

**UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ – UNIOESTE**  
**CAMPUS DE CASCAVEL**  
**CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AGRÍCOLA**

**EDILZA MARTINS DA SILVA**

**PRODUTIVIDADE DA SOJA ASSOCIADA A VARIÁVEIS AGROMETEOROLÓGICAS**  
**UTILIZANDO ESTATÍSTICAS CIRCULAR E ESPACIAL DE ÁREA,**  
**NO ESTADO DO PARANÁ**

**CASCAVEL**

**2019**

**UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ – UNIOESTE**  
**CAMPUS DE CASCAVEL**  
**CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AGRÍCOLA**

**EDILZA MARTINS DA SILVA**

**PRODUTIVIDADE DA SOJA ASSOCIADA A VARIÁVEIS AGROMETEOROLÓGICAS**  
**UTILIZANDO ESTATÍSTICAS CIRCULAR E ESPACIAL DE ÁREA,**  
**NO ESTADO DO PARANÁ**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Agrícola, em cumprimento parcial aos requisitos para obtenção do título de Mestre em Engenharia Agrícola, área de concentração em Sistemas.

Orientadora: Dra. Luciana Pagliosa Carvalho Guedes

Coorientador: Dr. Jerry Adriani Johann

**CASCAVEL**

**2019**

Ficha de identificação da obra elaborada através do Formulário de Geração Automática do Sistema de Bibliotecas da Unioeste.

Silva, Edilza Martins da  
Produtividade da soja associada a variáveis agrometeorológicas utilizando estatísticas circular e espacial de área, no estado do Paraná / Edilza Martins da Silva; orientador(a), Luciana Pagliosa Carvalho Guedes; coorientador(a), Jerry Adriani Johann, 2019.  
77 f.

Dissertação (mestrado), Universidade Estadual do Oeste do Paraná, Campus de Cascavel, Centro de Ciências Exatas e Tecnológicas, Programa de Pós-Graduação em Engenharia Agrícola, 2019.

1. Análise temporal. 2. Data do máximo vegetativo. 3. Dados Espaciais. 4. Decêndios. I. Guedes, Luciana Pagliosa Carvalho. II. Johann, Jerry Adriani . III. Título.

## **EDILZA MARTINS DA SILVA**

Produtividade da soja associada as variáveis agrometeorológicas utilizando estatísticas circular e espacial de área, no estado do Paraná

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Agrícola em cumprimento parcial aos requisitos para obtenção do título de Mestra em Engenharia Agrícola, área de concentração Sistemas Biológicos e Agroindustriais, linha de pesquisa Geoprocessamento, Estatística Espacial e Agricultura de Precisão, APROVADO(A) pela seguinte banca examinadora:



Orientador(a) - Luciana Pagliosa Carvalho Guedes

Universidade Estadual do Oeste do Paraná - Campus de Cascavel (UNIOESTE)



Gustavo Henrique Dalposso

Universidade Tecnológica Federal do Paraná (UTFPR)



Eriuelto Mercante

Universidade Estadual do Oeste do Paraná - Campus de Cascavel (UNIOESTE)

Cascavel, 14 de fevereiro de 2019.

## **BIOGRAFIA**

Edilza Martins da Silva nasceu em 15 de março de 1981, no município de Jesuítas, no Estado do Paraná. Possui graduação em Licenciatura em Matemática pela União Educacional do Médio Oeste do Paranaense – UNIMEO em 2011. Especialista em Matemática pela Universidade Tecnológica Federal do Paraná – UTFPR em 2013, foi professora substituta na mesma Universidade no período de 2013 a 2015. Em março de 2017, ingressou no mestrado do Programa de Pós-Graduação Stricto Sensu em Engenharia Agrícola, com área de concentração em Sistemas Biológicos e Agroindustriais.

“Dai-me um ponto de apoio que  
eu moverei o mundo”

(Arquimedes)

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, pelo dom da vida e por mais esta oportunidade;

Aos meus pais, meus irmãos e sobrinhos, enfim toda a minha família, que sempre esteve ao meu lado;

À profa. Dra. Luciana Pagliosa Carvalho Guedes, pela orientação, o incentivo e a amizade;

Ao meu coorientador, prof. Dr. Jerry Adriani Johann, pela contribuição significativa na realização deste trabalho;

Aos membros do Laboratório de Estatística Aplicada e GeoScience pelo auxílio computacional e pelo companheirismo, em especial ao Alex, à Franciele, à Priscila e ao Becker;

À Elizabete e à Rosângela, pela amizade que surgiu durante esse período;

À CAPES, pelo apoio financeiro;

A todos que diretamente e indiretamente contribuíram para a realização deste trabalho.

# PRODUTIVIDADE DA SOJA ASSOCIADA A VARIÁVEIS AGROMETEOROLÓGICAS UTILIZANDO ESTATÍSTICAS CIRCULAR E ESPACIAL DE ÁREA, NO ESTADO DO PARANÁ

## RESUMO

A produtividade da cultura da soja é influenciada por diversos fatores que intervêm diretamente em todos os estádios fenológicos da planta. Neste contexto, é importante avaliar a influência que ocorre na produtividade, em relação às variáveis agrometeorológicas, tais como: precipitação (mm), temperatura ( $^{\circ}\text{C}$ ), radiação solar ( $\text{MJ dia}^{-1}$ ) e evapotranspiração (mm), assim como sua distribuição temporal. O objetivo do trabalho foi analisar e relacionar os valores e a variabilidade temporal de algumas variáveis agrometeorológicas com a produtividade da soja nos municípios paranaenses nos anos-safras 2011/2012 e 2013/2014. Inicialmente, realizou-se uma análise descritiva da distribuição da produtividade da soja e das medidas associadas às variáveis agrometeorológicas. Então, analisou-se a distribuição espacial destas por meio dos Índices de Moran global e local, univariado e bivariado, da produtividade com as variáveis agrometeorológicas. Além disso, analisou-se o momento médio de maior concentração das variáveis agrometeorológicas e a presença da sazonalidade destas variáveis, utilizando-se a estatística circular. Os resultados mostraram que a produtividade e todas as medidas associadas às variáveis agrometeorológicas apresentaram autocorrelação espacial significativa. E que de forma geral, houve associação espacial direta e significativa da produtividade com os valores da precipitação, bem como uma associação espacial indireta e significativa da produtividade com os valores das seguintes variáveis: temperatura, radiação solar e evapotranspiração. Em relação à sazonalidade das variáveis agrometeorológicas, a precipitação foi a única variável que não apresentou uniformidade quanto a sua distribuição temporal nos intervalos avaliados, ocorrendo, em média, maiores volumes de chuva no decêndio da data do máximo vegetativo, bem como nos intervalos de um a dois decêndios em torno da data do máximo vegetativo. Concluiu-se que as regiões do estado do Paraná com maiores ou menores valores de produtividade da soja estão correlacionadas com o valor das variáveis agrometeorológicas no momento do máximo de desenvolvimento vegetativo (DMDV) e em intervalos próximos, bem como a ocorrência de sazonalidade destas mesmas nos intervalos considerados nos dois anos-safras em estudos.

