espaciais.

A atribuição de cada uma das variáveis do estudo para os polígonos do arquivo *shape* foram realizadas de maneiras diferentes. Por exemplo, os valores referentes ao rendimento médio do milho e altitude dos municípios, foram importados para o banco de dados e atribuído para seus respectivos polígonos (municípios). Os dados referentes a potencialidade agrícola do solo foram importadas para o banco de dados em meio a um arquivo *shape*, sendo utilizado consultas e funções espaciais do *Postgis* (versão 2.3.0) para verificar a classificação com maior influência no polígono e após isso atribuída ao mesmo, o critério de escolha da potencialidade agrícola do solo foi realizada pela maior ocupação das classificações: 8 – Boa, 7 - Boa a Regular, 6 - Regular a Boa, 5 – Regular, 4 - Regular a Restrita, 3 – Restrita, 2 - Restrita a Desfavorável, 1 – Desaconselhável na área do município

Após obter as medições diárias de Temperatura média [TMed] (°C), Radiação Solar [Rs] (KJ/m²) e Precipitação Pluvial [Prec] (mm) para os municípios do estado do Paraná, essas informações foram padronizadas da seguinte maneira:

- a) foi calculada a média aritmética da Temperatura média [TMed] (°C) e Radiação Solar [Rs] (KJ/m²) durante o período de safra de cada ano; e
- b) o valor da precipitação pluvial [Prec] (mm) foi somado no período de cada safra;

Os municípios que não possuíam estações meteorológicas ou que apresentarem problema de medição em um determinado dia para as variáveis: Precipitação Pluvial, Radiação Solar e Temperatura Média; tiveram os seus valores simulados aplicando a técnica de Polígonos de Thiessen. A criação dos polígonos de Thiessen foi realizada pela extensão espacial *Postgis*.

Para realizar os procedimentos descritos anteriormente foram utilizados o banco de dados PostgreSQL na versão 9.5 (POSTGRESQL, 2016) com a extensão espacial *Postgis* na versão 2.3.0 (POSTGIS, 2016). Os dados climatológicos também foram importados para o banco de dados e posteriormente calculados para serem atribuídos ao seu polígono.

Portanto, o arquivo *shape* contém os seguintes atributos não espaciais para cada polígono (município):

- Identificador do Polígono;
- Nome do município;
- Rendimento médio de milho primeira safra [Prod] (t ha⁻¹);

- Potencialidade agrícola do solo [Pot] (8 – Boa, 7 - Boa a Regular, 6 - Regular a Boa, 5 – Regular, 4 - Regular a Restrita, 3 – Restrita, 2 - Restrita a Desfavorável e 1 – Desaconselhável);
- Altitude [Alt] (m);
- Precipitação Pluvial [Prec] (mm);
- Temperatura Média [TMed] (°C);
- Radiação Solar [Rs] (KJ/m²);

Os municípios que não apresentaram produtividade de milho para algum dos anos: 2011/2012, 2012/2013 e 2013/2014, foram excluídos do estudo por não possuírem dados relevantes para aplicação dos modelos estatísticos.

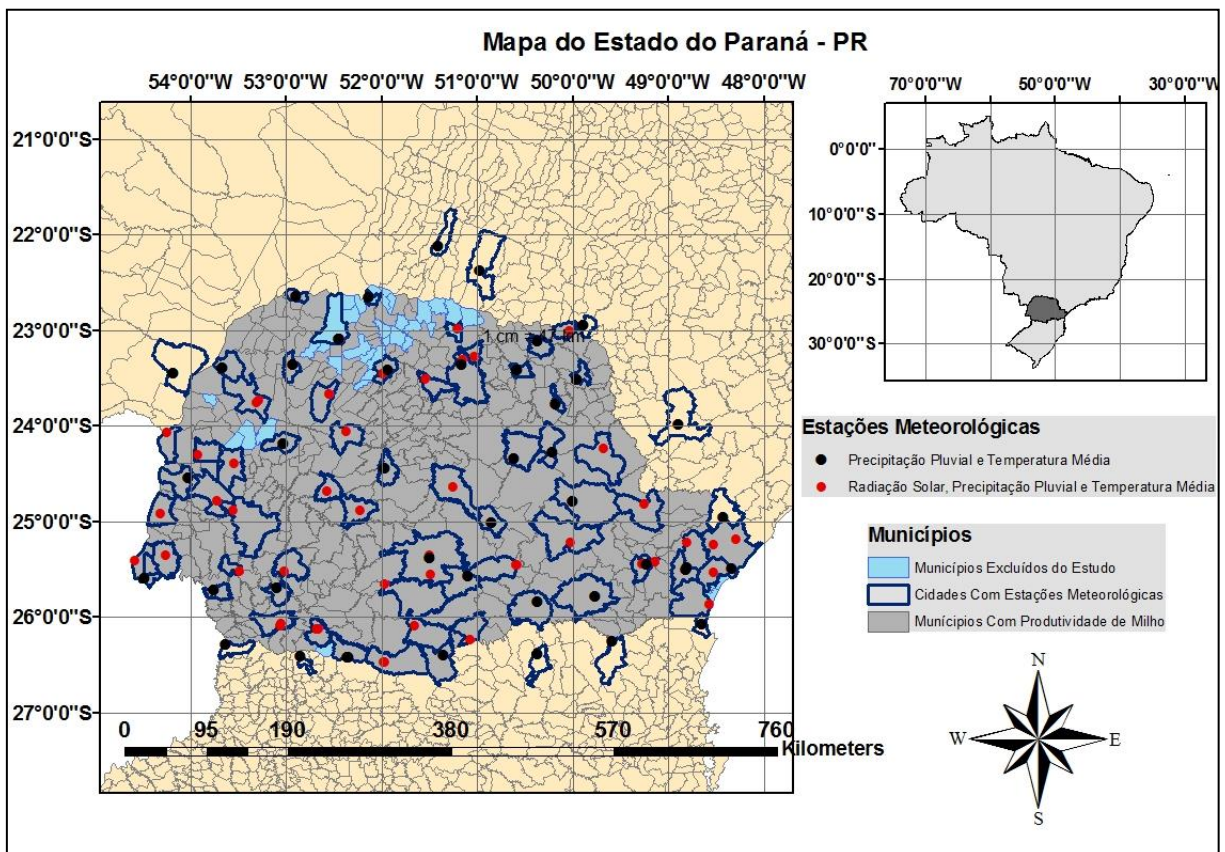


Figura 12 - Mapa dos municípios do estado do Paraná com suas respectivas estações meteorológicas

A Figura 12 apresenta os 364 municípios utilizados no estudo com a cor cinza e as 35 cidades excluídas da análise espacial com a cor azul claro, pode ser visto em destaque com a borda azul as cidades que possuem estações meteorológicas junto com a posição geográfica da mesma, o ponto com a cor vermelha é a localização das estações que possuem dados de radiação solar, temperatura média e precipitação pluvial, o ponto com a cor preta são as estações que possuem dados de precipitação

pluvial e temperatura média. As estações meteorológicas dos estados de Santa Catarina, Mato Grosso do Sul e São Paulo, foram selecionadas, por possuírem influência no valor climático de alguns municípios do estado do Paraná, em meio a aplicação dos polígonos de Thiessen.

3.3 POLÍGONOS DE THIESSEN

Os dados referentes as variáveis radiação solar, temperatura média e precipitação pluvial estão distribuídos para alguns municípios do estado do Paraná, neste contexto, foi necessário aplicar a técnica de polígonos de Thiessen para a simulação dos dados climatológicos para os municípios que não possuem estação meteorológica, segundo Moura (2003), o objetivo do Diagrama de Voronoi ou de Thiessen é de que, considerando que em um plano, existem pontos que estão mais próximos de uma fonte geradora do que de outra fonte, o resultado é um polígono cujas as distâncias entre a fonte e ponto são as menores possíveis.

Portanto, o valor simulado para cada uma das variáveis climáticas utilizadas no estudo foi realizado em meio a influência que as estações meteorológicas exercem sobre os municípios, por exemplo, caso ocorra de um determinado polígono (Município) ter influência de mais de uma estação meteorológica, então o valor será calculado conforme o peso que essas estações estão exercendo sobre ele.

O cálculo dos polígonos de Thiessen para as estações meteorológicas dos municípios do estado do Paraná para o presente estudo, pode ser visto na figura 13, percebe-se que em alguns casos existem estações meteorológicas exercendo influência para vários polígonos, e também, alguns municípios que sofrem influência de várias estações meteorológicas. Neste caso, o valor calculado de cada variável climática para seu respectivo polígono (município) foi realizado conforme a equação 13.

$$vatr_m = \sum_{i=1}^n \frac{A_m \cap A_{ei}}{A_m} Vatr_{ei} \quad \text{Eq. 13}$$

Em que: $vatr_m$ é o valor do atributo atribuído ao município; n é o número de estações meteorológicas; A_m é a área do município; A_{ei} é a área de influência da estação i ; $Vatr_{ei}$ é o valor do atributo da estação i .

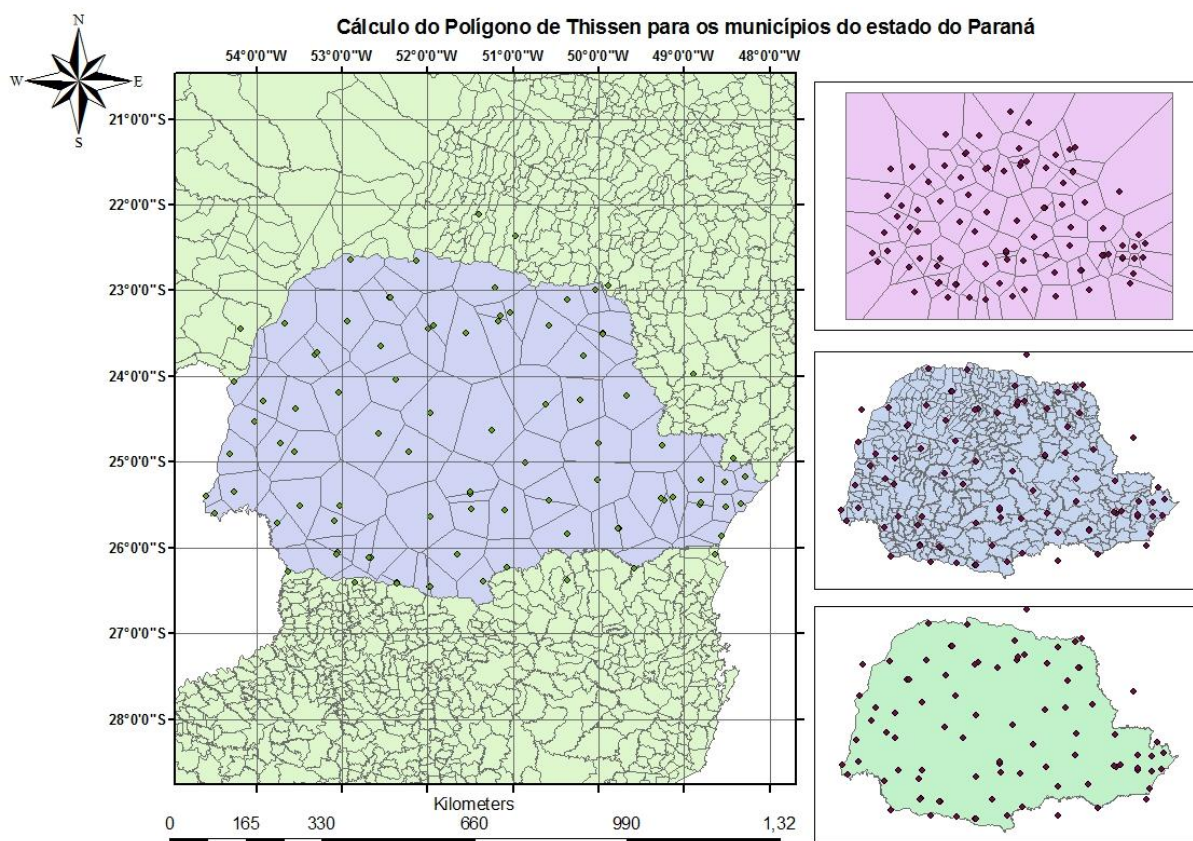


Figura 13 - Polígonos de Thiessen calculados para as estações meteorológicas dos municípios do PR

Com a realização deste procedimento, foram calculados os valores diários da precipitação pluvial, temperatura média e radiação solar para cada município do estado do Paraná, sendo que cada polígono recebeu os pesos proporcionais à área de influência de cada estação meteorológica.

3.4 MATRIZ DE PROXIMIDADE ESPACIAL

O primeiro passo antes de aplicar o índice de Moran para calcular a autocorrelação espacial dos dados é criar as matrizes de proximidade espacial para identificar as regras de vizinhança para cada polígono (município do estado do Paraná). Os critérios de vizinhança utilizados no presente estudo foram: contingência *queen*, distância entre centroides e vizinho mais próximo para a matriz de proximidade *W*.

Por exemplo, a matriz de vizinhança baseada no critério da contingência *queen* atribuiu o valor de 1 para os polígonos que são vizinhos e 0 para os polígonos que não compartilham fronteiras, neste caso, os municípios do estado do Paraná que são vizinhos de um determinado município X irão receber o valor 1, e o restante dos

municípios o valor 0. A figura 14 ilustra o histograma de conectividade da matriz de vizinhança contingência *queen* para os 364 municípios utilizados no estudo. Percebe-se que neste critério a maioria dos municípios possuem entre 5 a 7 fronteiras em comum, tendo poucos casos cidades com mais de 11 vizinhos e menos de 2 vizinhos.