**Palavra-chave:** Análise temporal; Data do máximo vegetativo; Dados Espaciais; Decêndios; Sazonalidade.

# SOYBEAN PRODUCTIVITY ASSOCIATED WITH AGROMETEOROLOGICAL VARIABLES USING CIRCULAR AND SPATIAL STATISTICS, IN THE STATE OF PARANÁ

## ABSTRACT

The productivity of the soybean crop is influenced by several factors, which interfere directly in all the phenological stages of the plant. In this context, agrometeorological variables, such as precipitation (mm), temperature (°C), solar radiation (MJ dia<sup>-1</sup>), and evapotranspiration (mm), as well as their spatial distribution should be evaluated as to how they interfere with the productivity of a crop. The objective of this study was to analyze and correlate the values and temporal variability of some agrometeorological variables with the soybean yield in the state of Paraná in the 2011/2012 and 2013/2014 crop years. Firstly, a descriptive analysis of the distribution of soybean yield and measures associated with agrometeorological variables was performed. Next, the spatial distribution of these variables was evaluated through the global and local Moran indices, univariate and bivariate, between productivity and agrometeorological variables. In addition, the mean moment of greatest concentration of the agrometeorological variables was analyzed, as well as the presence of their seasonality, employing the circular statistic. The results showed that the productivity and all the measures associated with the agrometeorological variables presented significant spatial autocorrelation. In general, there was a direct and significant spatial association of productivity with precipitation values, as well as an inverse and significant spatial association of productivity with values of temperature, solar radiation, and evapotranspiration. Regarding the seasonality of the agrometeorological variables, precipitation was the only variable that did not show uniformity as to its temporal distribution in the evaluated intervals, occurring, on average, a larger amount of rain in the 10-day of the date of maximum vegetative development, as well as in the one to two 10-day intervals around the date of the maximum vegetative. Thus, it can be concluded that regions of Paraná State with higher or lower soybean yield are correlated with the amount of agrometeorological variables at the time of moment of maximum vegetative development (DMDV) and at close intervals, as well as the occurrence of seasonality in the intervals considered in the two-year study periods.

**Keywords:** Temporal analysis; Spatial data; Maximum vegetative data; 10-day; Seasonality.

## SUMÁRIO

<b>LISTA DE FIGURAS .....</b>	<b>ix</b>
<b>LISTA DE TABELAS.....</b>	<b>xi</b>
<b>1. INTRODUÇÃO .....</b>	<b>1</b>
<b>2. OBJETIVOS .....</b>	<b>3</b>
2.1. Objetivo Geral.....	3
2.1.1. Objetivos Específicos .....	3
<b>3. REVISÃO BIBLIOGRÁFICA.....</b>	<b>4</b>
3.1. Produção de grãos da soja no Brasil e no Paraná .....	4
3.1.1. Estádios Fenológicos da Soja.....	5
3.1.2. Exigências agrometeorológicas da cultura da soja .....	6
3.2. Conceitos da Estatística Circular .....	9
3.2.1. Dados Circulares .....	10
3.2.2. Representação Gráfica de Dados Circulares.....	11
3.2.3. Medidas da Estatística Descritiva Circular .....	12
3.2.4. Direção Média para dados circulares.....	12
3.3. Análise de Dados Espaciais: Estatística espacial de área .....	16
3.4. Comparação de mapas temáticos .....	21
3.5. Análise Estatística Multivariada .....	22
<b>4. MATERIAL E MÉTODOS .....</b>	<b>25</b>
4.1 Área de estudo .....	25
4.2 Aquisição dos dados.....	26
4.2.1 Variáveis agrometeorológicas .....	26
4.2.2 Produtividade da soja.....	27
4.3 Análise descritiva das variáveis em estudos .....	28
4.3 Estatística Circular .....	29
4.4 Análise Espacial de Área .....	29
4.4 Análise Estatística Multivariada.....	30
<b>5. RESULTADOS E DISCUSSÃO.....</b>	<b>31</b>
5.1 Análise descritiva da produtividade da soja e das variáveis agrometeorológicas .....	31

5.2 Análise dos decêndios do máximo desenvolvimento vegetativo (DMDV) da soja no estado do Paraná .....	33
5.3. Dados agrometeorológicos no decêndio de máximo desenvolvimento vegetativo (DMDV).....	33
5.4 Dados agrometeorológicos em decêndios próximos ao máximo desenvolvimento vegetativo (DMDV) .....	36
5. 5 Sazonalidade e Decêndio médio da ocorrência das variáveis.....	42
5.6 Análise Espacial de Área .....	45
5.7 Análise Multivariada.....	52
<b>6. CONCLUSÃO.....</b>	<b>56</b>
<b>7. CONSIDERAÇÕES FINAIS.....</b>	<b>57</b>
<b>REFERÊNCIAS .....</b>	<b>58</b>

## LISTA DE FIGURAS

Figura 1 Área cultivada de soja no Brasil.....	4
Figura 2 Produtividade da soja no Brasil.....	5
Figura 3 Representação dos dados circulares: a) dados não agrupados, b) histograma circular, c) histograma linear, d) histograma de rosas .....	11
Figura 4 Amostra de dados circulares: a) Representação do i-ésimo ponto amostral, b) Centro de gravidade de uma amostra de tamanho $n=3$ .....	12
Figura 5 Representação Gráfica de uma amostra circular e vetor média (setas) .....	13
Figura 6 Representação do vetor resultante (R) e comprimento do vetor média (R).....	14
Figura 7 Variação do comprimento do vetor média em diferentes amostras de dados circulares, com tamanho da amostral $n= 5$ (seta representa o vetor média) .....	15
Figura 8 Representação Gráfica do Diagrama do Índice de Moran.....	19
Figura 9 Etapas do desenvolvimento da pesquisa .....	25
Figura 10 Área de estudo com as 10 mesorregiões do estado do Paraná .....	26
Figura 11 (a) e (b) Representações dos pontos geográficos da malha regular das estações virtuais do ECMWF; (c) Exemplo de uma EV selecionada com o percentual da área dos 3 municípios ( $a_1, a_2, a_3$ ) que a compõe. ....	28
Figura 12 Mapa da distribuição de produtividade ( $t\ ha^{-1}$ ) dos anos-safra: (a) 2011/2012 e (b) 2013/2014.....	31
Figura 13 Mapa da distribuição do decêndio de máximo desenvolvimento vegetativo no estado do Paraná: a) do ano-safra de 2011/2012 e b) do ano-safra de 2013/2014 .....	33
Figura 14 Mapa dos valores das seguintes variáveis no momento do DMDV: a) e b) temperatura média; c) e d) radiação solar; e) e f) evapotranspiração potencial; g) e h) precipitação; para os anos-safra do 2011/2012 e 2013/2014, respectivamente .....	35
Figura 15 Distribuição espacial das medidas associadas à temperatura média (TMed) considerando as variações decendiais: a) e c) do anos-safra do 2011/2012 e, b) e d) do ano-safra 2013/2014.....	37
Figura 16 Distribuição espacial das medidas associadas a radiação solar ( $R_s$ ) considerando as variações decendiais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014 .....	38
Figura 17 Distribuição espacial das medidas associadas à evapotranspiração potencial (ETo) considerando os intervalos decendiais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014.....	39
Figura 18 Distribuição espacial das medidas associadas à variável precipitação acumulada (Prec) considerando as variações decendiais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014.....	40
Figura 19 Distribuição espacial das variáveis em todo ciclo vegetativo da soja das seguintes variáveis agrometeorológicas: a) e b) temperatura média; c) e d) radiação solar acumulada;	

e) e f) evapotranspiração potencial acumulada; g) e h) precipitação acumulada; para os anos-safra do 2011/2012 e 2013/2014 .....	41
Figura 20 Distribuição espacial dos valores do R (comprimento do vetor média) e das variáveis: temperatura média, radiação solar acumulada, evapotranspiração potencial acumulada e precipitação acumulada do ano-safra do 2011/2012 e 2013/2014. ....	43
Figura 21 Distribuição espacial do volume de chuva do ano-safra 2011/2012 no intervalo decendial 1aDMDV1d. Histograma circular (unidade de medida em graus) do lado direito representa o pixel com a maior produtividade no estado do Paraná. Histograma circular (unidade de medida em graus) do lado direito representa o pixel com a menor produtividade no estado do Paraná, a seta aponta para a data média e o comprimento da seta é o valor de R (varia de 0 a 1) .....	44
Figura 22 Distribuição espacial da média circular das variáveis: temperatura média, radiação solar acumulada, evapotranspiração potencial acumulada e precipitação acumulada do ano-safra do 2011/2012 e 2013/2014 .....	45
Figura 23 Lisa Cluster Map da variável produtividade nos anos-safra: a) 2011/2012; b) 2013/2014.....	46
Figura 24 Lisa Cluster Map das medidas das variáveis agrometeorológicas dos anos-safra 2011/2012 e 2013/2014.....	51
Figura 25 Distribuição espacial do agrupamento nas variações decendiais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014 .....	52
Figura 26 Representação gráfica da classificação obtida para o ano-safra 2011/2012 (o 1) e 2013/2014 ( $\Delta$ 2): a) intervalo decendial 1aDMDV1d e b) 2aDMDV2d. ....	54

## LISTA DE TABELAS

Tabela 1 Descrição dos estádios de desenvolvimento da cultura da soja .....	6
Tabela 2 Evapotranspiração média nos diversos subperíodos fenológicos da soja .....	9
Tabela 3 Matriz de erros de ordem $k \times k$ .....	21
Tabela 4 Qualidade da classificação do Índice Kappa ( $K$ ) e Tau ( $\tau$ ) .....	22
Tabela 5 Tabela genérica das frequência dos erros de classificação.....	24
Tabela 6 Medidas associadas às variáveis agrometeorológicas .....	27
Tabela 7 Estatística descritivas para média ponderada da produtividade (PMP) da soja e as medidas associadas às variáveis agrometeorológicas.....	32
Tabela 8 Índice de Moran Global Univariado ( $I$ ) e Bivariado ( $I_{xy}$ ) entre produtividade da soja e as medidas associadas à evapotranspiração potencial (ETp).....	47
Tabela 9 Índice de Moran Global Univariado ( $I$ ) e Bivariado ( $I_{xy}$ ) entre produtividade da soja e as medidas associadas à radiação solar ( $R_s$ ).....	47
Tabela 10 Índice de Moran Global Bivariado entre produtividade da soja e as medidas associadas à Temperatura média (TMed).....	48
Tabela 11 Índice de Moran Global Univariado e Bivariado entre produtividade da soja e as medidas associadas à precipitação (Prec).....	48
Tabela 12 Componentes Principais; Proporção da variância explicada; Proporção acumulada da variância total para as variáveis agrometeorológicas dos anos-safras em estudos.....	53
Tabela 13 Índices de classificação de qualidade da função quadrática .....	54

## 1. INTRODUÇÃO

O desenvolvimento da soja, principalmente na região sul do Brasil, ocorreu devido à semelhança com o ecossistema do Estados Unidos, onde a cultura se desenvolveu inicialmente, sendo utilizadas cultivares adaptadas às baixas latitudes de climas tropicais (EMBRAPA SOJA, 2003).

A consolidação da soja como uma das principais culturas do agronegócio brasileiro e com uma melhor rentabilidade em relação a outras culturas faz com que os produtores se sintam estimulados a continuar apostando nesta cultura (CONAB, 2018).

As variações a longo prazo da produtividade da cultura da soja, são causadas pela introdução de novas técnicas de manejo, de variedades e de insumos. Por outro lado, as variações de curto prazo na produtividade, isto é, de um ano para outro, são causadas principalmente pelas alterações climáticas (FERREIRA; COSTA; SOUZA, 2007).

Dessa forma, diversos trabalhos sobre a relação entre as variáveis agrometeorológicas e a produtividade da soja foram realizados no Brasil, sendo verificada uma grande influência destas variáveis no rendimento da produtividade (FERREIRA; COSTA; COSTA, 2007; SILVA; SARTORI; WOLLMANN, 2014; VIVAN et al., 2015; RADIN, SCHÖNHOFEN; TAZZO, 2017).

Vivan et al. (2015) identificaram e caracterizaram o efeito da restrição hídrica sobre o rendimento de grãos da soja, na região de Santiago no Rio Grande do Sul. Verificaram uma redução da produção e significativa redução de lucro em função de variabilidade climática.

Em um estudo realizado por Radin, Schönhofen e Tazzo (2017), foi constatado que a variação do rendimento de grãos pode ser explicada pela quantidade da chuva, principalmente nos estádios de floração e enchimento de grãos.

Pode-se determinar a distribuição das variáveis ao longo do tempo, sendo possível determinar um padrão da ocorrência das variáveis, permitindo uma previsão do comportamento climático de uma região, utilizando, assim, para planejamento e gestão de inúmeras atividades humanas (ASSIS et al., 2004).

Em relação à distribuição das variáveis agrometeorológicas, na atividade agrícola, a distribuição da chuva é um dos mais importantes fatores que influencia na produção agrícola (BRAIDO; TOMMASELLI, 2010).

Dentre as variáveis agrometeorológicas, a precipitação é a variável que mais influencia o rendimento dos grãos da soja, ao considerar a variação temporal, sendo este o período de primavera-verão, isto na região sul do Brasil (FONTANA et al., 2015).

Ressalta-se que, especificamente no estado do Paraná, estudos mostraram que existe uma diferença de produtividade da soja entre diferentes anos-safras, que são influenciados por fatores climáticos durante as fases do ciclo da cultura como, por exemplo, falta de precipitação, alta radiação global, evapotranspiração, temperatura e umidade relativa do ar, tanto na fase da semeadura como na fase de desenvolvimento da cultura, afetando a sua

produtividade final (ANJOS; NERY, 2005; GRZEGOZEWSKI et al., 2016; ARAÚJO; URIBE-OPAZO; JOHANN, 2016).

Desta maneira, os agricultores podem utilizar as informações relacionadas ao cultivo da soja, data de semeadura e área semeada, associando às condições meteorológicas ocorridas para avaliar os riscos da soja (RADIN; MATZENAUER, 2016).

Neste sentido, justifica-se a importância de verificar a influência das variáveis agrometeorológicas na produtividade da soja no estado do Paraná, identificando as regiões propícias a um melhor rendimento de grãos e a relação da variabilidade temporal e espacial das variáveis agrometeorológicas com a produtividade da soja.

Desta forma, a presente dissertação foi elaborada com a seguinte estrutura: no primeiro capítulo é apresentada uma introdução sobre o assunto abordado. No segundo capítulo, os objetivos geral e específicos. O terceiro consiste em uma revisão bibliográfica sobre a soja e as metodologias adotada: estatística circular e a estatística espacial de área. No quarto capítulo é apresentada a área de estudo, a aquisição dos dados e as metodologias utilizadas. No capítulo seguinte é relatado sobre as análises encontradas, como a distribuição espacial das variáveis produtividade e as medidas associadas às variáveis agrometeorológicas no estado do Paraná para o anos-safras 2011/2012 e 2013/2014 e os resultados obtidos sobre a análise espacial de área, por meio dos Índices de Moran Global e Local univariado e bivariado. Ao fim deste, temos a apresentação do momento das ocorrências decendiais e a sazonalidade das medidas associadas às variáveis agrometeorológicas, por meio da estatística circular. Finalizando a dissertação, têm-se as conclusões, as considerações finais sobre os resultados encontrados e a lista bibliográfica utilizada.

## **2. OBJETIVOS**

### **2.1. Objetivo Geral**

Analisar e relacionar os valores e a variabilidade temporal de variáveis agrometeorológicas com a produtividade da soja nos municípios paranaenses nos anos-safras 2011/2012 e 2013/2014.

#### **2.1.1. Objetivos Específicos**

- Analisar a distribuição espacial das variáveis agrometeorológicas no momento do máximo desenvolvimento vegetativo (DMDV) da soja e nas variações de intervalos próximos ao DMDV durante os anos-safras;
- Verificar a ocorrência de sazonalidade e determinar o momento médio de ocorrência das maiores concentrações das variáveis agrometeorológicas, calculado por medidas da estatística circular, nas variações decendiais consideradas;
- Avaliar a distribuição espacial da produtividade da soja e das medidas associadas às variáveis agrometeorológicas, utilizando técnicas de Índices de Moran Global e Local univariado;
- Analisar a correlação espacial entre a variável produtividade e as medidas associadas às variáveis agrometeorológicas, por meio do Índice de Moran Global bivariado.