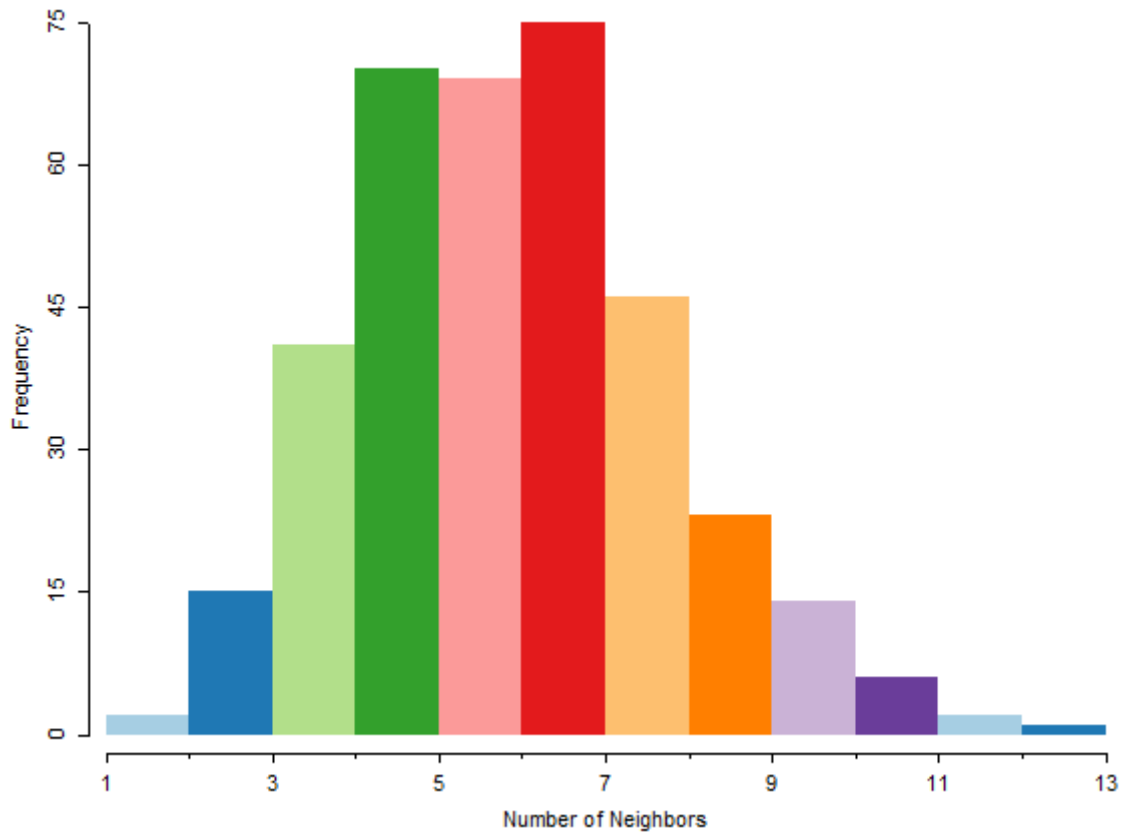


Figura 14 - Histograma de conectividade para o critério contingência *queen*

A matriz da distância entre o centroide considera como vizinhos os polígonos que forem menor que a distância mínima atribuída no cálculo, por exemplo, a matriz utilizada no estudo foi considerado como limite o valor de 35 Quilômetros, o cálculo é realizado a partir do centroide de cada polígono (município), neste caso, os municípios que possuem uma distância menor do que 35 Quilômetros são considerados vizinhos, e os maiores não são atribuídos como vizinhos. A figura 15 mostra o histograma de conectividade para esse critério, pode-se verificar que houve uma distribuição parecida de vizinhos para os 364 municípios, poucas cidades tiveram mais do que 18 vizinhos.

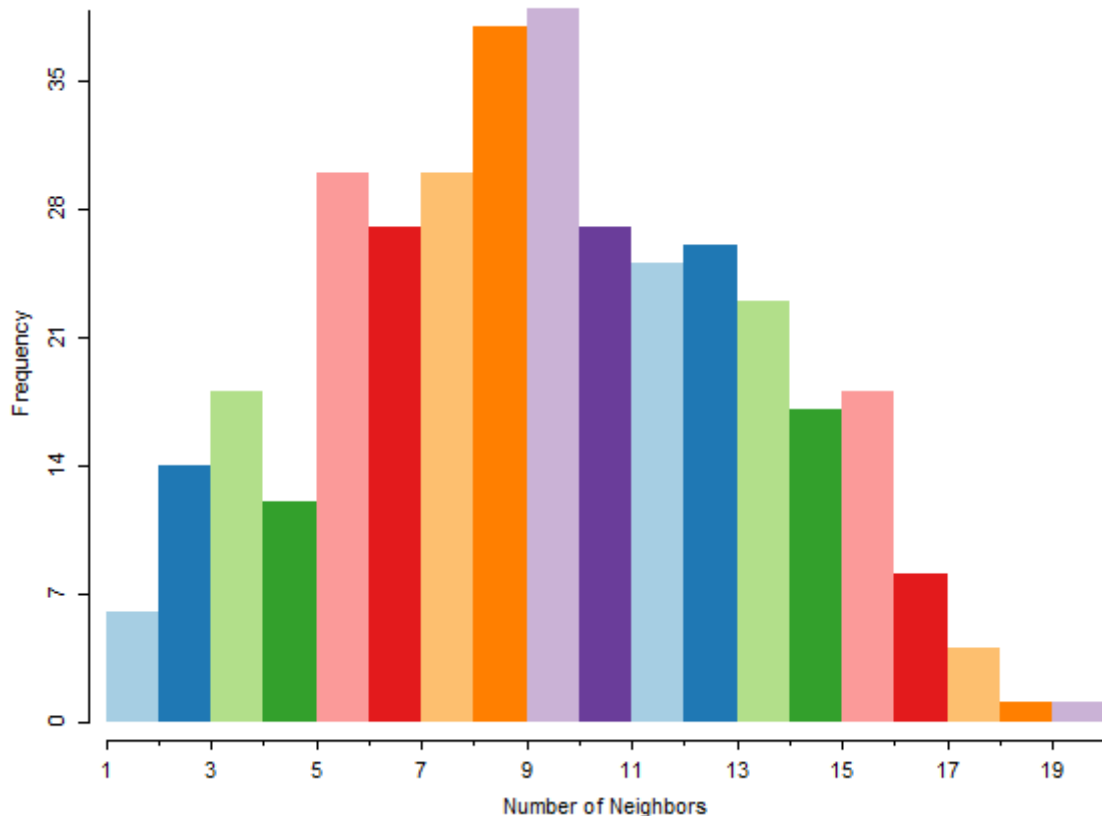


Figura 15 - Histograma de conectividade para o critério da distância entre o centroide

O critério do vizinho mais próximo para a matriz de proximidade W assegura que cada observação tem exatamente o mesmo número de vizinhos. No caso da regra aplicada para este estudo, atribuiu para o número de vizinhos o valor 2, portanto, foi considerado para cada polígono (município) os dois vizinhos mais próximos. Esse valor de vizinhança foi utilizado devido aos resultados possuírem maior autocorrelação espacial significativa ao nível de 1%. Não foi gerado o histograma de conectividade, pois o mesmo não é muito significativo, pelo fato de que cada local ter exatamente 2 vizinhos.

Todas as matrizes abordadas neste capítulo foram criadas pelo software Geoda (Versão 1.6.7).

3.5 ANÁLISE DE DADOS DE ÁREA

Como apresentado no capítulo 2.6.4 a estatística espacial de área ou também conhecida como análise de dados de área, são eventos aonde a localização espacial do dado está associada a área, isso ocorre quando está se trabalhando com fenômenos agregados, como é o caso do presente estudo, aonde os atributos estão

vinculados a cada polígono que representa um município no conjunto de dados, de acordo com Braga et. al. (2010), para essas análises, procura-se identificar padrões espaciais de distribuição nos valores observados. As técnicas de análise de dados de área foram desenvolvidas para tentar identificar regiões onde a distribuição dos valores possa apresentar um padrão específico associado à sua localização espacial. A resposta buscada é o quanto o valor de uma determinada variável em uma área é semelhante com os valores da mesma variável em regiões. Portanto, foi verificada a existência de padrões espaciais (autocorrelação espacial) dos seguintes atributos: Produtividade de Milho, Classificação do Solo, Temperatura Média, Radiação Solar, Precipitação pluvial e altitude entre os municípios do estado do Paraná.

Para a verificação da autocorrelação espacial global univariada foi utilizado o índice de Moran global, que tem o objetivo de analisar se os dados estão distribuídos aleatoriamente no espaço, ou , se os dados estão autocorrelacionados, neste caso, a distância entre os municípios influencia no valor dos municípios mais próximos .

Utilizando como exemplo a produtividade de Milho, caso ocorra autocorrelação espacial positiva (agrupamento *High-High*), indica que os municípios próximos obtiveram resultados semelhantes na produtividade de milho, ou seja, cidades com alta produtividade tendem a ficarem próximas de cidades que também tiveram uma alta produtividade, o mesmo ocorre com os municípios que tiveram uma baixa produtividade de milho (agrupamento *Low-Low*). Quando a autocorrelação espacial é negativa (agrupamento *High-Low* ou *Low-High*), indica a diferença entre os valores dos atributos obtidos nos municípios com localização próxima, neste caso a influência da distância é negativa, concluindo que municípios com baixa produtividade de milho estão próximos de municípios com alta produtividade, ou ao inverso, municípios com alta produtividade de milho rodeados por cidades com baixa produtividade. O mesmo ocorre para as demais variáveis utilizadas no estudo. O índice de Moran global foi calculado para todas as variáveis utilizadas no estudo (Rendimento médio de milho, Classificação do Solo, Temperatura Média, Radiação Solar, Precipitação pluvial e altitude) e pode ser calculado conforme a equação 1.

A autocorrelação espacial global bivariada verifica a correlação espacial entre duas variáveis de interesse. O objetivo é responder se os valores da variável observada em certa localização tendem a ter uma relação com outra variável em locais vizinhos. Por exemplo, no caso da produtividade de milho, foi calculada a correlação entre os dados de produtividade de cada ano safra com as variáveis dependente:

Classificação do Solo, Temperatura Média, Radiação Solar, Precipitação pluvial e altitude. O índice de Moran bivariado global pode ser calculado conforme a equação 4.

Como área do estudo é considerada grande é necessário aplicar o índice LISA, conhecido como indicador local de associação espacial. Essa análise é importante, pois o índice de Moran global não revela padrões de associação local, que podem ser ocultadas nas estatísticas de autocorrelação global. O índice LISA foi utilizado na análise de autocorrelação espacial univariada e bivariada, as equações 5 e 6 detalham como calcular, respectivamente.

Para visualizar a autocorrelação espacial da produtividade de milho em cada município foi necessário criar mapas de agrupamento para apresentar os padrões espaciais. O mapa de espalhamento de Moran dividiu os municípios em sua determinada classificação, que são do tipo: H-H (*High-High*), L-L (*Low-Low*) L-H (*Low-High*) e H-L (*High-Low*). Os agrupamentos H-H, são para as áreas com valores altos para um determinado atributo, sendo cercado por municípios que apresentam valores elevados para o mesmo atributo; LL, para áreas com valores baixos para um determinado atributo, sendo rodeado por áreas que apresentam valores baixos para o mesmo atributo; HL, para áreas com valores altos para um atributo particular, cercado por vizinhos exibindo valores baixos para o mesmo atributo; e por último LH, para áreas com valores baixos para um determinado atributo e rodeado por vizinhos com valores altos para este atributo.

As análises exploratórias dos dados espaciais foram realizadas pelos softwares: Geoda na versão 1.6.7 (GEODA, 2016) e ArcMap 9.3 (ESRI, 2016).

3.6 MODELOS DE REGRESSÃO ESPACIAL

Modelos de Regressão são técnicas estatísticas que fazem o uso do relacionamento existente entre duas ou mais variáveis, de maneira que o valor de uma delas possa ser estimado pelas outras variáveis, como é o caso do atual trabalho, em que a produtividade de milho que é a variável dependente foi estimada pelas variáveis explicativas: classificação do solo, temperatura média, radiação solar, precipitação pluvial e altitude. E quando é estudado comportamento de dados espaciais e está presente a autocorrelação espacial, as estimativas do modelo devem utilizar essa estrutura espacial, pelo fato da dependência entre as observações alterar o poder explicativo do modelo.

Como a maioria das variáveis geográficas são espacialmente autocorrelacionadas, modelos de regressão espacial são mais adequados do que modelos que não levam em consideração a autocorrelação espacial para analisar as relações entre as variáveis dependentes e independentes, como é o caso do modelo OLS - *Ordinary Least Square*. Métodos de regressão espacial podem ser divididos em SAR – *Spatial LAG Model* e CAR – *Spatial Error Model*. O primeiro leva em conta a correlação espacial das variáveis dependente e pode ser explicado pela equação 10, enquanto o segundo considera os efeitos de autocorrelação dos erros aleatórios, conforme equação 12. Para este estudo, a variável dependente (Y) é a que representa a produtividade média de milho, e as independentes (X1, X2, X3, X4 e X5), são as variáveis classificação do solo, altitude, temperatura média, precipitação pluvial e radiação solar.

Para determinar o modelo mais apropriado (OLS, SAR ou CAR) para a estimativa da produtividade de milho, o processo de decisão conduzido por (Anselin, 2005 e Song *et al.*, 2014) foi aplicado, sendo ilustrado na Figura 16.

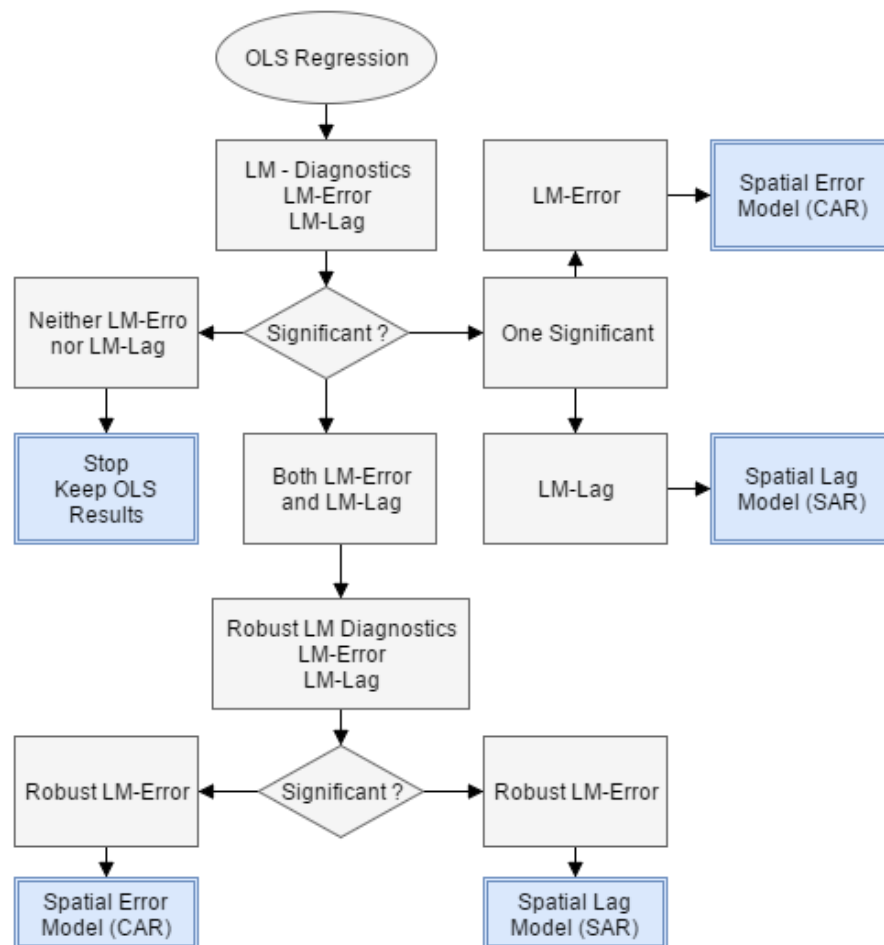


Figura 16 - Decisão do Processo de Regressão Espacial (Anselin, 2005; Song, 2014)

O modelo OLS é primeiro ajustado para se obter diagnósticos de regressão para a dependência espacial dos resíduos, e em seguida, quatro testes estatísticos são conduzidos para detectar a presença de dependência espacial em modelos lineares, neste caso, foi verificado se os dados que estão sendo utilizados para a explicação da produtividade de milho possuem autocorrelação espacial. Os testes estatísticos referem-se ao LM Simple e LM Robust Lagrange Multiplier, incluindo os testes: LM Lag, LM Error, Robust LM Lag, e Robust LM Error. Se o teste LM Lag é significativo, a presença de dependência espacial está implícita, e, assim, o modelo SAR deve ser selecionado para a estimativa da produtividade de milho; se o teste LM Error é significativo, o modelo CAR deve ser escolhido para a explicação da produtividade de milho. Se ambos os testes (LM Lag ou LM Error) forem significativos, o modelo espacial cujo o resultado do teste LM Robust for mais significativo deve ser escolhido para estimar a produtividade de milho. No presente estudo, o modelo de regressão espacial SAR foi recomendado pelo diagnóstico nos anos 2011/2012, 2012/2013, 2013/2014 e o modelo de regressão espacial CAR foi recomendado para os anos de 2013/2014.