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1. Produção de grãos da soja no Brasil e no Paraná

A soja (*Glycine max (L.) Merrill*) foi introduzida no Brasil pelo estado da Bahia em 1882, mas a cultura não se desenvolveu inicialmente. Em 1890 novos genótipos foram testados e aprovados no estado do Rio Grande do Sul, desta vez desenvolvendo-se. Porém, até o final da década de 1940, a produção era apenas destinada à alimentação de animais (INMET, 2009). A partir da década de 1960, a soja se estabeleceu como cultura economicamente importante e responsável pelo surgimento da agricultura comercial no Brasil.

Segundo a Embrapa Soja (2003), no estado do Paraná a produção da soja se tornou importante na agricultura, como lavoura comercial, em meados da década de 50. Destacando-se que na região Sul do estado, o cultivo iniciou-se como alternativa para rotação de cultura, enquanto que no Nordeste, Norte e Sudoeste, o cultivo da soja iniciou-se após a destruição dos cafezais pela grande geada de 1953. Devido às terras férteis, baixo preço e o clima favorável, as regiões sudoeste e oeste do Paraná foram alvo de disputa na década de 1960 por gaúchos e catarinenses, sendo responsáveis pela introdução e expansão da soja nessas regiões (CAMPOS, 2011).

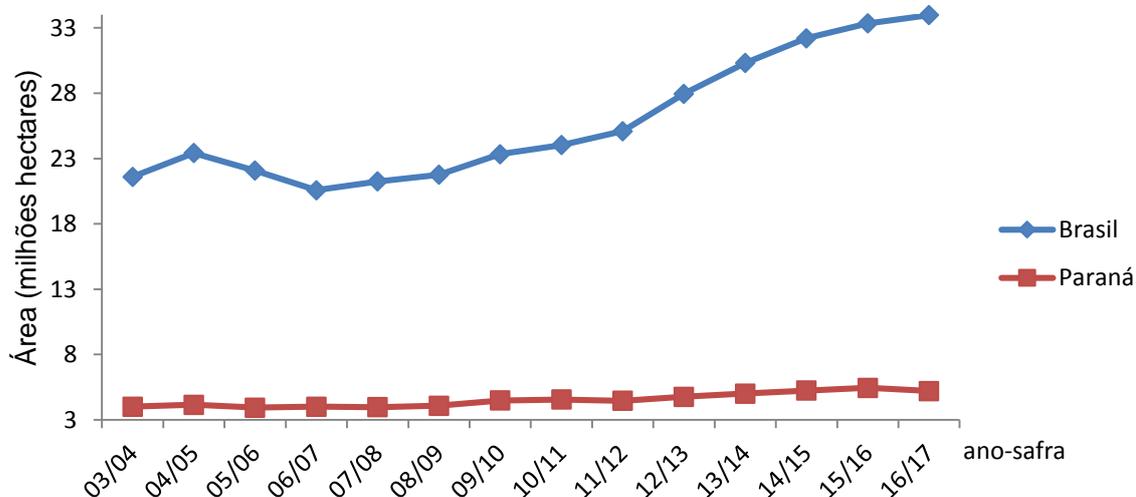


Figura 1 Área cultivada de soja no Brasil

Fonte: CONAB (2018)

De acordo com a CONAB (2018) soja tem sido a protagonista no aumento da produção de grãos no Brasil nos últimos anos. A área plantada foi de 35 milhões de hectares (Mha) no ano-safra de 2017/2018 e o estado do Paraná participou com mais de 5,4 Mha (Figura 1). Destaca-se, na Figura 1, que no ano-safra 2011/2012 houve uma área planta no Brasil de 25 Mha e no estado do Paraná de 4,46 Mha, sendo uma área menor do que no ano-safra 2013/2014, de 30,3 Mha e 5 Mha, respectivamente.

Além disso, a produtividade da soja do país do ano-safra de 2017/2018 (Figura 2) foi de 3,39 t ha<sup>-1</sup> e no estado do Paraná de 3,36 t ha<sup>-1</sup> (CONAB, 2018). Com uma produtividade em média superior ao ano-safra anterior (2016/2017), isso foi possível devido ao clima do estado e à utilização de tecnologias por parte do setor produtivo (EMBRAPA SOJA, 2017).

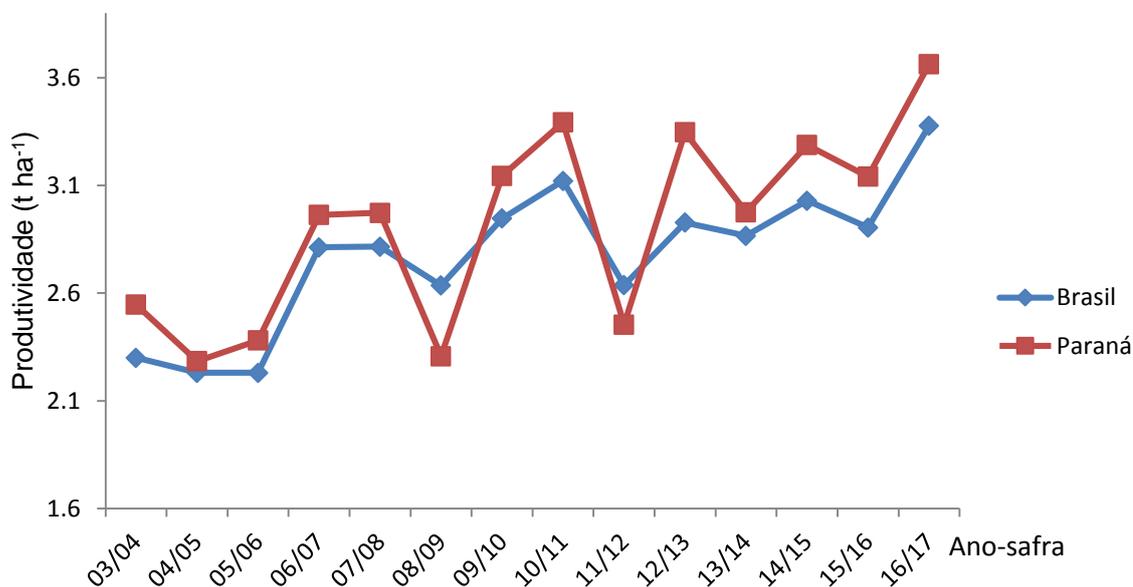


Figura 2 Produtividade da soja no Brasil  
Fonte: CONAB (2018) e IBGE (2018)

Para ano-safra 2011/2012 (Figura 2) a produtividade da soja no país foi de 2,63 t ha<sup>-1</sup>, e no estado de Paraná foi de 3,45 t ha<sup>-1</sup>. Já no ano-safra 2013/2014, houve uma produtividade de 2,86 t ha<sup>-1</sup> e 2,97 t ha<sup>-1</sup>, respectivamente (IBGE, 2018).

No estado do Paraná, a época do plantio da soja é determinada por meio da portaria nº 202, que estabelece o período do vazio sanitário e determina as datas limites para a semeadura e a colheita da soja para o controle da ferrugem asiática (*Phakopsora pachyrhizi*). Nesta portaria é estabelecido que o período do vazio sanitário vegetal no território paranaense entre 10 de junho e 10 de setembro de cada ano, ou seja, é o período que é proibido cultivar, manter ou permitir a existência de plantas vivas da soja em qualquer estágio vegetativo (ADAPAR, 2017).

### 3.1.1. Estádios Fenológicos da Soja

Os estádios fenológicos da soja no estado são definidos de acordo com o tipo da variedade de cada cultivar (EMBRAPA SOJA, 2013). Estes dividem-se em dois estádios fenológicos: o vegetativo (V) e o reprodutivo (R) (Tabela1). O estágio vegetativo se inicia no estágio emergência (VE), seguido do estágio Cotilédone (VC). A partir deste, o estágio vegetativo (V) é subdividido sequencialmente (V1, V2,...,V<sub>r</sub>), em que *r* é o número de nós

existentes, acima do nó cotilédone, com folha completamente desenvolvida (EMBRAPA, 2007).

**Tabela 1 Descrição dos estádios de desenvolvimento da soja**

Estádios Vegetativos	Estádios Reprodutivos
VE- Emergência	R1- Início da floração
VC- Cotilédone	R2- Florescimento Pleno
V1- Primeiro nó	R3- Início da formação de vagem
V2- Segundo nó	R4- Vagem completamente desenvolvida
V3- Terceiro nó	R5- Início do enchimento do grão
V4- Quarto nó	R6- Grão cheio ou completo
:	R7- Início da maturação
V <sub>r</sub> - r -ésimo nó	R8- Maturação plena

Fonte: Fehr e Caviness (1977) e Embrapa (2007)

Já o estágio reprodutivo (R) abrange quatro distintos grupos, sendo esses: o florescimento da planta (R1 e R2), o desenvolvimento da vagem (R3 e R4), o desenvolvimento do grão (R5 e R6) e a maturação da planta (R7 e R8) (FEHR; CAVINESS, 1977).

### **Exigências agrometeorológicas da cultura da soja**

De acordo com Klosowski (1997), a relação existente entre os elementos agrometeorológicos e a produção agrícola é muito complexa, pois os elementos agrometeorológicos afetam diretamente o crescimento e o desenvolvimento das plantas nas diversas fases do ciclo da cultura. Este autor menciona que os elementos que mais influenciam no desenvolvimento e na produtividade da cultura da soja são: disponibilidade hídrica, temperatura, fotoperíodo e evapotranspiração.

Apesar do vasto conhecimento tecnológico para o cultivo da soja, a disponibilidade hídrica durante a fase de crescimento da planta é a principal limitação do potencial de produtividade da cultura (EMBRAPA, 2007), assim como corresponde a uma das principais causas da variabilidade da produtividade da cultura, de um ano para o outro, ou de uma região para outra (EMBRAPA SOJA, 1999).

De acordo com Bonato (2000), tanto o excesso quanto a falta de chuva no estágio vegetativo da soja podem acarretar em uma redução drástica da produção de grãos. Enquanto a deficiência de água reduz o crescimento da planta e sua área foliar, a água em demasia resulta em plantas de pequena estatura, comprometendo também a produção dos grãos.

Desta maneira, a água é um dos elementos climáticos mais importantes para o desenvolvimento da soja, principalmente nos períodos de germinação-emergência e de floração-enchimento de grãos. No período de germinação-emergência, a planta necessita absorver no mínimo 50% do seu peso em água para a ocorrência de uma boa germinação. Conforme se desenvolve, a planta necessita absorver uma quantidade maior de água até

atingir a quantidade máxima, que ocorre durante o período de floração-enchimento, sendo de 7 a 8 mm por dia, e somente após esse período a quantidade de absorção de água diminui (EMBRAPA, 2007).

Segundo a Embrapa Soja (2003) o consumo mais elevado de água coincide com o período em que a cultura apresenta maior altura e índice de área foliar, sendo um total de água por ciclo da cultura de 450 a 800 mm. Contudo, esse valor do consumo de água depende do clima, do manejo da cultura e da duração do seu ciclo. Para um bom desempenho, além de um volume de água adequado, deve ocorrer uma boa distribuição de chuvas ao longo de todo o ciclo, ou seja, a distribuição precipitação pluviométrica sem uniformidade é um fator limitante para a obtenção de alta produtividade de grãos (EMBRAPA, 2007).

Em trabalhos realizados pela Embrapa (2007), ao longo de 15 safras, em Londrina, PR, verificou-se que, para um melhor rendimento de grãos da soja, é necessário um volume de água em torno de 650 a 700 mm, bem distribuído em todo o ciclo. Observou-se também que ao ocorrer o déficit hídrico durante os dois estádios da cultura, mesmo que o volume de água seja de 700 mm, a produtividade dos grãos foi comprometida, pois o volume da chuva não estava distribuído uniformemente, principalmente no estágio reprodutivo.

Além disso, nestes trabalhos realizados pela EMBRAPA (2007) constatou-se que o volume de água ideal para a fase crítica, ou seja, para a fase reprodutiva (desde o início do enchimento do grão até o grão cheio), é de 120 a 300 mm, devendo ser bem distribuído ao longo desse período. Desta forma, verificou-se que para a obtenção de alta produtividade de soja é necessário um volume de água distribuído ao longo de todo o ciclo.

Para solucionar o problema causados pela deficiência ou pelo excesso hídricos, deve-se semear apenas cultivares adaptadas à região e à condição do solo. No caso do déficit hídrico, indica-se a adoção de práticas que favoreçam o armazenamento de água no solo, como a irrigação, que deve ocorrer principalmente no estágio vegetativo. Já para o excesso hídrico, recomendam-se as práticas que favoreçam a drenagem e a aeração do solo sujeito ao encharcamento (BONATO, 2000).

Em relação à temperatura recomenda-se que o cultivo da soja esteja em um ambiente com temperatura entre 20 a 30 °C, sendo considerado como temperatura ideal em torno de 30 °C (EMBRAPA, 2007). Além disso, ao cultivar a soja em regiões com temperaturas baixas (< 15 °C) ocorre a diminuição no tamanho da planta, atrasando o seu desenvolvimento e, como consequência, o seu florescimento e o seu ciclo são estendidos (BONATO, 2000). Para as regiões com temperaturas acima de 40 °C, ocorrem estragos na floração, diminuição e capacidade de retenção das vagens, sendo agravados com a ocorrência de déficit hídrico (EMBRAPA, 2010).

Para o momento da semeadura não é recomendável que a temperatura do solo esteja abaixo de 20 °C, pois isto prejudica a germinação e a emergência da semente, tendo como faixa de temperatura do solo adequada para a semeadura entre 20 e 30 °C, sendo 25 °C a

temperatura ideal para uma emergência rápida e uniforme (EMBRAPA, 2007).

De acordo com Embrapa Soja (2003), a floração da soja é somente induzida quando ocorre temperatura ambiente acima de 13 °C. Este resultado foi observado com a análise de diferentes datas de floração apresentadas por uma mesma cultivar e semeada em uma mesma época em diversos anos.

Bonato (2000) menciona que, ao ocorrer temperatura acima de 30 °C durante os estádios vegetativos, ocorre a diminuição do crescimento da planta e o adiantamento do seu desenvolvimento, causando florescimento precoce, o que pode ocasionar um comprometimento da produtividade.

Havendo altas temperaturas (>40 °C) no momento do estágio de maturação, associadas a períodos de alta umidade, a qualidade dos grãos pode ser afetada (EMBRAPA SOJA, 2003). Ao ocorrer temperatura baixa associada a períodos chuvosos ou alta umidade, ocorrerá um atraso na data da colheita (FARIAS; NEUMAIER; NEPOMUCENO, 2009).

O outro elemento agrometeorológico que influencia o ciclo fenológico da soja é o fotoperíodo (EMBRAPA, 2007). Segundo Varejão-Silva (2005), esse é o intervalo de tempo transcorrido entre o nascer e o pôr do sol em determinado local e data, não sendo considerado apenas o período no qual o lugar está iluminado, mas também os períodos do crepúsculo matutino e vespertino.

De acordo com Embrapa (2010), a adaptação de diferentes cultivares da soja em determinadas regiões, depende não somente das disponibilidades hídrica e térmica, mas também do fotoperíodo, sendo que a sensibilidade ao fotoperíodo é uma característica variável entre as cultivares, ou seja, cada cultivar possui o seu fotoperíodo crítico. Por essa razão, a soja é denominada como planta de dia curto.

Thomas (2017) menciona que o fotoperiodismo influencia no florescimento das plantas, na dormência de gemas e na iniciação de bulbos ou tubérculos, interferindo na produtividade da cultura.

De acordo com Rodrigues et al. (2002), o baixo nível de radiação solar acumulada durante os estádios reprodutivos da soja tem sido a principal causa da perda do rendimento da soja.