3.7 SOFTWARES UTILIZADOS

O quadro 1 detalha as técnicas de estatística espacial e os softwares que foram utilizados no trabalho para a realização de cada etapa.

Etapas	Procedimentos	Software
Análise da autocorrelação espacial	I de Moran Global Diagrama de Espalhamento de Moran	OpenGeoda 1.6.7
Produção de indicadores Globais (Variáveis Espaciais Globais)	Diagrama de Espalhamento de Moran(Quadrantes para cada ZT)	OpenGeoda 1.6.7
Produção dos Indicadores Locais	I Moran Local	OpenGeoda 1.6.7
Construção dos Mapas	<i>Moran Map</i> e os resíduos padronizados dos modelos de regressão	ArcView 9.3
Aplicação do modelo de Regressão	Modelo SAR e CAR	OpenGeoda 1.6.7
Matriz de Vizinhança	Contingencia <i>Queen</i> Vizinho mais próximo Distância entre o centroide	OpenGeoda 1.6.7
Simulação de dados climatológicos	Polígonos de <i>Thiessen</i>	PostgreSQL 9.5 (Postgis 2.3.0)

Organização dos dados espaciais e não espaciais	Organização dos dados	PostgreSQL 9.5 (Postgis 2.3.0)
---	-----------------------	--------------------------------

Quadro 1- Softwares utilizados para o desenvolvimento do trabalho**Fonte: Autoria Própria**

Esta abordagem é importante para saber as ferramentas que foram aplicadas para o desenvolvimento do trabalho.

4. RESULTADOS E DISCUSSÕES

A Tabela 1 apresenta o índice global de Moran para as variáveis utilizadas no estudo, podendo-se observar que todas apresentaram autocorrelação espacial em um nível de significância de 1% de probabilidade, segundo os critérios de contiguidade *queen*, distância entre o centroide e vizinhos mais próximos. Os valores de z-score foram maiores que 2,71 para todos os anos e critérios de vizinhança, corroborando com a significância da autocorrelação espacial das variáveis do estudo. Wrublack (2013) utilizou o índice de Moran global para verificar a dependência espacial da produtividade de Canola no estado do Paraná, a autora concluiu que a cultura também apresenta autocorrelação espacial significativa a 5%, Bohórquez (2011) usou o índice de Moran para identificar áreas com possíveis ocorrências de queimadas que ocasionaram desmatamento no Parque Nacional de Tuparro na Colômbia, já Zhang (2008) utilizou o índice de Moran para verificar a autocorrelação espacial da presença de carbono orgânico no solo no sudoeste da Irlanda e também investigou os pontos críticos de contaminação do solo por presença de Chumbo (Pb) na cidade de Galway na Irlanda, respectivamente. Portanto, o índice de Moran é utilizado em diversos estudos que envolvem a descoberta de algum padrão espacial que está relacionado com a localização geográfica dos eventos nas áreas de interesse.

Os três critérios de vizinhança utilizados para o cálculo do índice de Moran global resultaram em valores semelhantes em todos os anos e variáveis estudadas, com destaque para a regra dos vizinhos mais próximos que obteve a maior autocorrelação em todas as variáveis. Percebe-se que os anos de 2013/2014 para a variável Prod (produtividade de milho) tiveram os melhores resultados de autocorrelação espacial, sendo que com o critério dos vizinhos mais próximos obteve o valor de 0,8306. Os anos de 2011/2012 apresentou o menor índice de autocorrelação entre os três anos, mais mesmo assim, resultou em uma forte autocorrelação espacial, concluindo que municípios próximos possuem uma produtividade de milho parecida.

A variável PSolo (potencialidade agrícola do solo) também apresentou autocorrelação espacial positiva, portanto, a estatística corroborou com a lei da geográfica de que regiões próximas tendem a ter características semelhantes, o mesmo ocorreu para a variável Alt (altitude).

As variáveis climáticas Prec (precipitação pluvial), TMed (temperatura média) e Rs (Radiação solar global média) apresentaram os maiores índices de autocorrelação

espacial, isso pode ser explicado facilmente, pelo fato de regiões próximas possuíram características agrometeorológicas semelhantes, por exemplo, dificilmente os vizinhos de um determinado município x terão uma temperatura média ou radiação solar muito discrepante de x, mais a medida que a distância vai aumentando essas variáveis podem se alterar, o mesmo ocorre para a precipitação, pois a probabilidade de chover também nos municípios vizinhos é maior do que em uma região mais distante.

Tabela 1 - Índice de Moran global aplicado para as variáveis do estudo

Variáveis	Índice Global de Moran								
	Contiguidade (<i>queen</i>)			Distância entre centroides			Vizinhos mais próximos		
	2011/ 2012	2012/ 2013	2013/ 2014	2011/ 2012	2012/ 2013	2013/ 2014	2011/ 2012	2012/ 2013	2013/ 2014
Prod	0,6263** (18,87)	0,6970** (20,37)	0,7885** (24,36)	0,6180** (21,93)	0,6879** (25,49)	0,7651** (27,42)	0,6856** (13,87)	0,7497** (14,85)	0,8306** (17,37)
PSolo	0,6757** (21,19)	0,6757** (21,19)	0,6757** (21,19)	0,6000** (21,22)	0,6000** (21,22)	0,6000** (21,22)	0,7315** (15,10)	0,7315** (15,10)	0,7315** (15,10)
Alt	0,6900** (20,32)	0,6900** (20,32)	0,6900** (20,32)	0,6369** (22,74)	0,6369** (22,74)	0,6369** (22,74)	0,6945** (13,86)	0,6945** (13,86)	0,6945** (13,86)
Prec	0,8880** (27,54)	0,7898** (23,75)	0,8423** (25,83)	0,8579** (30,81)	0,7599** (27,15)	0,8029** (27,89)	0,9088** (18,94)	0,8691** (17,95)	0,8927** (17,70)
TMed	0,9364** (28,41)	0,9299** (27,68)	0,9261** (28,94)	0,9091** (31,21)	0,9027** (31,10)	0,8997** (32,55)	0,9533** (19,51)	0,9500** (20,07)	0,9501** (20,01)
Rs	0,8656** (26,58)	0,8827** (26,59)	0,8904** (27,28)	0,8500** (28,53)	0,8732** (30,81)	0,8994** (32,47)	0,9043** (18,56)	0,9257** (19,26)	0,9297** (20,03)

Prod: produtividade de milho ($t\ ha^{-1}$); PSolo: Potencialidade agrícola do Solo; Alt: Altitude (m); Prec: precipitação pluvial (mm); TMed: temperatura média do ar ($^{\circ}C$); Rs: radiação solar global média (KJ/m^2). Entre parênteses tem-se o Z-score. * Significância ao nível de 0,05. ** Significância ao nível de 0,01.

Na Figura 17, apresenta-se o mapa de espalhamento de Moran Global para a produtividade de milho ($t\ ha^{-1}$). Para esse cálculo utilizou-se a matriz de vizinhança dos vizinhos mais próximos para a matriz de proximidade (W). Observa-se que os três anos possuem agrupamentos semelhantes, tendo um padrão espacial do tipo *High-High* e *Low-Low*, se repetindo para vários municípios nos três mapas referentes aos anos 2011-2012, 2012-2013 e 2013-2014, o mesmo ocorre, em alguns casos, para os municípios que apresentaram características diferentes dos seus vizinhos, como nos agrupamentos do tipo: *High-Low* e *Low-High*. A região oeste em todos os anos apresentou autocorrelação alta (*High-High*) para produtividade de milho, ou seja, os municípios situados nessa localidade obtiveram uma produtividade maiores que o restante do estado, tendo um grande destaque nos mapas, principalmente nos anos de 2012-2013 e 2013-2014. Já uma parte da região norte e o litoral apresentaram autocorrelação baixa (*Low-Low*) nessas áreas, portanto, os municípios não tiveram uma boa produtividade. Para saber a relevância e significância dos padrões espaciais

encontrados na análise global é necessário aplicar o índice LISA, nos próximos parágrafos será abordado essa etapa.

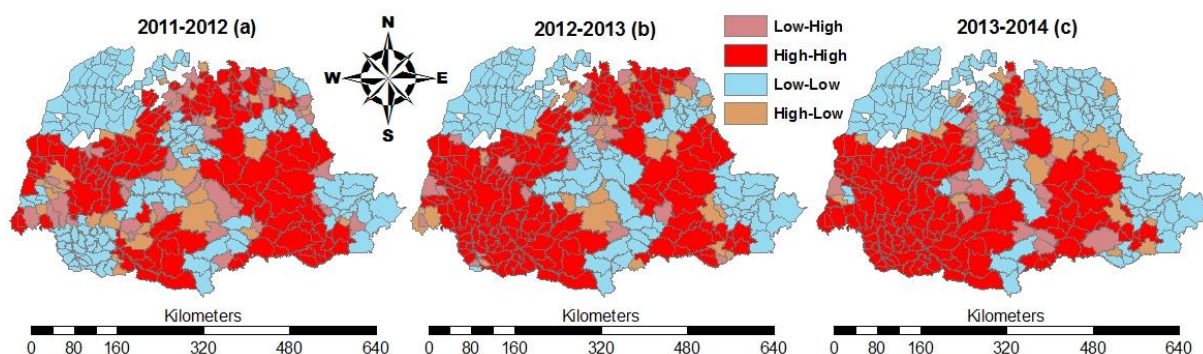


Figura 17 - Mapa de Espalhamento de Moran para a produtividade de Milho

A Figura 18 ilustra a autocorrelação espacial das variáveis climáticas (radiação solar (KJ/m^2), temperatura média ($^{\circ}\text{C}$) e precipitação pluvial (mm)) para os anos de 2011/2012, 2012/2013 e 2013/2014. Utilizou-se como critério de vizinhança para a geração do mapa a regra de vizinhança da distância entre o centroide. Observa-se que as três variáveis possuem padrões de agrupamento espacial semelhante nos anos analisados (2011/2012, 2012/2013 e 2013/2014). Por exemplo, a radiação solar (Figura 18(a, b e c)) apresentou um agrupamento do tipo *High-High* (Municípios com alta incidência de radiação solar) nas regiões sudoeste, oeste, noroeste e norte abrangendo também um pouco da região sul e central do Paraná. As regiões leste, sudeste e nordeste foram classificadas com agrupamento *Low-Low* (Municípios com baixa incidência de radiação solar).

A temperatura média ($^{\circ}\text{C}$) apresentou características semelhantes ao longo dos anos, conforme o mapa da Figura 18 (d,e,f). Verifica-se que os agrupamentos do tipo *High-High* (Municípios com alta temperatura média), *Low-Low* (Municípios com baixa temperatura média), *Low-High* (Municípios com baixa temperatura média rodeados por municípios com alta temperatura média) e *High-Low* (Municípios com alta temperatura média rodeados por municípios com baixa temperatura média) se repetiu para quase todos os municípios, portanto, os padrões de temperatura se alteraram pouco durante os anos.

A precipitação pluvial (mm) foi a variável climática que mais sofreu alterações de agrupamento para os anos estudados, observa-se nos três mapas (Figura 18(g,h,i)) que alguns municípios da região oeste, sudoeste, sul apresentaram alta quantidade de chuva nos três anos, o restante das regiões sofreu alterações em

relação a quantidade de chuvas, isso se explica, pois a precipitação pode variar muito em um determinado período de um ano para o outro, ao contrário da radiação solar e temperatura média que tendem a serem semelhantes.

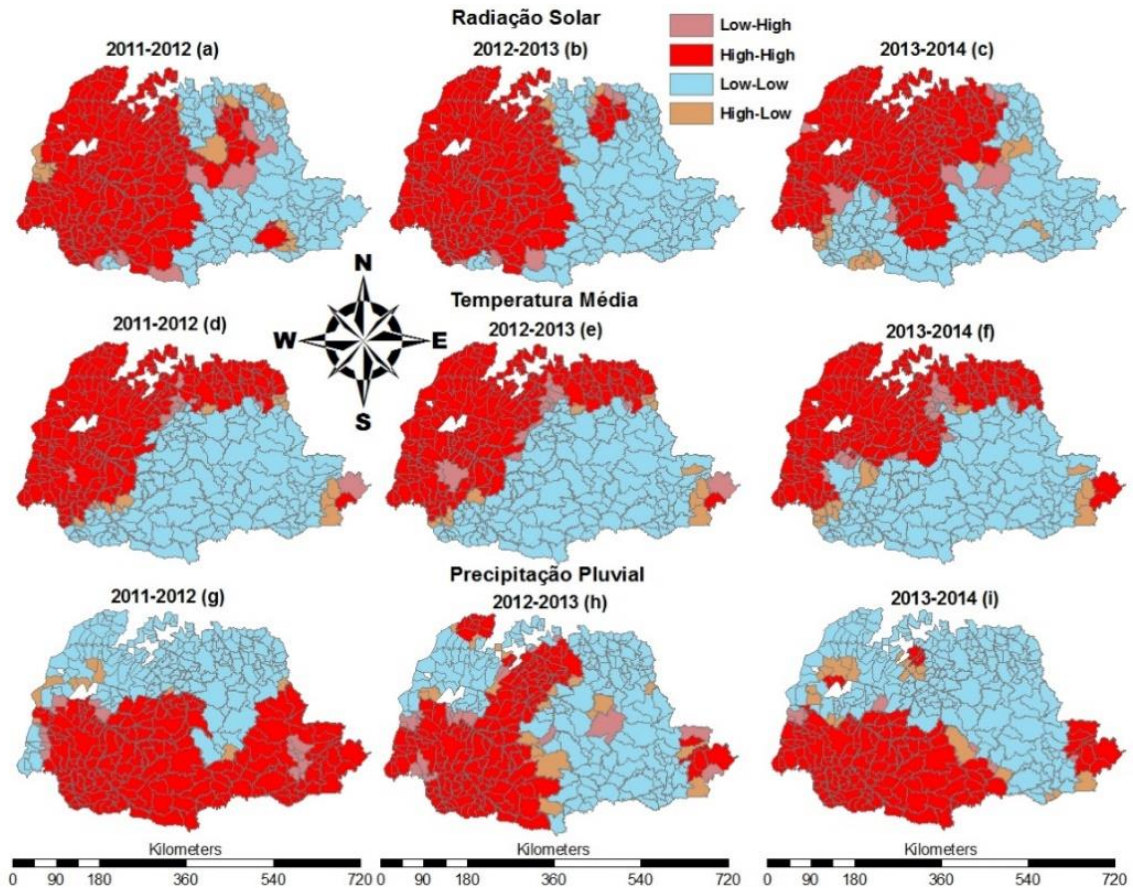


Figura 18 - Mapa de Espalhamento de Moran para as variáveis climáticas (Radiação Solar, Temperatura Média, Precipitação Pluvial)

A Figura 19 apresenta o mapa de espalhamento de Moran para as variáveis altitude e potencialidade agrícola do solo para os municípios do estado do Paraná. Utilizou-se para o cálculo do índice de Moran global a matriz de vizinhança da distância entre o centroide. Observa-se que os municípios que apresentam altitude elevada (agrupamento do tipo *High-High*) estão situados nas regiões sul e central do estado, alguns municípios da região leste e nordeste também são considerados altos. Os municípios das regiões oeste e noroeste são classificados com baixa altitude (agrupamento do tipo *Low-Low*), existem alguns casos de pontos discrepantes, ou seja, municípios com altitude elevada rodeado por municípios com baixa altitude, o inverso também ocorre. Já variável potencialidade agrícola do solo apresentou agrupamento do tipo *High-High* (Solo recomendado para a pratica agrícola) nas regiões oeste, noroeste e norte. Muitos municípios foram classificados como *Low-Low*

(Solo não recomendado para a pratica agrícola), em alguns casos, estão próximos de municípios com classificação alta para a pratica agrícola, nessas situações, houve bastante agrupamento do tipo *High-Low*, ou seja, municípios com solo adequado para a pratica agrícola próximo de municípios que possuem solo não recomendado.

Analisando os mapas da Figura 19, é possível verificar que as regiões mais altas do estado do Paraná não possuem um solo bom para a pratica agrícola, os agrupamento *High-High* ocorrem em áreas inversas nos mapas, portanto, podem existir casos em que a produtividade milho pode ser influenciada apenas pela altitude, mesmo tendo um solo com classificação baixa para a pratica agrícola, o inverso também pode ocorrer.

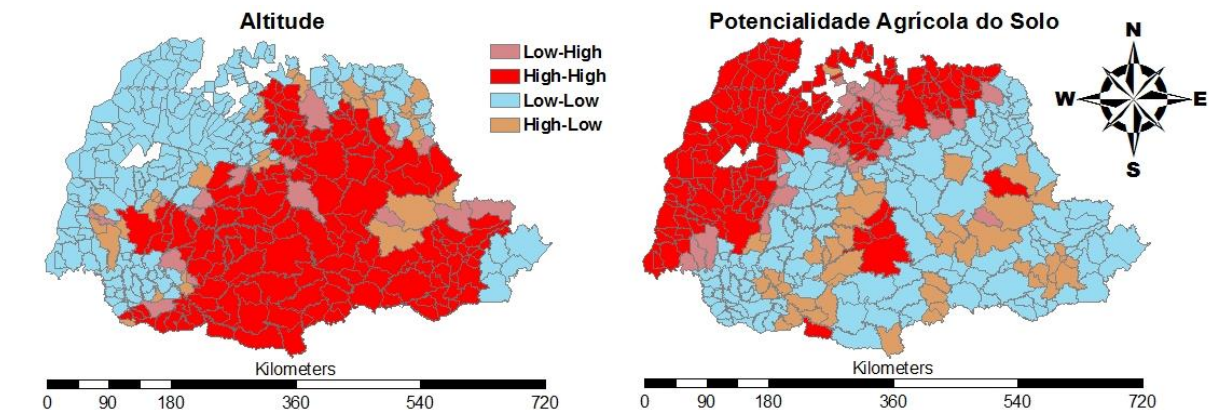


Figura 19 - Mapa de Espalhamento de Moran para as variáveis altitude e potencialidade agrícola do solo

A Figura 20 ilustra os agrupamentos locais para a produtividade de milho calculado pelo Indicador Local de Associação Espacial (LISA), sendo os mapas gerados com o critério de vizinhança da distância entre o centroide. O *Moran Map* representa o Gráfico de espalhamento de Moran, determinando os agrupamentos do tipo H-H (*High-High*), L-L (*Low-Low*), L-H (*Low-High*) e H-L (*High-Low*) que são significativos ao nível de 5%, neste caso, pode-se concluir estatisticamente que são as regiões que possuem padrões locais de associação espacial, e necessitam serem analisadas com atenção. Analisando os agrupamentos apresentados no *Moran Map*, percebe-se que para todos os anos a região noroeste e leste do estado do Paraná apresentaram agrupamento do tipo *Low-Low*, que neste caso, são cidades com baixa produtividade de milho, sendo assim, a estatística corrobora com as características de relevo (altitude próxima ao nível do mar), solo (limitado para a pratica agrícola) e chuva (quantidade baixa de chuva para região noroeste e excesso de chuva para a região leste) dos municípios localizados nesta área do estado, dificultando o bom rendimento médio da cultura do milho. Para os anos de 2011/2012 a região nordeste

não apresentou agrupamento do tipo *Low-Low* como nos anos de 2012/2013 e 2013/2014, um dos motivos que pode ter influenciado é a frequência maior de chuva que ocorreu nos anos de 2011/2012 nos municípios desta área. Os municípios pertencentes ao agrupamento do tipo *High-High* (alta produtividade de milho) estão localizados na região oeste do Paraná em todos os anos, novamente, a estatística comprova as boas características dessa região para prática da cultura do milho, pois, o solo é recomendado para atividade agrícola, a altitude em alguns municípios é considerada alta e a chuva possui frequência alta e bem distribuída. Os anos de 2011/2012 a região sudoeste apresentou baixa produtividade de milho, característica divergente que os demais anos (2012/2013 e 2013/2014) que apresentaram alta produtividade, os dados climáticos foram semelhantes (chuva, temperatura média e radiação solar) e o solo e altitude são iguais, portanto, é necessário um estudo mais detalhado nessa região nos anos de 2011/2012 para entendimento desta diferença. As regiões centrais dos três mapas não apresentaram nenhum padrão significativo de associação espacial, alguns municípios possuíram característica diferentes que seus vizinhos, por isso, foram classificados como *Low-High*, municípios com baixa produtividade de milho próximos de locais com alta produtividade, e *High-Low*, municípios com alta produtividade rodeado por regiões com baixa produtividade.

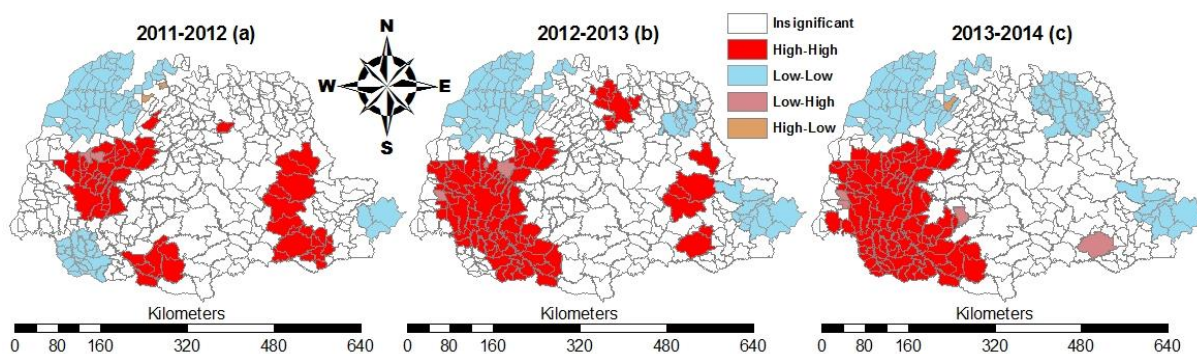


Figura 20 - Mapa de agrupamento para a produtividade de milho (*Moran Map*)

A Figura 21 representa o *Moran Map* para as variáveis climáticas (radiação solar (KJ/m^2), temperatura média ($^{\circ}\text{C}$) e precipitação pluvial (mm)) utilizadas no estudo para os anos de 2011/2012, 2012/2013 e 2013/2014. O índice LISA aplicada para esses dados utilizou a matriz de vizinhança da distância entre o centroide. Observe-se que os padrões espaciais descobertos para cada variável e ano são muito semelhantes, por exemplo, a variável temperatura média apresentou agrupamentos muito parecidos para todos os anos, já a variável radiação solar teve algumas mudanças nos agrupamentos, como é o caso do ano de 2013/2014 (c) que resultou

em um grande agrupamento *High-High* (municípios com alta incidência de radiação solar) na região norte do estado, sendo que neste local alguns municípios resultaram em um agrupamento do tipo *Low-Low* (municípios com baixa incidência de radiação solar) nos anos anteriores. Os anos de 2012/2013 (b) apresentaram um pequeno agrupamento do tipo *High-High* na região Sudoeste e 2011/2012(a) obteve um agrupamento do tipo *High-High* em quatro municípios da região oeste e em alguns municípios situados no centro até o sul do estado.

A variável precipitação pluvial também apresentou diferentes padrões de agrupamentos espaciais nos anos de 2011/2012 (g), 2012/2013 (h) e 2013/2014 (l), principalmente nas regiões aonde ocorreram pouca chuva, percebe-se que no ano de 2011/2012 toda a região norte do estado resultou em um agrupamento do tipo *Low-Low* (municípios com baixa quantidade de chuva), diferente dos outros anos, que apenas alguns municípios tiveram pouca chuva. Já alguns municípios situados nas regiões leste e sudeste apresentaram agrupamento *Low-Low* nos anos de 2012-2013 (h) e 2013/2014 (l), ao contrário dos anos de 2011/2012 (g), que não obteve agrupamento *Low-Low* nessa região. Os agrupamentos do tipo *High-High* (municípios com alta quantidade de chuva) apresentaram padrões de agrupamento semelhantes nos três mapas (2011/2012 (g), 2012/2013 (h) e 2013/2014 (l)), tendo algumas diferenças em áreas pequenas no estado, como na região leste dos anos 2011/2012 (g) e 2013/2014 (l), na região norte dos anos 2012/2013 (h) e nas regiões sul e sudeste dos anos 2011/2012 (g).

Analisando as características climáticas de cada região do Paraná, é possível observar que uma parte da região sul e litoral, apresentam uma baixa radiação solar e temperatura média, isso pode influenciar negativamente na produtividade de milho, pois o desenvolvimento do milho é melhor em temperaturas mais elevadas e com muita radiação solar, mesmo tendo uma boa frequência de chuvas. Alguns municípios da região oeste tiveram alta radiação solar, temperatura média e precipitação pluvial, isso pode ter influenciado positivamente para a produtividade de milho nessa região.

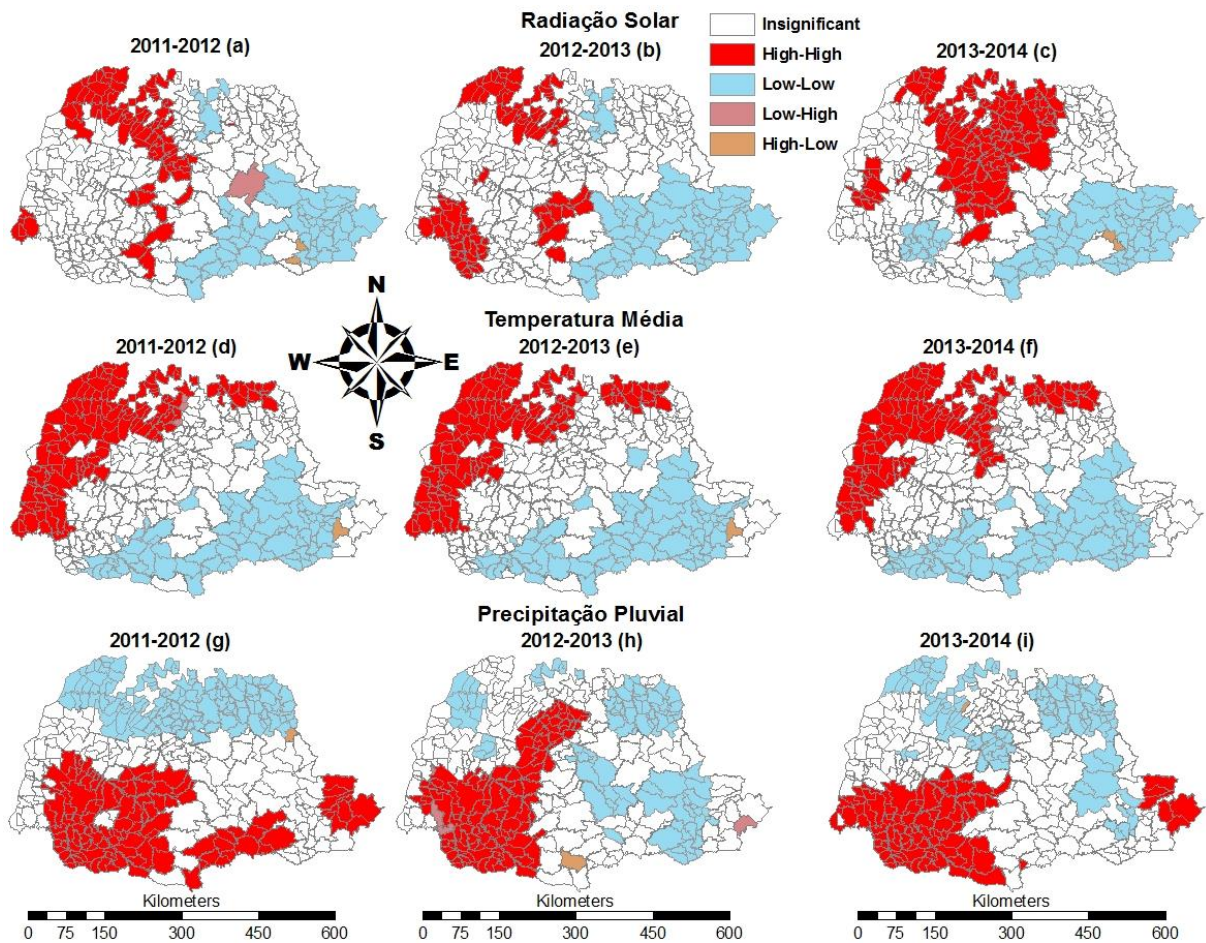


Figura 21 - Mapa de agrupamento para os dados climáticos (*Moran Map*)

A Figura 22 é o *Moran Map* para as variáveis altitude e potencialidade agrícola do solo, foi utilizado a matriz de vizinhança da distância entre o centroide para o cálculo do índice LISA. A variável altitude apresentou agrupamento do tipo *High-High* (municípios com alta altitude) na região sul e sudeste do estado, e um pequeno agrupamento do tipo *Low-Low* (municípios com baixa altitude) em algumas cidades da região leste. Observa-se que a região oeste e noroeste resultaram em um forte agrupamento do tipo *Low-Low*, ou seja, os municípios localizados nestas áreas são considerados com baixa altitude em relação aos demais.

A variável potencialidade agrícola do solo apresentou pequenos agrupamentos em alguns municípios, como nas regiões oeste e norte (agrupamento do tipo *High-High*, municípios com solo recomendado para a prática agrícola), e também nas regiões sudoeste, nordeste, leste e central do estado (agrupamento do tipo *Low-Low*, municípios com solo não recomendado para a prática agrícola), o restante dos municípios apresentaram insignificância (nenhum padrão espacial foi detectado), ou foram considerados diferentes dos seus vizinhos, com agrupamento do tipo *High-Low*

(municípios com solo recomendado para a prática agrícola rodeado por municípios com solo não adequado) e *Low-High* (municípios com solo não recomendado para a prática agrícola rodeado por municípios com solo adequado).