Por fim, a evapotranspiração potencial (ETp) também influencia no ciclo vegetativo da soja, sendo esta ação conjunta entre a evaporação (fenômenos da natureza física que transforma em vapor a água da superfície do sol, lagos e mares) com a transpiração (evaporação da ação fisiológica dos vegetais) (PINTO et al., 1976).

De acordo com Allen et al. (1998), existem alguns fatores que afetam o processo de evapotranspiração, tais como: características da cultura e seu o estágio de desenvolvimento; características climáticas (radiação, temperatura do ar, umidade atmosférica e vento); aspectos ambientais, considerando a salinidade do solo; a baixa fertilidade da terra; aplicação de fertilizante; e ausência de controle de doenças e pragas, limitando o desenvolvimento da

cultura e reduzindo a evapotranspiração.

Em relação à quantidade de evapotranspiração diária durante o desenvolvimento da cultura da soja, ocorre uma maior quantidade diária de evapotranspiração no início do estágio de florescimento (R1) até o início do estágio de enchimento dos grãos (R5) (Tabela 2).

Tabela 2 Evapotranspiração média nos diversos subperíodos fenológicos da soja

Subperíodo	Evapotranspiração diária (mm)
Semeadura – Emergência	2,2
Emergência - Início da fase reprodutiva	5,1
R1 – R3. (Florescimento até início da frutificação)	7,4
R3 – R7. (Frutificação até maturação fisiológica)	6,6
R7 - R8. (Maturação fisiológica até maturação do grão)	3,7
Média do ciclo	5,8

Fonte: EMBRAPA (2007)

De acordo com Embrapa (2007), é importante o conhecimento da evapotranspiração máxima nos diversos períodos de desenvolvimento da soja, para o manejo da água em plantação irrigada. Além disso, tem-se que, para o cultivo da lavoura da soja não irrigada, essa informação é útil para decisões sobre práticas de culturas que permitam o melhor aproveitamento das disponibilidades hídricas naturais de cada região. Desta maneira, é possível decidir a época de semeadura, de modo a evitar que os períodos críticos em relação à água coincidam com períodos de menor disponibilidade de água para cultura.

Embrapa (2004) menciona que a soja necessita de uma umidade adequada do solo para poderem absorver os nutrientes necessários ao crescimento e ao desenvolvimento, conseqüentemente com maiores rendimentos de grãos, sendo essencial um fornecimento adequado da água pelas cultivares. Desta maneira, os estádios fenológicos que a evapotranspiração mais influencia são os estádios em que a precipitação influencia, sendo eles emergência e florescimento até enchimento de grãos.

### 3.2. Conceitos da Estatística Circular

A estatística circular (ou direcional) é considerada um ramo particular da estatística, na qual seus dados são denominados como dados direcionais, angulares ou circulares (PEWSEY; NEUHÄUSER; RUXTON, 2013). Em uma amostra de dados circular, cada observação de uma variável circular tem sua medida variando entre  $0^\circ$  e  $360^\circ$  ou equivalente de 0 até  $2\pi$ , ou seja, são medidas cíclicas (BATSCHELET, 1981). Cada observação pode ser representada, geometricamente como um ponto na circunferência unitária, no caso bidimensional (MARDIA, 1972).

De acordo com Mardia (1972), a estatística circular é aplicada em diversas áreas do conhecimento, como, por exemplo, na Meteorologia, tal como no estudo da direção e velocidade do vento, sendo considerada a representação no sentido horário como Norte ( $0^\circ$ ) Leste ( $90^\circ$ ) Sul ( $180^\circ$ ) e Oeste ( $270^\circ$ ), e em seguida volta-se para o Norte ( $360^\circ$ ). Ou, ainda,

a estatística circular pode ser aplicada em dados provenientes de precipitação, evaporação de bacia hidrográfica ou outros ciclos da hidrologia.

Neste contexto, Beskow et al. (2014) avaliaram a sazonalidade do regime temporal e espacial de vazões de estiagem em diferentes bacias hidrográficas no estado do Rio Grande do Sul, por meio de três métodos de sazonalidade, sendo um deles o Índice de Sazonalidade, que faz uso de dois parâmetros da estatística circular, calculados a partir das datas do calendário Juliano, considerando todos os dias do período de observação, na qual as vazões foram iguais ou inferiores à vazão de referência de  $Q_{90}$  (vazão que é superada a 90% do tempo). Assim, determinaram o dia médio da ocorrência das vazões de estiagem do estado, a existência (considerando a vazão de referência) de uma distribuição espacial dos parâmetros da estatística circular, e ainda identificaram uma boa alternativa para delinear regiões hidrologicamente homogêneas no que se refere às vazões de estiagem no estado.

Na área agrícola, pesquisadores aplicaram a estatística circular para avaliar o efeito da orientação da semente sobre o ângulo da folha do milho em relação à linha plantada e à taxa de emergência (TORRE et al., 2011). Além disso, Koller et al. (2016) verificaram a orientação da semente em relação à ganhos de produtividade devido a luz otimizada pela planta.

Desta forma, a estatística circular é aplicada em dados que apresentam padrão periódico e em medidas angulares nas diversas áreas do conhecimento.

### 3.2.1. Dados Circulares

Os dois principais instrumentos de medição dos dados circulares são a bússola e o relógio. Por meio da bússola pode-se medir, por exemplo, a direção do vento ou da migração dos pássaros. Por meio do relógio, é possível medir observações que consideram o tempo como, por exemplo, a hora de chegada de um paciente em um determinado hospital ou o dia do mês de um determinado ano da ocorrência de um evento (MARDIA; JUPP, 1999).

Jammalamadaka e Seagupta (2001) citam que algumas observações dos dados circulares podem muitas vezes se referir aos eixos, não atribuindo um sentido à direção observada, sendo considerado como dados axiais.

Por padrão, deve-se converter os dados axiais em dados vetoriais, transformando um ângulo ( $\theta$ ) em  $2\theta$ , removendo, assim, a ambiguidade das direções, podendo ser expressos em grau ou em radiano, com intervalos de variação  $[0; 2\pi]$  ou  $[-\pi; \pi]$  dependendo do interesse da análise (MARDIA; JUPP, 1999).

Para a obtenção dos dados circulares representados por um ângulo direcional ( $\theta$ ) utilizam-se as equações:

$$\theta = \frac{(360^\circ)(D)}{k} \text{ (em graus), ou } \theta = \frac{(2\pi)(D)}{k} \text{ (em radianos).} \quad \text{Eq. (1)}$$

em que  $D$  é uma medida de tempo (com uma determinada unidade), observada num período de tamanho  $k$  (com mesma unidade de medida de  $D$ ). Como, por exemplo, para a conversão do tempo  $D$  (em horas) observado dentro de um período de um dia, tem-se  $k = 24$  horas; ou para converter  $D$  dias observados num período de um ano, tem-se  $k = 365$  (ou  $k = 366$  para ano bissexto) (ZAR, 2010).

A transformação de um ângulo ( $\theta$ ) para um momento no tempo ( $D$ ) é dado por (BESKOW, 2014):

$$D = \theta \frac{365}{360} \quad \text{ou} \quad D = \theta \frac{365}{2\pi}. \quad \text{Eq. (2)}$$

Os dados circulares podem ser representados por diagramas e gráficos, como será descrito posteriormente.

### 3.2.2. Representação Gráfica de Dados Circulares

Ao considerar os dados circulares não agrupados, a sua representação no caso bidimensional é um ponto em uma circunferência unitária, para cada observação (Figura 3a).

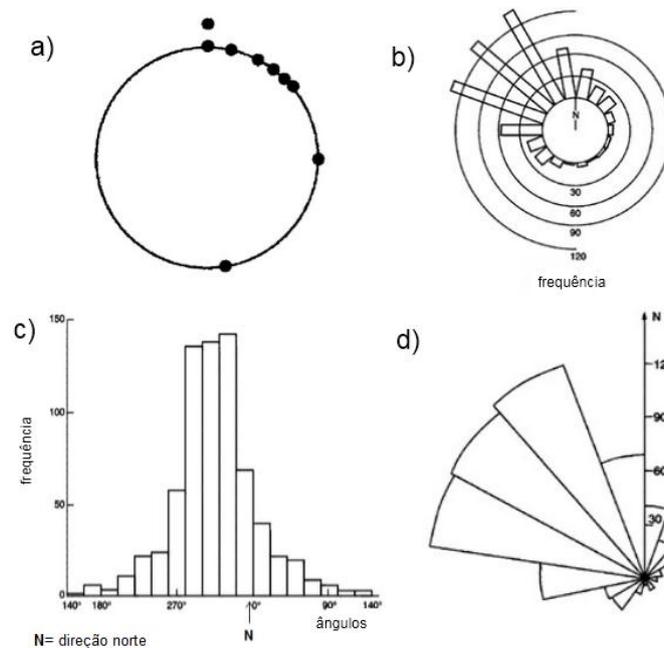


Figura 3 Representação dos dados circulares: a) dados não agrupados, b) histograma circular, c) histograma linear, d) histograma de rosas.

Fonte: Mardia e Jupp (1999)

De acordo com Mardia (1972), quando os dados são agrupados eles podem ser representados por uma das três seguintes maneiras: histograma circular (Figura 3b), histograma linear (Figura 3c) ou diagrama de rosa (Figura 3d). Esse autor menciona que o histograma circular (Figura 3b) é formado por um conjunto de barras que representa as frequências de classe, cuja base está sobre a circunferência de um círculo unitário, sendo dividida de acordo com o tamanho de cada classe

Ao realizar um corte no eixo do gráfico do histograma circular e transportá-lo ao eixo  $X$ , o gráfico será transformado em um histograma semelhante ao de um conjunto de dados

linear (Figura 3c). Já no diagrama de rosas (Figura 3d), as barras são substituídas por setores circulares, nos quais a área de cada setor é proporcional ao comprimento do seu respectivo raio (que é a frequência de um determinado evento ocorrido na sua classe) (MARDIA, 1972).

Após a coleta dos dados circulares e sua representação gráfica, é necessário estabelecer algumas estatísticas descritivas, pois os métodos de análise utilizado para os dados lineares não são adequados para os dados circulares (JAMMALAMADAKA; SENGUPTA, 2001).

### 3.2.3. Medidas da Estatística Descritiva Circular

De acordo com Mardia (1972), uma amostra aleatória de tamanho  $n$  proveniente de uma população de dados circulares é composta por  $\theta_i$  ( $i = 1, \dots, n$ ), definido como  $i$ -ésimo ângulo direcional (Figura 4a) (definido a partir do eixo na direção positiva), sendo que este ângulo é também uma representação do vetor unitário  $\overrightarrow{OP_i} = e_i$  (Figura 4a), em que  $O$  é a origem da circunferência unitária e  $P_i = (x_i, y_i)$  ( $i = 1, \dots, n$ ) é um ponto localizado nesta circunferência.

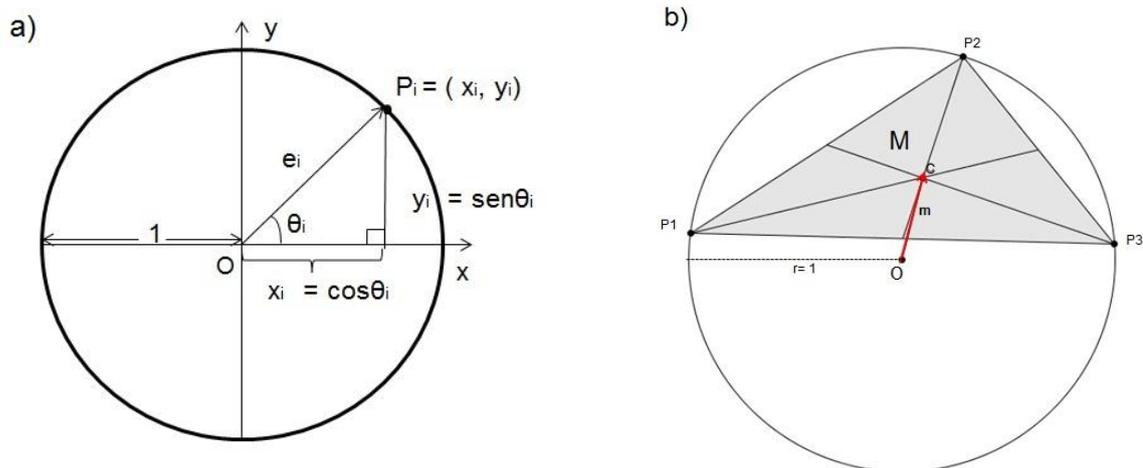


Figura 4 Amostra de dados circulares: a) Representação do  $i$ -ésimo ponto amostral, b) Centro de gravidade de uma amostra de tamanho  $n=3$ .

O ponto  $P_i$  pode ser representado pelas coordenadas retangulares, com  $y_i = R \text{sen}\theta_i$  e  $x_i = R \text{cos}\theta_i$  (Figura 4a). Como há interesse apenas na direção do vetor e não na sua magnitude, considera-se que o comprimento do vetor é unitário ( $R=1$ ). Logo,  $x_i = \text{sen}\theta_i$  e  $y_i = R \text{cos}\theta_i$  ou em coordenadas polares tem-se  $P_i = (R, \theta_i)$ , considerando  $R$  a distância da origem  $O$  até  $P_i$  (JAMMALAMADAKA; SENGUPTA, 2001).

### 3.2.4. Direção Média para dados circulares

Jammalamadaka e Sengupta (2001) ressaltam que, para definir a direção média ou o vetor médio de um determinado conjunto de dados circulares unimodais, não é correto utilizar a média aritmética dos ângulos.

Para ilustrar este cálculo do ângulo médio em dados circulares, calculou-se a média aritmética dos ângulos, considerando uma amostra de três direções dados pelos ângulos  $80^\circ$ ,  $350^\circ$  e  $50^\circ$  (Figura 5).

Ao realizar os cálculos encontraremos  $1/3(80 + 350 + 50) = 160^\circ$ , que não é valor aceitável, pois se espera que o ângulo médio assuma um valor entre  $0^\circ$  e  $50^\circ$  (Figura 5). Por outro lado, se substituirmos o ângulo  $350^\circ$  pelo seu ângulo correspondente igual  $-10^\circ$ , então o valor do ângulo médio será  $1/3(80 + (-10) + 50) = 40^\circ$ . A partir desse exemplo, fica claro que a média aritmética para ângulos não é uma estatística adequada para representar a direção média.

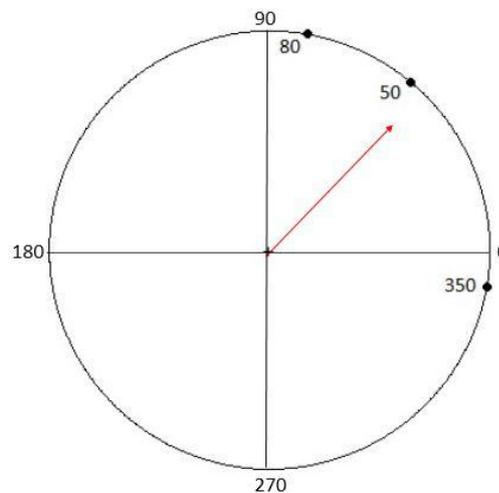


Figura 5 Representação Gráfica de uma amostra circular e vetor média (setas).  
Fonte: Elaborado pelo autor (2019)

De acordo com Batschelet (1981), a estatística apropriada para o cálculo da direção média dos dados circulares é diferente da utilizada para dados lineares. Primeiramente, considera-se um círculo unitário e, para cada observação dos dados circulares, representa-o pelo ponto  $P_i$  ( $i = 1, \dots, n$ ) nesta circunferência. O autor menciona o valor de massa denominado de  $M_i$  ( $i = 1, \dots, n$ ) para seu respectivo ponto  $P_i$  e, sendo associado o valor da massa para cada ponto de uma amostra tamanho  $n$ , em uma circunferência unitária, então tem-se  $M_1 = M_2 = \dots = M_n = M$ . Na Figura 4b é representada uma amostra de tamanho ( $n=3$ ), cujos os pontos estão localizados em uma circunferência unitária. Desta maneira, tem-se  $M_1 = M_2 = M_3 = M$ , podendo, assim, encontrar o centro da massa ( $C$ ) (Figura 4b) chamado de centro gravitacional. Caso o centro for diferente da origem, então a direção  $\overrightarrow{OC}$  define a direção média amostral.