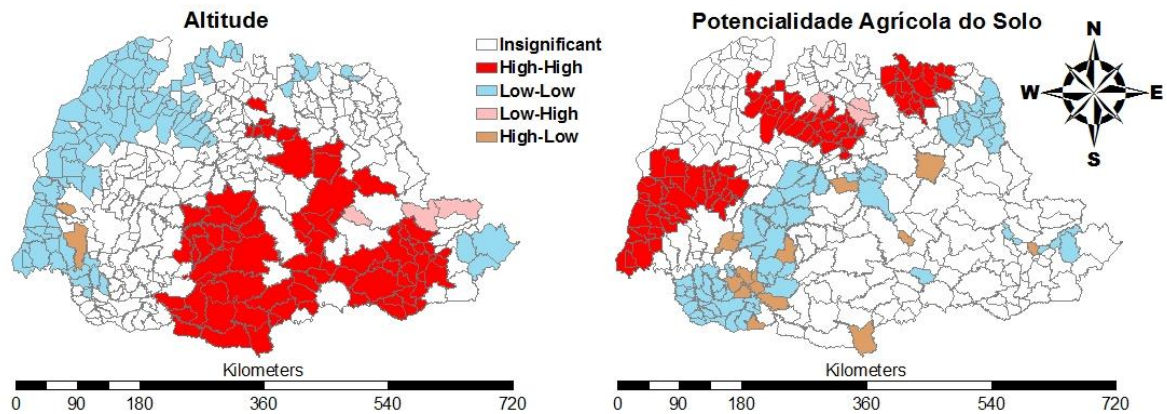


Figura 22 - Mapa de agrupamento para a altitude e potencialidade agrícola do solo (*Moran Map*)

Se for comparado o mapa de espalhamento de Moran para a produtividade de milho (Figura 20) com o mapa das variáveis climáticas (Figura 21), altitude e potencialidade agrícola do solo (Figura 22), pode-se observar que os municípios que apresentaram um padrão espacial alto para a produtividade média de milho estão localizados em regiões que tiveram temperatura média e precipitação pluvial alta. Outra característica é a potencialidade agrícola do solo que também foi classificado como boa em algumas dessas áreas, em algumas regiões apenas a altitude pode ter influenciado na produtividade. Em relação aos municípios que apresentaram baixa produtividade média do milho no norte do Paraná, pode-se observar, que nessas regiões não houve um padrão espacial significativo alto para a quantidade de chuva, esse cenário pode ter influenciado na produtividade média, além de o solo e altitude em algumas dessas cidades não serem adequados para a prática agrícola, já o litoral paranaense teve uma quantidade de chuva alto mais a produtividade de milho foi baixo, isso pode ser explicado, pois nessa região a altitude, solo, temperatura média e radiação solar, não são adequados para a prática da cultura do milho.

Para ter uma melhor explicação da produtividade média de milho, foi aplicado o índice de Moran bivariado para verificar a correlação espacial das variáveis independentes (radiação solar (KJ/m^2), temperatura média ($^{\circ}\text{C}$) e precipitação pluvial (mm), altitude (m) e potencialidade agrícola do solo) com a variável dependente (produtividade média de milho).

A correlação espacial da produtividade de milho ($t\ ha^{-1}$) com as demais variáveis do estudo, podem ser observadas na Tabela 2. De acordo com o valor calculado para o índice de Moran bivariado, as variáveis Alt, Prec e TMed apresentaram correlação espacial significativa ao nível de 1% em todos os anos (2011/2012, 2012/2013 e 2013/2014) e critérios de vizinhança (Contiguidade Queen, Distância entre o centroide e Vizinhos mais próximos), sendo que a variável TMed apresentou correlação espacial negativa e as demais positiva. A variável Psolo apresentou correlação espacial positiva e significativa ao nível de 1% e 5% para os anos de 2011/2012 e 2012/2013, com exceção no critério de proximidade espacial dos vizinhos mais próximos, no qual não houve significância ao nível de 1% ou 5%, no entanto para os anos de 2013/2014 a mesma resultou em uma correlação espacial negativa e significativa, diferente que os demais que obtiveram correlação positiva. A variável Rs apresentou correlação espacial negativa e significativa ao nível 1% para os anos de 2011/2012, no período de 2012/2013 houve correlação espacial positiva e significativa apenas para o critério de vizinhança da distância entre o centroide, nos demais não houve significância. Nos anos de 2013/2014 houve correlação espacial significativa ao nível de 1% para os critérios de vizinhança: contiguidade queen e distância entre o centroide, nos demais não ocorreu significância.

Tabela 2 - Índice de Moran Global Bivariado aplicado para a produtividade de milho com as demais variáveis do estudo

Variáveis	Índice de Moran Global Bivariado								
	Contiguidade (<i>queen</i>)			Distância entre centroides			Vizinhos mais próximos		
	2011/ 2012	2012/ 2013	2013/ 2014	2011/ 2012	2012/ 2013	2013/ 2014	2011/ 2012	2012/ 2013	2013/ 2014
PSolo	0,0841** (3,69)	0,0590* (2,39)	-0,0890** (-3,67)	0,0781** (3,81)	0,0475* (2,35)	-0,0968** (-4,70)	0,0784** (2,07)	0,0499 (1,30)	-0,1141** (-3,07)
Alt	0,3043** (11,63)	0,2225** (9,33)	0,2654** (10,27)	0,2915** (12,95)	0,2134** (9,80)	0,2532** (11,88)	0,3072** (7,60)	0,2200** (5,70)	0,2722** (6,98)
Prec	0,1752** (7,04)	0,2838** (11,96)	0,4796** (18,67)	0,1794** (8,03)	0,2842** (12,57)	0,4611** (20,26)	0,1705** (4,40)	0,2950** (7,74)	0,4712** (11,40)
TMed	-0,3049** (-11,83)	-0,2450** (-9,89)	-0,3713** (-14,22)	-0,3078** (-14,31)	-0,2486** (-11,39)	-0,3673** (-16,99)	-0,3183** (-8,46)	-0,2558** (-6,76)	-0,3746** (-9,79)
Rs	-0,1421** (-5,66)	0,0234 (0,99)	-0,0754** (-3,17)	-0,1183** (-5,65)	0,0463* (2,28)	-0,0659** (-3,20)	-0,1267** (-3,49)	0,0385 (1,09)	-0,0507 (-1,37)

Psolo: Potencialidade agrícola do Solo; Alt: Altitude (m); Prec: precipitação pluvial (mm); TMed: temperatura média do ar (°C); Rs: radiação solar global média (KJ/m²). Entre parênteses tem-se o Z-value. * Significância ao nível de 0,05. ** Significância ao nível de 0,01.

As Figuras 23 e 24 apresentam o *Moran Map* da correlação espacial das variáveis radiação solar (KJ/m²), temperatura média (°C) e precipitação pluvial (mm), altitude (m) e potencialidade agrícola do solo com a produtividade média de milho para os anos 2011/2012, 2012/2013 e 2013/2014. Para o cálculo do índice de Moran

bivariado e a construção dos mapas, foi utilizado a matriz de vizinhança da distância entre o centroide.

A Figura 23 ilustra os padrões espaciais encontrados entre a correlação das variáveis climáticas com a produtividade de milho, é possível analisar que houve algumas alterações entre os anos, a variável temperatura média (Figuras 2011/2012 (d), 2012/2013 (e), 2013/2014 (f)) foi a que apresentou os agrupamentos mais parecidos, aonde a região oeste e norte resultaram em um agrupamento do tipo *High-High* (Municípios com alta produtividade média de milho rodeado por municípios com uma alta temperatura média) e a região sul, sudeste e leste com agrupamentos do tipo *Low-Low* (Municípios com baixa produtividade média de milho rodeado por municípios com uma baixa temperatura média) e *High-Low* (Municípios com alta produtividade média de milho rodeado por municípios com uma baixa temperatura média), a região norte e noroeste obteve agrupamento do tipo *Low-High* (Municípios com baixa produtividade média de milho rodeado por municípios com uma alta temperatura média).

A radiação solar (Figuras 2011/2012 (a), 2012/2013 (b), 2013/2014 (c)) apresentou agrupamento do tipo *Low-Low* (Municípios com baixa produtividade média de milho rodeado por municípios com uma baixa radiação solar) na região leste e sudeste, *High-Low* (Municípios com alta produtividade média de milho rodeado por municípios com uma baixa radiação solar) na região sudeste, *Low-High* (Municípios com baixa produtividade média de milho rodeado por municípios com uma alta radiação solar) na região norte e noroeste, *High-High* (Municípios com alta produtividade média de milho rodeado por municípios com uma alta radiação solar) na região central, oeste e norte.

A variável precipitação pluvial (2011/2012 (g), 2012/2013 (h), 2013/2014 (i)) apresentou agrupamentos do tipo *High-High* (Municípios com alta produtividade média de milho rodeado por municípios com uma alta precipitação pluvial) na região oeste, Sudoeste para todos os anos, na região central para os anos 2012/2013 (h) e na região sudeste para os anos de 2011/2012 (g), a região norte resultou em agrupamentos do tipo *Low-Low* (Municípios com baixa produtividade média de milho rodeado por municípios com uma baixa precipitação pluvial) e *High-Low* (Municípios com alta produtividade média de milho rodeado por municípios com uma baixa precipitação pluvial) para todos os anos com destaque no período de 2011/2012(g) que preencheu todos os municípios. Os anos de 2012/2013 (h) teve agrupamentos

tipo *High-Low*, *Low-Low* e *Low-High* (Municípios com baixa produtividade média de milho rodeado por municípios com uma alta precipitação pluvial) nas regiões sudeste e leste. A região leste obteve agrupamento do tipo *Low-High* nos anos de 2011/2012 (g) e 2013/2014 (i), a região sul e oeste apresentou agrupamento do tipo *Low-High* apenas para os anos de 2011/2012 (g).

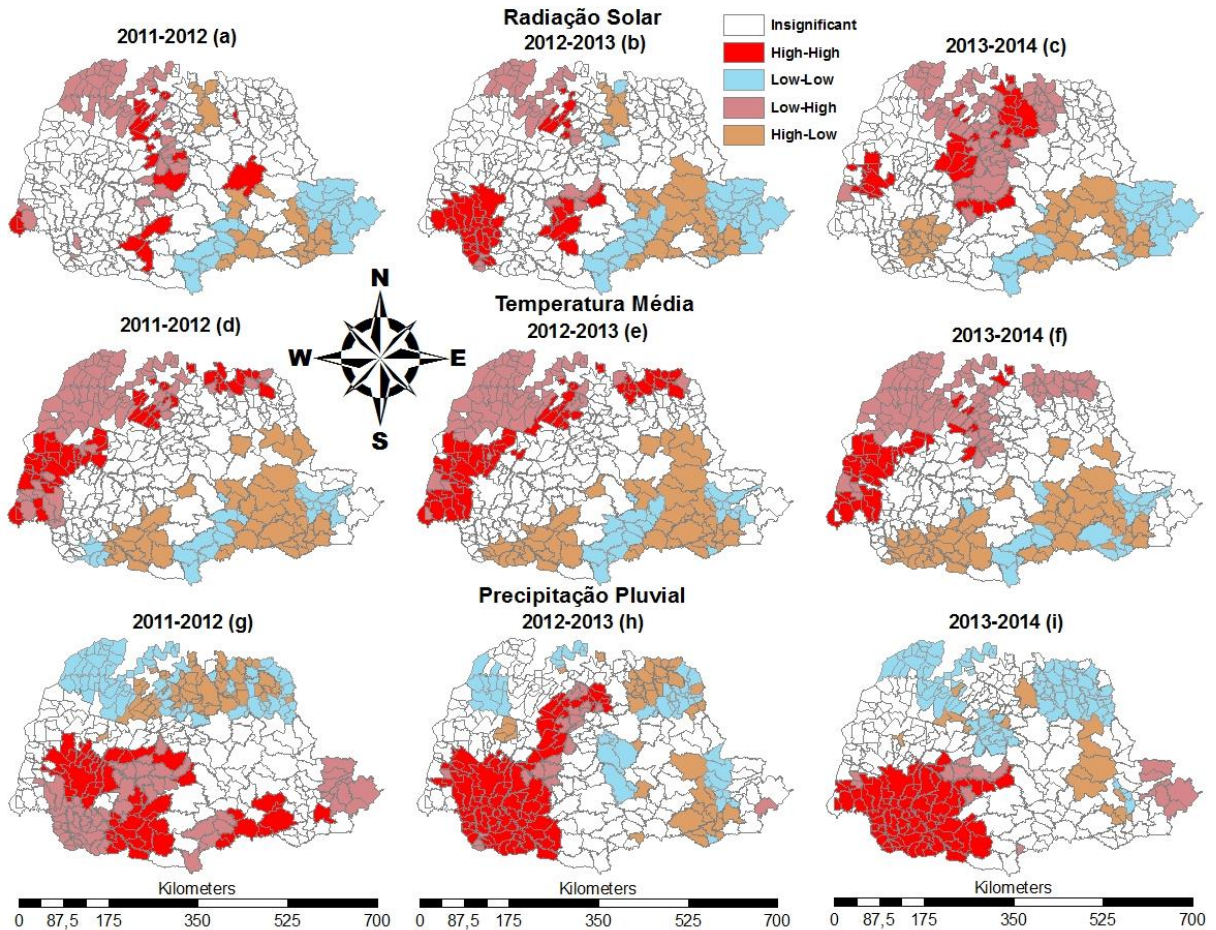


Figura 23 - Mapa da correlação espacial das variáveis climáticas com a produtividade de Milho (Moran MAP)

A figura 24 apresenta os agrupamentos encontrados entre a correlação espacial da produtividade milho com as variáveis altitude e a potencialidade agrícola do solo. Ao contrário das variáveis climáticas, pode-se perceber que os agrupamentos são semelhantes entre os anos, isso pode ser explicado pelo fato da altitude e a potencialidade agrícola do solo não sofrerem alterações em seus valores como as variáveis agrometeorológicas. A altitude apresentou agrupamento do tipo *Low-Low* (Municípios com baixa produtividade média de milho rodeado por municípios com baixa altitude) nas regiões oeste, noroeste, norte e leste em todos os anos (2011/2012 (a), 2012/2013 (b), 2013/2014 (c)). O mesmo ocorreu para as regiões sul, central, norte e sudeste apresentaram agrupamento do tipo *High-High* (Municípios com alta

produtividade média de milho rodeado por municípios com alta altitude) e *Low-High* (Municípios com baixa produtividade média de milho rodeado por municípios com alta altitude), a região oeste obteve agrupamento do tipo *High-Low* (Municípios com alta produtividade média de milho rodeado por municípios com baixa altitude).

A variável potencialidade agrícola do solo apresentou agrupamento do tipo *High-High* (Municípios com alta produtividade média de milho rodeado por municípios com alta potencialidade agrícola do solo) e *Low-High* (Municípios com baixa produtividade média de milho rodeado por municípios com alta potencialidade agrícola do solo) nas regiões oeste, norte e nordeste para todos os anos (2011/2012 (d), 2012/2013 (e), 2013/2014 (f)). As regiões centrais, sul, sudoeste e nordeste resultaram em agrupamento *High-Low* (Municípios com alta produtividade média de milho rodeado por municípios com baixa potencialidade agrícola do solo). As regiões centrais, nordeste e leste apresentaram agrupamento do tipo *Low-Low* (Municípios com baixa produtividade média de milho rodeado por municípios com baixa potencialidade agrícola do solo) em todos os anos (2011/2012 (d), 2012/2013 (e), 2013/2014 (f)), a região sudoeste obteve apenas agrupamento do tipo *Low-Low* para os anos 2011-2012(d) e 2012-2013(e) em alguns municípios.

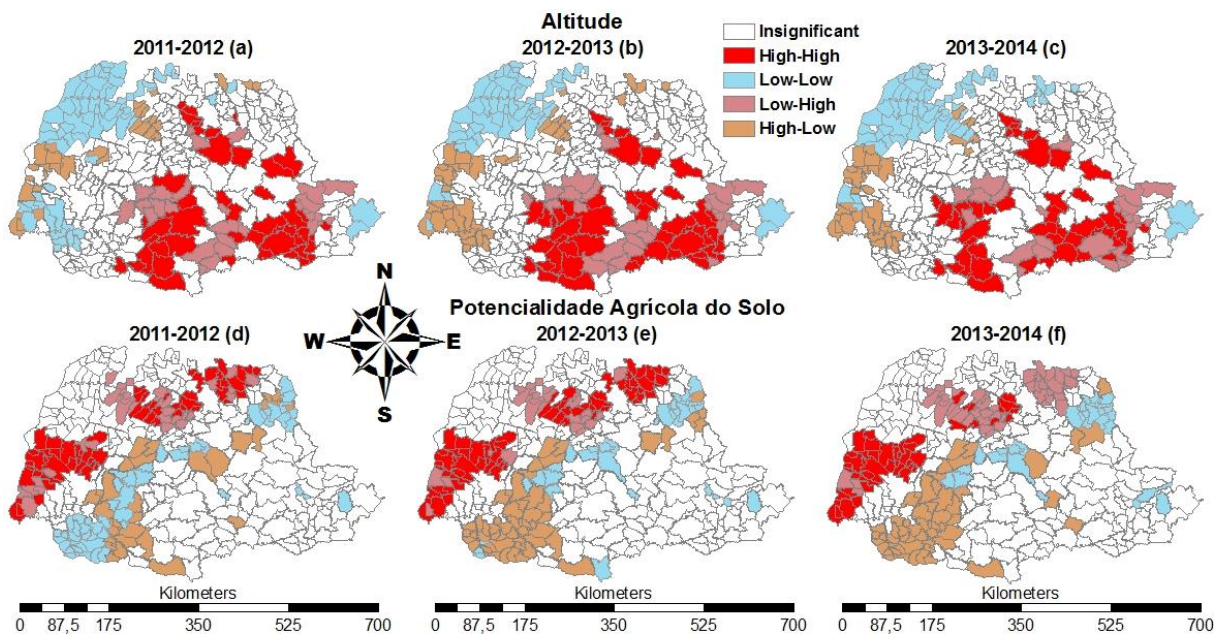


Figura 24 - Mapa da correlação espacial das variáveis altitude e potencialidade agrícola do solo com a produtividade de milho (*Moran MAP*)

Confirmado estatisticamente a correlação espacial entre as variáveis, seguiu-se em busca da explicação da produtividade média de milho com as demais variáveis.

A Tabela 3 apresenta o Diagnóstico OLS para descobrir o modelo de regressão mais recomendado para ser aplicado na predição da variável dependente \widehat{Prod} (produtividade de milho), conforme o fluxo ilustrado na Figura 2, sendo estimada pelas seguintes variáveis independentes: P_{Prec} , P_{TMed} , P_{RS} , P_{PSolo} , P_{Alt} . Percebe-se que o modelo OLS não obteve um resultado satisfatório, pois as variáveis independentes não conseguiram obter um valor aceitável para a predição da produtividade de Milho, isso pode ser explicado pelo fato desse modelo não incorporar a dependência espacial entre as variáveis, verifica-se que o valor de R^2 está menor que 50% para os três anos, e que houve uma melhora nos critério BIC e MVLFV, e também na estimativa das variáveis independentes nos modelos de regressão espacial aplicados na tabela 4.

Os testes estatísticos LM Lag, LM Error, Robust LM Lag, e Robust LM Error, indicaram para cada um dos anos o melhor modelo de regressão espacial. Para os anos de 2011/2012 e 2012/2013 foram recomendados o modelo SAR (Robust LM Error: 0,89 > 0,05; Robust LM Error: 0,051 > 0,05, respectivamente). Para os anos de 2013/2014, tanto o modelo SAR como o CAR foram recomendados, neste caso, aplicou-se os dois modelos. Esse critério da escolha do melhor modelo de regressão é abordado também nos trabalhos de Anselin (2005) e Song (2014).

Tabela 3 - Diagnóstico OLS para verificar o modelo adequado para aplicar a regressão conforme diagrama da Figura 1

Anos-Safra	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	LM-Error	LM-Lag	R LM-Error	R LM-Lag	R ²	MVLFV	BIC
2011/2012	8060,49	0,57	-320,45	0,99	258,15	2,02	0,00	0,00	0,89	0,00	20,61%	-3246,45	6528,29
2012/2013	8808,30	2,33	-503,63	3,04	229,29	1,01	0,00	0,00	0,05	0,00	28,00%	-3231	6497,38
2013/2014	6979,09	3,84	-771,29	7,07	173,66	0,17	0,00	0,00	0,00	0,00	40,55%	-3266,5	6568,38

$\hat{\beta}_0$: estimativa do coeficiente linear $\hat{\beta}_1$: estimativa do parâmetro associado à precipitação pluvial (mm); $\hat{\beta}_2$: estimativa do parâmetro associado à temperatura média do ar (°C); $\hat{\beta}_3$: estimativa do parâmetro associado à radiação solar global média (KJ/m²); $\hat{\beta}_4$: estimativa do parâmetro associado à potencialidade agrícola do solo; $\hat{\beta}_5$: estimativa do parâmetro associado à altitude (m); LM-Error: Significance Lagrange Multiplier (error); LM-Lag: Significance Lagrange Multiplier (lag); R LM-Error: Significance Robust LM (error); R LM-Lag: Significance Robust LM (lag); R²: coeficiente de determinação; MVLFV: máximo valor do logaritmo da função verossimilhança (Log likelihood); BIC: critério de informação bayesiano (Schwarz criterion).

A Tabela 4 apresenta o resultado do modelo SAR para os anos de 2011/2012, 2012/2013 e 2013/2014 sendo estimado por $\widehat{Prod} = \hat{\beta}_0 + \hat{\beta}_1 P_{PREC} + \hat{\beta}_2 P_{Tmed} + \hat{\beta}_3 P_{RS} + \hat{\beta}_4 P_{Solo} + \hat{\beta}_5 P_{Alt} + \hat{\rho}_{WProd}$, e do modelo CAR para os anos de 2013/2014, sendo explicado por $\widehat{Prod} = \hat{\beta}_0 + \hat{\beta}_1 P_{PREC} + \hat{\beta}_2 P_{Tmed} + \hat{\beta}_3 P_{RS} + \hat{\beta}_4 P_{Solo} + \hat{\beta}_5 P_{Alt} + \hat{\lambda}_{WE}$. Observando as repostas obtidas nos modelos de regressão espacial, houve uma melhora significativa nos resultados, verifica-se que o coeficiente de determinação R² para os anos de 2011/2012 (64,06%), 2012/2013 (71,94%) e 2013/2014 (CAR: 79,63% e SAR: 79,54%) obteve uma melhor predição.

Ocorreu melhora significativa no valor do critério de informação bayesiano (BIC) e no máximo valor do logaritmo da função verossimilhança (Log-likelihood), que são as medidas recomendadas em estudos envolvendo regressão espacial, de acordo com Anselin (2005) e Song (2014) o coeficiente de determinação R^2 não é apropriado em um modelo de regressão espacial, o valor listado na saída não é um verdadeiro R^2 , mais pode ser chamado de um pseudo – R^2 , no qual não pode ser comparado com o resultado obtido na regressão OLS.

Analisando a predição das variáveis independentes, a temperatura média foi a que obteve maior influência na estimativa produtividade de milho, o valor do seu coeficiente é negativo, isso indica que quanto maior for a temperatura média, menor será a produtividade de milho. As demais variáveis resultaram em um coeficiente positivo, indicando o inverso da temperatura média, ou seja, a produtividade de milho será maior na medida que as variáveis obtiverem valores maiores. Para todos os anos o coeficiente espacial auto regressivo (CAR: $\hat{\lambda}$ e SAR: $\hat{\rho}$) foi significativo ao nível de 1%, tendo seu maior valor para os anos de 2013/2014 para o modelo CAR (0,89) e menor valor para os anos de 2011/2012 (0,81).

A predição da produtividade de milho dos anos de 2011/2012 foi a que obteve os piores resultados para o coeficiente de determinação R^2 , critério de informação bayesiano BIC e o logaritmo da função verossimilhança Log-likelihood. Um dos motivos pode ser visto na Tabela 1 no valor do Índice de Moran global para a produtividade de milho. Se comparado com os demais anos é o que resultou na menor autocorrelação espacial, já o ano de 2013/2014 teve a maior autocorrelação espacial, resultando nas melhores respostas nos modelos de regressão espacial SAR e CAR.

O teste Breusch-Pagan não foi significativo ao nível de 1% em nenhum dos anos, isso confirma a rejeição da hipótese de heterocedasticidade nos dados analisados, sugerindo autocorrelação espacial. O teste Likelihood Ratio foi significativo ao nível de 1%, confirmando a forte importância da utilização do coeficiente auto regressivo espacial nos dados do estudo para todos os anos.

Tabela 4 - Modelo de Regressão Espacial aplicado para a predição da produtividade de milho

Modelo	Anos-Safra	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\frac{\text{CAR: } \hat{\lambda}}{\text{SAR: } \hat{\rho}}$	R^2	BP	LR	MVLFV	BIC
SAR	2011/2012	1204.02	0.02	-77.82	0.43	74.10	0.87	0.81	64.06%	0.89	0.00	-3128.31	6297.90
SAR	2012/2013	714.02	0.63	-107.76	0.74	83.47	0.69	0.84	72.94%	0,84	0.00	-3081.91	6205.11

CAR	2013/2014	6016.56	1.51	-314.49	3.48	63.90	0.76	0.89	79,63%	0,47	0,00	-3106.15	6247.69
SAR	2013/2014	-608.26	0.87	-139.40	1.92	65.56	0.63	0.84	79,54%	0,22	0,00	-3101.78	6244.83

β_0 : estimativa do coeficiente linear β_1 : estimativa do parâmetro associado à precipitação pluvial (mm); β_2 : estimativa do parâmetro associado à temperatura média do ar (°C); β_3 : estimativa do parâmetro associado à radiação solar global média (KJ/m²); β_4 : estimativa do parâmetro associado à potencialidade agrícola do solo; β_5 : estimativa do parâmetro associado à altitude (m); λ : estimativa do coeficiente exponencial autoregressivo (CAR); λ : estimativa do coeficiente exponencial autoregressivo (SAR); R²: coeficiente de determinação; BP: Significância do teste Breusch-Pagan; LR: Significância do teste Likelihood Ratio; MVLFV: máximo valor do logaritmo da função verossimilhança (Log likelihood); BIC: critério de informação bayesiano (Schwarz criterion).

A figura 25 apresenta o mapa dos modelos residuais padronizados da regressão espacial gerados pelos modelos SAR e CAR para os anos de 2011/2012, 2012/2013, 2013/2014.

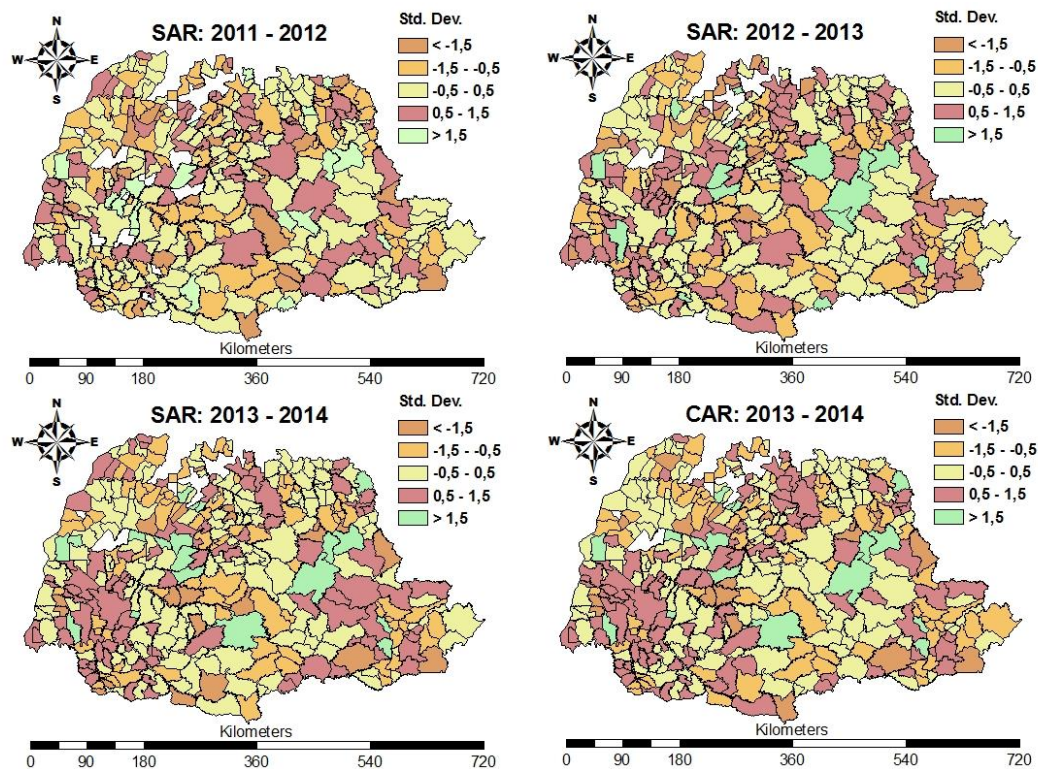


Figura 25 - Mapa dos resíduos padronizados da regressão espacial gerada pelos modelos Spatial Error e Spatial Lag, considerando o método do desvio-padrão.

Para todos os anos (Para 2011/2012: (I)=-0,0031, E(I)=-0,0028, p-value=0,498 e z-score=0,0042. Para 2012/2013: (I)=0,0020, E(I)=-0,0028, P-value=0,442 e Z-score=0,1784. Para 2013/2014 modelo SAR: (I)=0,0192, E(I)=-0,0028, P-value=0,202 e Z-score=0,8302. Para 2013/2014 modelo CAR: (I)=-0,0051, E(I)=-0,0028, P-value=0,471 e Z-score=-0,0917) o índice de Moran indicou que os resíduos estão distribuídos aleatoriamente no espaço, e o z-score não é significativo ao nível de 5%, aceitando a Hipótese nula que se refere a distribuição aleatória dos resíduos. Isso indica que a inclusão do componente WY nos modelos, praticamente, eliminou a dependência espacial, fazendo com que a inclinação da reta, que representa o índice de Moran, no diagrama de espalhamento, fosse muito pequena. Portanto, os modelos

de regressão geraram resíduos distribuídos aleatoriamente pela área de estudo, como pode ser observado na Figura 5, que representa o mapa dos resíduos padronizados, gerado pelo método do desvio-padrão resultante da aplicação do modelo SAR (2011/2012, 2012/2013 e 2013/2014) e CAR (2013/2014).

5. CONCLUSÃO

Em meio ao índice de Moran global, verificou-se a existência de autocorrelação espacial em todos os anos, variáveis e critérios de vizinhança. O método LISA (Indicador Local de Associação Espacial) apresentou padrões locais de associação no espaço, sendo destacadas as regiões do estado do Paraná que possuíram características semelhantes em todos os anos utilizados no estudo.

As variáveis independentes: temperatura média, precipitação e altitude apresentaram correlação espacial significativa ($P < 5\%$) com a produtividade de milho para todos os anos e critérios. Para os anos 2012/2013 (Matriz de vizinhança: contiguidade *queen*, vizinho mais próximo) e 2013/2014 (Matriz de vizinhança: vizinho mais próximo) a variável radiação solar não possuiu correlação espacial significativa, o mesmo ocorreu para os anos de 2013/2014 (Matriz de vizinhança: vizinho mais próximo) para a variável potencialidade agrícola do solo. Para os anos restantes e critérios de vizinhança, houve correlação espacial significativa para essas variáveis (radiação solar e potencialidade agrícola do solo).

O uso de modelos de regressão que não levam em consideração a autocorrelação espacial, pode ser inapropriado em estudos envolvendo variáveis que possuem dependência espacial. No caso da predição da produtividade de milho, os modelos de regressão espacial SAR (2011/2012, 2012/2013, 2013/2014) e CAR (2013/2014) foram os mais adequados quando comparado ao modelo de regressão OLS. O coeficiente de determinação R^2 , os critérios de informação bayesiano (BIC) e o máximo valor do logaritmo da função verossimilhança (Log-likelihood), apresentou melhora significativa na estimação da produtividade do milho quando utilizado SAR e CAR.

A estatística espacial de área mostrou-se eficiente na aplicabilidade de suas técnicas para a descoberta de padrões espaciais no presente estudo. Os métodos estatísticos espaciais aplicados neste trabalho apresentaram-se também de forma eficiente na identificação de padrões de área, na quantificação da autocorrelação espacial, na correlação espacial.

O diagnóstico OLS mostrou-se eficaz na recomendação dos modelos de regressão para a predição da produtividade de milho em todos os anos utilizados no estudo (2011/2012, 2012/2013, 2013/2014). Houve melhora significativa quando utilizado o modelo SAR (2011/2012, 2012/2013, 2013/2014) e CAR (2013/2014), pois,

os dois incorporam a autocorrelação espacial para a explicação da produtividade de milho e como os dados possuem dependência espacial, comprovou-se a eficiência desses modelos (SAR e CAR).

REFERÊNCIAS

- ALBUQUERQUE, P. H. M. **Curso de Econometria Espacial Aplicada**. Fortaleza: IPECE, 2009. 88p.
- ALMEIDA, E. S. **Curso de econometria espacial aplicada**. Piracicaba: Esalq-USP, 127 p. 2004.
- ALMEIDA, E. S. **Econometria Espacial Aplicada**. Juiz de Fora, (Mimeo), FEA/UFJF, 2006.
- ALMEIDA, Eduardo Simões de. **Econometria Espacial Aplicada**. Apostila, FEA/UFJF, 2007a.
- ALMEIDA, Marco A. S. **Análise exploratória e modelo explicativo da criminalidade no Estado de São Paulo: interação espacial (2001)**. 2007b. 85 f. Dissertação (Mestrado em Economia) - Programa de Pós Graduação em Economia da Unesp, Araraquara, São Paulo, 2007b.
- ALVES, F. D.; SILVEIRA, V. C. P. **Evolução das desigualdades regionais no Rio Grande do Sul: espaço agrário, imigração e estrutura fundiária**. Caminhos de Geografia, Uberlândia, v. 9, n. 26, p. 1-15, jul. 2008.
- ANDREOLI, Cleverson V.; ANDREOLI, Fabiana N. de; JUSTI, Jorge J. **Formação e características dos solos para o entendimento de sua importância agrícola e ambiental**. Coleção Agrinho, 2013.
- ANSELIN, L. **Exploring Spatial Data with GeoDa: A Workbook**. Center for Spatially Integrated Social Science, 2005.
- ANSELIN, L. **Under the hood issues in the specification and interpretation of spatial regression models**. Agricultural Economics, v. 27, n. 3, p. 247–267, 2002.
- ANSELIN, L., BERA, A., FLORAX, R., YOON, M. **Simple Diagnostics Tests for Spatial Dependence**, *Regional Science and Urban Economics* 26, 77-104. 1996
- ANSELIN, L. **SpaceStat Tutorial**. 1992. Disponível em: <<http://www.spacestat.com>>. Acesso em: 10 set. 2012.
- ANSELIN L. **Exploratory Spatial Data Analysis and Geographic Information Systems**. DOSES/Eurostat Workshop on New Tools for Spatial Analysis, Lisboa, Portugal, (West Virginia University, Regional Research Institute, Research Paper 9329), 1993.
- ANSELIN, L. **Local indicators of spatial association - LISA**. Geographical Analysis, Ohio/USA, 27:91-115, 1995.

ANSELIN, L. **Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences**. 1992. Disponível em <http://www.ncgia.ucsb.edu/Publications/Tech_Reports/92/92-10.PDF>. Acesso: 16 de set 2015.

ANSELIN, L. **The Moran scatterplot as ESDA tool to assess local instability in spatial association**. In: M. Fisher, H. J. Scholten and D. Unwin (ed). *Spatial Analytical Perspectives on GIS*. London, Taylor & Francis, p111-126, 1996.

ANSELIN, L. **Exploratory spatial data analysis and geographic information systems**. In: PAINHO, M. (Ed.) *New tools for spatial analysis: proceedings of the workshop*. Luxemburgo: EuroStat. p.45-54. 1994.

ANSELIN, L. **Spatial Econometrics: Methods and Models**, Dordrecht, Kluwer Academic Publishers. 1988.

ANSELIN L.; SYABRI, I.; SMIRNOV, O. **Visualizing multivariate spatial correlation with dynamically linked windows**. Urbana-Champaign: Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, 2004. 13 p. Disponível em: <<http://www.real.illinois.edu/d-paper/01/01-t-10.pdf>>. Acesso em: 14 out. 2015.

ARAÚJO, E. C. **Análise da produtividade da soja associada a fatores agrometeorológicos, por meio de estatística espacial de área, na região oeste do estado do paraná**. 2012.

ARAÚJO, E. C., URIBE-OPAZO, M. A., & JOHANN, J. A. **Modelo de regressão espacial para estimativa da produtividade da soja associada a variáveis agrometeorológicas na região oeste do estado do paraná**. 2013

ASSUNÇÃO, R. M. **Estatística Espacial com aplicações em Epidemiologia**. Economia, Sociologia. 7 Escola de Modelos de Regressão, São Carlos, SP. 2001.

AYOADE, J.O. **Introdução a climatologia para os trópicos**. São Paulo: Difel, 1986, 332p.

BARBOSA, J. V. A. **Fisiologia do milho**. In: EMPRESA BRASILEIRA DE ASSISTÊNCIA TÉCNICA E EXTENSÃO RURAL – EMBRATER. *Cultura do milho*. Brasília: EMBRATER, 1983. p. 7-12.

BARROS, J. F. C, CALADO J. G. **A Cultura do Milho**. Universidade de Évora, 2014.

BAILEY, T.C. **A review of statistical spatial analysis in geographical information systems**. Em: Fotheringham, S., Rogerson, P. *Spatial Analysis and GIS*. London: Taylor and Francis. p.13-44. 1994.

BAILEY, T. C., GATRELL, A. C. **Interactive spatial data analysis**. Essex: Longman Scientific and Technical, 1995.

BRAGA, Altemir da S. et al. **Estudo de Dependência Espacial utilizando Análise de Dados de Área Aplicada na Mesorregião Metropolitana de Belo Horizonte por meio do Indicador Econômico PIB 19ª SINAPE**, 2010.

BRASIL. Ministério da Saúde. **Abordagens espaciais na saúde pública / Ministério da Saúde**, Fundação Oswaldo Cruz; Simone M. Santos, Christovam Barcellos, organizadores. -Brasília : Ministério da Saúde, 2007. 136 p. - (Série B. Textos Básicos de Saúde) (Série Capacitação e Atualização em Geoprocessamento em Saúde; 3)

BURROUGH, P. A.; MCDONNELL, R.A. **Principles of geographical information systems**. Oxford: Oxford University Press, 1998.

BOHÓRQUEZ, L.; GÓMEZ, I.; SANTA, F. **Methodology for the discrimination of areas affected by forest fires using satellite images and spatial statistics**. Procedia Environmental Sciences, v. 7, p. 389–394, 2011.

CÂMARA, G.; CASANOVA M.; HEMERLY A.; MAGALHÃES G. e MEDEIROS C. **Anatomia de Sistemas de Informações Geográficas**. Instituto de Computação, UNICAMP, Campinas, São Paulo, 1996.

CÂMARA, G., MEDEIROS, J.S. **Geoprocessamento para projetos ambientais**. São José dos Campos, 1996. p39. (Relatório do INPE)

CÂMARA, G., CARVALHO, M. S., CRUZ, O. G., & CORREIA, V. **ANÁLISE ESPACIAL DE ÁREAS**, 2005. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/analise/>, acesso em 27 de outubro de 2015.

CÂMARA, G., Monteiro, A. M., FUCKS, S. D., & CARVALHO, M. S. **ANÁLISE ESPACIAL DE DADOS GEOGRÁFICOS**. Brasília: EMBRAPA. 2014. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/analise/>, acesso em 27 de outubro de 2015.

CÂMARA, G.; MONTEIRO, A M.V.; CARVALHO, M. S.; DRUCK, S (2002) **Análise Espacial de dados Geográficos**. 2a edição (online), disponível <http://www.dpi.inpe.br/gilberto/livro/analise/>, acesso em 10 de set. 2015

CAMARGO, A.P.; PINTO, H.S.; PEDRO JR., M.J.; et al. **Aptidão climática de culturas agrícolas**: In: São Paulo – Secretaria da Agricultura. Zoneamento agrícola do Estado de São Paulo. São Paulo, SA-SP, 1974, V.1, p.109-149.

CAMARGO, A.P.; ALFONSI, R.R.; PINTO, H.S.; CHIARINI, J.V. **Zoneamento da aptidão climática para culturas comerciais em zonas de cerrado**. In: Simpósio sobre o Cerrado, 6. São Paulo, Ed. da USP, 1977. p.89-105.

CARMO NETO, O. V.; LUI, J. J.;PIRES, L. P. M.; CANCELLIER, L. L.; PELUZIO, J. M. **Desempenho de genótipos de cana-de-açúcar em três cortes na região sul do estado do tocantins**. Revista verde de agroecologia e desenvolvimento sustentável, v. 6, n.4, p.19-27, 2011.

CAMURÇA, Paulo J. **Cálculo da precipitação média de forma automática pelo método de thiesen e avaliação da precipitação na bacia do jaguaribe**. 2011. 76 f. Monografia - Licenciatura em Física do Centro de Ciências e Tecnologia da Universidade Estadual do Ceará, Ceará, 2011.

CARVALHO, M. S. **Aplicação de métodos de análise espacial na caracterização de áreas de risco a saúde**. 149f.Tese. Universidade Federal do Rio de Janeiro. Programa de Pós-Graduação em Ciências em Engenharia Biomédica. 1997.

CELEBIOGLU, F.; DALL'ERBA, S. **Spatial Disparities across the regions of Turkey: an exploratory spatial data analysis**, The Annals of Regional Sciences. v.45, n.2, p.379-400, 2009.

CLEMENT, G.; LAROUCHE, C.; GOUIN, D.; MORIN, P. & KUCERA, H. **OGDI: toward interoperability among geoespacial databases**. SIGMOD Records, 26:18-23, 1997.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da Safra Brasileira - Grãos**. Observatório Agrícola, v. 2, n. 5, p. 1–116, 2015.

CRESSIE, N. **Statistics for Spatial Data Chichester**, John Wiley. 1992.

CRESSIE, N. **Bayesian smoothing of rates in small geographic areas**. Journal of Regional Science. 35, 659-673. 1995.

CRESSIE, N. **Change of support and the modifiable areal unit problem**. Geographical System. 3, 159-180. 1996.

CRUZ, José Carlos et al. **Manejo da Cultura do Milho**. Ministério da Agricultura, Pecuária e Abastecimento. Sete Lagoas: Embrapa, 2006.

CRUZ, José Carlos et al. **Cultivo do milho**. Sete Lagoas: Embrapa Milho e Sorgo, 2008. (Sistema de Produção, 2). Disponível em: <http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Milho/CultivodoMilho_4ed/manejomilho.htm>. Acesso em: 15 set. 2015.

CRUZ, José Carlos et al. **Cultivo do milho**. 6 ed. Sete Lagoas: Embrapa Milho e Sorgo, 2010. (Sistema de Produção, 1). Disponível em: <http://www.cnpms.embrapa.br/publicacoes/milho_6_ed/referencias.htm>. Acesso em: 15 set. 2015.

DUARTE, J. de O.; CRUZ, J. C.; GARCIA, J. C.; MATTOSO, M. J. **Economia da produção**. In: CRUZ, J. C. (Ed.). Cultivo do milho. 6. ed. Sete Lagoas: Embrapa Milho e Sorgo, 2010. (Embrapa Milho e Sorgo. Sistema de produção, 1). Disponível em: <http://www.cnpms.embrapa.br/publicacoes/milho_6_ed/economia.htm>. Acesso em: 12 ago. 2015.

DANTAS, A.; TACO P. W. G. e YAMASHITA, Y. **Sistema de Informação geográfica em Transportes**. Anais do X Congresso de Pesquisa e Ensino em Transportes, ANPET, Brasília, DF, Vol. 1, p 211-219, 1996.

DEMERS, M. N. **Fundamentals of Geographic Information Systems**. Ed. John Wiley & Sons, Inc., New York, USA. 1997.

DEMERS, M. N. **Fundamentals of geographic information systems**. New Jersey, J. Wiley, 2009.

DRUCK, S.; CARVALHO, M.S.; CÂMARA, G.; MONTEIRO, A.V.M. (eds) **Análise Espacial de Dados Geográficos**. Brasília, EMBRAPA, 2004 (ISBN: 85-7383-260-6).

DURÃES, F.O.M.; CORRÊA, L.A.; CRUZ, J.C.; OLIVEIRA, A.C.; MAGALHÃES, .C.; SHANAHAN, J.F. **Estabilidade do rendimento e tolerância a estresses em genótipos de milho cultivados nas regiões Centro e Sul do Brasil**. In: CONGRESSO NACIONAL DE MILHO E SORGO, 24., 2002, Florianópolis, SC. Meio ambiente e a nova agenda para o agronegócio de milho e sorgo. Resumos expandidos. Sete Lagoas: ABMS/Embrapa Milho e Sorgo/EPAGRI, 2002.

EMBRAPA. **Cultivo do Milho**. Embrapa Milho e Sorgo, Sistema de Produção 1, ISSN 1679-012X Versão Eletrônica - 7ª edição, Setembro, 2011. Disponível em: <http://www.cnpms.embrapa.br/publicacoes/milho_7_ed/manejomilho.htm>. Acesso em: 02 set. 2016.

EMBRAPA. **Sistema Brasileiro de Classificação de Solos**. 2. ed. Rio de Janeiro, 2009

ESRI. **Arc GIS Spatial Analyst**. 2016. Disponível em: <<http://www.esri.com/software/arcgis/extensions/spatialanalyst/surface.html>>. Acesso em: 9 set. 2016.

FANCELLI, A. L.; DOURADO NETO, D. **Produção de milho**. Agropecuária Guaíba, 2000.

FOOK, K. D. **INTEGRAÇÃO DA ESTATÍSTICA ESPACIAL EM AMBIENTES GIS**.2005.

FOTHERINGHAM, A.S.; BRUNSDON, C.; CHARLTON, M. **Quantitative Geography; Perspectives of Spatial Data Analysis**. London Sage Publications. Ch4,.2000.288p.

GAMA, R. G.; STRAUCH, J. C. M. **Construção e análise de indicadores de desenvolvimento sustentável para a bacia hidrográfica do Paraíba do Sul**. In: ENCONTRO NACIONAL DE GEOGRÁFOS: O espaço não pára, por uma AGB em movimento, XV. Anais... São Paulo: AGB, 2008.

GEHLKE, C. E; BIEHL, K. **Certain effects of grouping upon the size of the correlation coefficient in census tract material**. Journal of the American Statistical Association Supplement, 29, 169-170, 1934.

GEODA. (s.d.). GeoDa. 2016. Disponível em GeoDa: <https://geodacenter.asu.edu/>. Acesso em: 3 set. 2016

GONÇALVES, E. **A distribuição espacial da atividade inovadora brasileira: uma análise exploratória**. Belo Horizonte: UFMG/Cedeplar. 33p. 2005.

GOODCHILD, M. , ANSELIN, L., APPELBAUM, R. AND HARTHORN, B. **Towards spatially integrated social science**. International Regional Science Review 23, 139-159. 2000.

HADDAD, E. A.; PIMENTEL, E. **Análise da Distribuição Espacial da Renda no Estado de Minas Gerais: Uma Abordagem Setorial**. Texto para discussão, NEREUS. São Paulo. 2004.

HAINING, R., **Spatial Data Analysis in the Social and Environmental Sciences**, Cambridge University Press. 1997 INSTITUTO NACIONAL DE ESTATÍSTICA. Cartografia da pobreza. Cabo Verde. 2004

HOOGENBOOM, G. **Contribution of agrometeorology to the simulation of crop production and its application**. Agricultural and Forest Meteorology, 103, 137-157, 2000.

IAPAR. **Instituto Agrônomo do Paraná**. 2016.

IBAMA - Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis. **GEO BRASIL 2002 Perspectivas do Meio Ambiente no Brasil**. Brasília: Edições IBAMA, 2002.

IBGE - Instituto Brasileiro de Geografia e Estatísticas. **Manual Técnico de Pedologia**. 2. ed. Rio de Janeiro, 2007.

IBGE - Instituto Brasileiro de Geografia e Estatísticas. **Potencialidade Agrícola dos Solos**. 2002. Disponível em <<http://mapas.ibge.gov.br/tematicos/solos>>. Acesso em: 21 nov. de 2015.

IBGE. **Instituto Brasileiro de Geografia e Estatística**. 2016.

INMET. **Manual de Observações Meteorológicas**. 3a ed., Brasília: Ministério da Agricultura e do Abastecimento, 1999, 84p.

INMET. **Instituto Nacional de Meteorologia**. 2016.

IPARDES. **Instituto Paranaense de Desenvolvimento Econômico e Social**. 2016.

KREMPI, Ana P. **EXPLORANDO RECURSOS DE ESTATÍSTICA ESPACIAL PARA ANÁLISE DA ACESSIBILIDADE DA CIDADE DE BAURU**. 2004. 80 f. Dissertação (Mestrado em Engenharia Civil) - Escola de Engenharia de São Carlos, Universidade de São Paulo USP, São Carlos, 2004.

LIMA, W. P. **Precipitação. Notas de aula**: Piracicaba-SP, 2000.

LEPSCH, I. F. **Formação e Conservação dos Solos**. São Paulo: Oficinas de Textos, 2002.

LEPSCH, I.F. **Influência dos fatores edáficos na produção**. In: CASTRO, P.R.C.; FERREIRA, S.O.; YAMADA, T. (Coord.) *Ecofisiologia da produção*. Piracicaba: Associação Brasileira para Pesquisa da Potassa e do Fosfato, 1987. p.8398.

MAGALHÃES, P. C.; DURÃES, F. O. M. **Fisiologia da produção de milho**. Sete Lagoas: Embrapa Milho e Sorgo, 2006. 10 p. Embrapa Milho e Sorgo. (Circular Técnica, 76).

MAGUIRE, D.; GOODCHILD, M.F.; RHIND, D.W. **Geographic information systems: principles, and applications**. Harlow. Longman Scientific and Technical, 1997.

MAVI, H.S.; TUPPER, G.J. **Agrometeorology – Principles and application of climate studies in agriculture**. New York: Food Products Press. 2004. 364p.

MENESES, H. B. **Interface Lógica em ambiente SIG para Bases de Dados de Sistemas Centralizados de Controle do Tráfego Urbano em Tempo Real**. 182 fl. Dissertação de Mestrado, Universidade Federal do Ceará, Fortaleza, CE, 2003.

MELO, A. S.; HEPP, L. U. **Ferramentas estatísticas para análises de dados provenientes de biomonitoramento**. *Oecologia Brasiliensis* 12(3):463-486, 2008.

MILLER, H. J. **Potential contribution of spatial analysis to Geographic Information System for Transpotation (GIS-T)**. *Geographical Analysis*, 31, 373-399. 1999.

MONASTERIO, L. M.; ÁVILA, R. P. **Uma análise espacial do crescimento econômico do Rio Grande do Sul (1939-2001)**. *Revista Economia*, Brasília, v. 5, n. 2, p. 269-296, jul./dez. 2004.

MONASTERIO, L. M.; SALVO, M.; DAMÉ, O. M. **Estrutura espacial das aglomerações e determinação dos salários industriais no Rio Grande do Sul**. *Ensaio FEE*, v. 28, p. 801-824, 2008.

MORAN, P.A.P. **The interpretation of Statistical Maps**. *Biometrika* 35, 255-260. 1947

MOURA, A. C. M.. **Geoprocessamento na gestão e planejamento urbano**. *Belo Horizonte* : Ed. Da autora, 2003, 294p

NEVES, M. C. et al. **Análise exploratória espacial de dados socioeconômicos de São Paulo**. In: GISBRASIL 2000, Salvador (CD-ROM).

NEVES, M. C; RAMOS, F. R; CAMARGO, E. C. G; CÂMARA, G; MONTEIRO, A. M. **Análise exploratória espacial de dados sócios-econômicos de São Paulo**, Campinas (SP): Embrapa Meio Ambiente e Instituto Nacional de Pesquisa Espacial - INPE, 2001.

OSTENSEN, O. **Mapping the future of geomatics**, ISO/TC 211. Paris, Prentice Hall International, 1995. p.49. (Bulletin, 1)

PEDRAZZI, J.A. **FACENS – Hidrologia Aplicada**, 1999. Disponível em: <<http://www.facens.br/site/alunos/download/hidrologia>>. Acesso em 05 set. 2015

PEREIRA, A.R.; ANGELOCCI, L.R.; SENTELHAS, P.C. **Agrometeorologia – fundamentos e aplicações práticas**. Guaíba: Ed. Agropecuária. 2002. 478p

PEREIRA FILHO, I. A.; CRUZ, J. C. **Produção de milho em sistema irrigado**. In: RESENDE, M.; ALBUQUERQUE, P. E. P.; COUTO, L. A cultura do milho irrigado. Brasília: EMBRAPA, 2003.

PEROBELLI, F. S. et al. **Produtividade do setor agrícola brasileiro, 1991-2003: uma análise espacial**. Nova Economia, Belo Horizonte, v. 17, n. 1, p. 65-91, jan./abr. 2007.

POSTGRESQL. **PostgreSQL**. 2016. Disponível em: <<https://www.postgresql.org/about/>>. Acesso em: 06 julho de 2016.

POSTGIS. **PostGIS 2.3.0 dev Manual**. 2016. Disponível em <<http://postgis.net/docs/manual-dev/postgis-br.html>>. Acesso em: 06 julho de 2016.

PINA, M. F. **Modelagem e Estruturação de Dados Não-Gráficos em Ambiente de Sistemas de Informação Geográfica: estudo de caso na área de saúde pública**. p.5 a 21. Dissertação de Mestrado, Rio de Janeiro: Instituto Militar de Engenharia, 1994.

PINTO, N. L. de S. et al. **Hidrologia Básica**. 1ª ed. São Paulo: Edgard Blücher, 1976. 279 p.

QUEIROZ, G. R.; FERREIRA, K. R. **SGBDs com extensões espaciais**. In: Casanova, M. A.; Câmara, G.; Davis-Jr, C. A.; Vinhas, L.; Queiroz, G. R. (Org.). Bancos de Dados Geográficos. Curitiba: Editora MundoGEO, 2005. cap. 8, p. 267-303.

R. (2013). disponível em <http://www.r-project.org/>, acesso em 10 de 10 de 2014.

RAMANATHAN, R. **Introductory econometrics with applications**. San Diego: The Dryden Press. 664 p. 1998.

REIS, A. C.; CASTRO, S. G. **Incorporação da Dependência Espacial na Identificação do Cluster Potencial Agropecuário-Florestal no Estado do Pará**. Apresentação Oral-Desenvolvimento Rural, Territorial e regional. UNIVERSIDADE DA AMAZÔNIA - UNAMA, BELÉM - PA - BRASIL, 2008.

RIGOTTI, J. I. R.; VASCONCELLOS, I. R. P. **Uma análise espacial exploratória dos fluxos populacionais brasileiros nos períodos 1986-1991 e 1995-2000**. Anais do IV Encontro Nacional Sobre Migrações da Abep. Rio de Janeiro, 2005.

ROBINSON, W. S. **Ecological correlation and the behavior of individuals**. American Sociological Review, 15, 351-357. 1950.

ROCHA, M. M. **Modelagem da Dispersão de Vetores Biológicos com emprego da Estatística Espacial**. 93f. Dissertação de Mestrado, Instituto Militar de Engenharia-IME, Rio de Janeiro, 2004.

ROSA, R. **Análise Espacial em Geografia**. Spatial Analysis in Geography. Universidade Federal de Uberlândia. Revista da Ampege, v.7, n.1, p. 275-289, out. 2011.

SANS, L. M. A.; GUIMARÃES, D. P. Zoneamento Agrícola. In: CRUZ, J. C. **Cultivo do milho**. Sete Lagoas: Embrapa Milho e Sorgo, 2011. (Embrapa Milho e Sorgo. Sistema de Produção, 1). Disponível em: <http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Milho/CultivodoMilho_7ed/zoneamento.htm>. Acesso em: 14 jul. 2015.

SANTOS, R. F.; CARLESSO, R. **Enrolamento e expansão das folhas de milho submetidas a déficit hídrico em diferentes solos**. Revista Brasileira de Engenharia Agrícola e Ambiental, v. 3, n. 1, p. 1-6, 1998.

SEAB - Secretaria de Estado da Agricultura e do Abastecimento. **Milho análise da conjuntura**. 2015. Disponível em <<http://www.agricultura.pr.gov.br/modules/conteudo/conteudo.php?conteudo=32>>. Acesso em: 21 nov. de 2015.

SENTELHA, P. C., MONTEIRO J. E. B. de A. **Agrometeorologia dos Cultivos – O Fator Meteorológico na Produção Agrícola**. INPE, 2009.

SERRANO, R. M.; VALCARCE, E. V. **Técnicas econométricas para el tratamiento de datos espaciales: la econometria espacial**. Barcelona, Edicions Universita de Barcelona. 2000.

SHEKKAR, S.; CHAWLA, S. **Spatial databases - a tour**. Upper Saddle River, NJ, USA, Prentice-Hall, 2003.

SIMEPAR. **Sistema Meteorológico do Paraná**. 2016.

SILVA, J.SV. **Análise multivariada em zoneamento para planejamento ambiental**. Estudo de caso: Bacia Hidrográfica do Rio Taquari MS/MT. 2003. 307 f. Tese (Doutorado em Engenharia Agrícola). Universidade Estadual de Campinas, Campinas, 2003.

SILVA, A. T. A. da. **Aspectos meteorológicos e balanço hídrico em um aterro de resíduos sólidos urbanos**. 141f. Dissertação de mestrado. COPPE/UFRJ, 2008.

SILVA, A. M. A.; RESENDE, G. M. **Crescimento econômico comparado dos municípios alagoanos e mineiros: uma análise espacial**. Brasília: Ipea. 31p. 2006.

SILVEIRA, Thyago de A. **Geoprocessamento aplicado ao planejamento dos transportes urbanos**. João Pessoa, p.20-22, 2006.

SONG, J. et al. **The relationships between landscape compositions and land surface temperature**: Quantifying their resolution sensitivity with spatial regression models. *Landscape and Urban Planning*, v. 123, p. 145–157, 2014.

TAIZ, L.; ZEIGER, E. **Fisiologia vegetal**. Tradução E. R. Santarém et al. 4. ed. Porto Alegre, Artmed, 2009. 848p.

TEIXEIRA, A.; MATIAS, L.; NOAL, R.; MORETTI E. **Qual a Melhor Definição de SIG**. Fator GIS - A Revista do Geoprocessamento, Ano 2, Nº 11, P. 20-22, Dezembro-1995.

TEIXEIRA G. L. **Uso de Dados Censitários para Identificação de Zonas Homogêneas para Planejamento de Transportes Utilizando Estatística Espacial**. Dissertação de Mestrado. Departamento de Engenharia Civil e Ambiental - Universidade de Brasília - DF, 155p. 2003.

TOBLER, W. R. **A computer movie simulating urban growth in the Detroit region**. *Economic Geograpy*, 46, 234-240, 1970.

TOMLIN, D. **Geographic information systems and cartographic modeling**. New Jersey, Prentice Hall, Englewood Cliffs, 1990.

TOMLINSON, R. F. **Thinking about GIS**: Geographic information system planning for managers. ESRI Press, 2005.

UNWIN A.; UNWIN D. **Spatial Data Analysis with Local Statistics**. *Journal of the Royal Statistical Society: Series D (The Statistician)*, Londres/Inglaterra, v.47, n.3, p.415–421, 1998

USDA - United States Department of Agricultura. **Keys to Soil Taxonomy**. 11. ed., 2012.

XAVIER da SILVA, J.; ZAIDAN, R.T. **Geoprocessamento & análise ambiental: aplicações**. Rio de Janeiro: Bertrand Brasil, 2004.

VIANELLO, R. L.; ALVES, A. R. **Meteorologia básica e aplicações**. 3. ed. Viçosa: Ed. da Universidade Federal de Viçosa, 2001, 449p.

WRUBLACK, S. C. et al. **Spatial distribution of Canola culture in the State of Paraná (Brazil) between the agricultural years of 2005 and 2009.** *Ciencia e investigación agraria*, v. 40, n. 3, p. 523–535., 2013.

ZHANG, C. et al. **Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland.** *Science of the Total Environment*, v. 398, n. 1–3, p. 212–221, 2008.