Batschelet (1981) definiu duas maneiras de determinar o vetor média ( $\bar{\theta}$ ), pela álgebra vetorial e pelas funções trigonométricas. Por meio da álgebra vetorial, tem-se que para cada ponto  $P_i$ , existe um vetor de comprimento unitário  $v_i$  ( $i = 1, \dots, n$ ).

Ao considerar  $n$  pontos ( $P_1, \dots, P_n$ ) de massas  $M_1, M_2, \dots, M_n$ , e, conseqüentemente,  $n$  vetores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ , respectivamente, tem-se:

$$\mathbf{m} = \frac{1}{\sum_{i=1}^n M_i} (M_1 \mathbf{v}_1 + M_2 \mathbf{v}_2 + \dots + M_n \mathbf{v}_n), \quad \text{Eq. (3)}$$

em que  $\mathbf{m}$  é o vetor médio da amostra.

Ao assumir que  $M_1 = M_2 = \dots = M_n$ , pois a norma dos vetores  $\|\mathbf{v}_i\|=1$  ( $\forall i = 1, \dots, n$ ), obtém-se  $\sum_{i=1}^n M_i = nM$ . Simplificando a Equação 3, tem-se:

$$\mathbf{m} = \frac{1}{n} (\mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_n). \quad \text{Eq. (4)}$$

Desta forma, o vetor média de um conjunto de dados circular é definido pela Equação 5:

$$\mathbf{m} = \frac{\sum_{i=1}^n M_i}{n}. \quad \text{Eq. (5)}$$

Também se define algebricamente o comprimento do vetor resultante  $\bar{R}$  e o comprimento do vetor média ( $R$ ) (Figura 6), sendo expressos por  $\|\sum_{i=1}^n \mathbf{v}_i\|$  e  $\|\mathbf{m}\|$ , respectivamente. Como essas medidas são definidas em uma circunferência de um círculo unitário, tem-se que  $0 \leq \bar{R} \leq n$  e  $0 \leq R \leq 1$ .

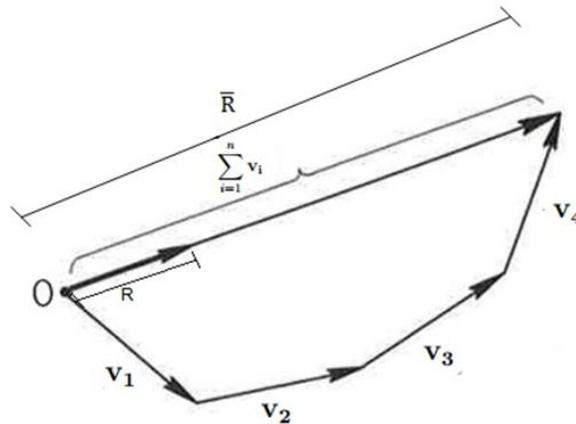


Figura 6 Representação do vetor resultante ( $\bar{R}$ ) e comprimento do vetor média ( $R$ ).

Fonte: Batschelet (1981)

O comprimento do vetor média ( $R$ ) pode ser definido também utilizando as funções trigonométricas como descrito pela Equação 6 (BESKOW et al., 2014).

$$R = \sqrt{\bar{x}^2 + \bar{y}^2}, \quad \text{Eq. (6)}$$

em que  $\bar{x}$  e  $\bar{y}$  são as coordenadas retangulares do centro de massa  $C$ , descrito pela Equação 7.

Na Figura 7 têm-se diferentes exemplos de amostras de dados circulares, todas com tamanhos de  $n=5$ . Constata-se que com o aumento da dispersão dos pontos na circunferência do círculo unitário (Figuras 7a a 7f), o comprimento do vetor média diminui (BATSCHULET, 1981), sendo que o valor de  $R$  é adimensional, indicando a uniformidade da ocorrência de determinado evento, em um período considerado. Seu valor varia de 1 (indicando assim o

máximo de concentração dos dados circulares, Figura 7a) a 0 (quando não ocorre uma concentração em torno de uma única direção Figura 7f) (BESKOW, 2014).

Para determinar o vetor média aplicando as funções trigonométricas, Batschelet (1981), Mardia e Jupp (1999) e Jammalamadaka e Sengupta (2001) consideraram o sistema de coordenadas retangulares. Logo, sejam  $x_i$  e  $y_i$  as componentes retangulares do vetor unitário  $v_i$  ( $i = 1, \dots, n$ ), expressas por  $x_i = \cos\theta_i$  e  $y_i = \sin\theta_i$ .

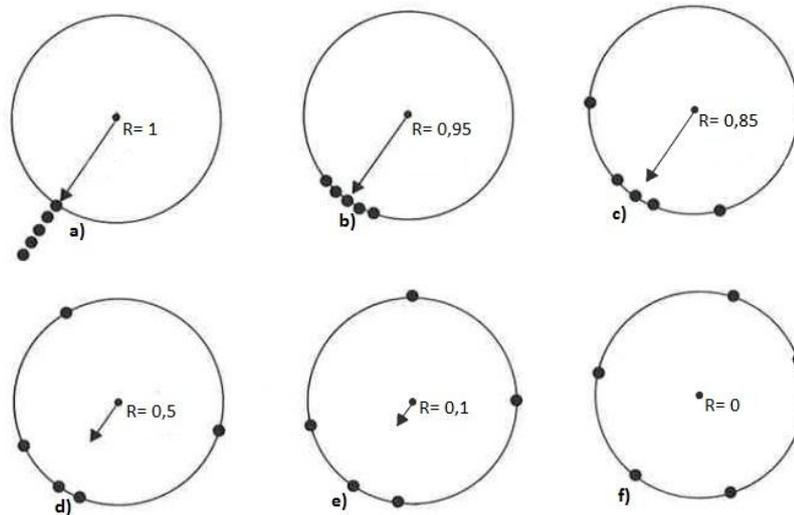


Figura 7 Variação do comprimento do vetor média em diferentes amostras de dados circulares, com tamanho da amostral  $n=5$  (seta representa o vetor média).

Fonte: Batschelet (1981)

Logo, o ponto  $P_i$  pode ser escrito como  $P_i = (\sin\theta_i, \cos\theta_i)$ , podendo, desta maneira, denotar  $(\bar{x}, \bar{y})$  as coordenadas retangulares do centro de massa  $C$ , tal que (Equação 7):

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) \quad \text{e} \quad \bar{y} = \frac{1}{n} (y_1 + y_2 + \dots + y_n). \quad \text{Eq. (7)}$$

Substituindo  $x_i$  e  $y_i$  nas Equações descritas em 7 têm-se (Equação 8):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n \cos\theta_i \quad \text{e} \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n \sin\theta_i \quad \text{Eq. (8)}$$

Desta maneira é possível definir a média circular ( $\bar{\theta}$ ) para  $R > 0$ , como descrito pela Equação 9 (JAMMALAMADAKA; SENGUPTA, 2001).

$$\bar{\theta} = \begin{cases} \arctan\left(\frac{\bar{y}}{\bar{x}}\right), \text{ se } \bar{x} > 0, \bar{y} \geq 0 \\ \frac{\pi}{2} \text{ se } \bar{x} = 0, \bar{y} > 0 \\ \left(\arctan\frac{\bar{y}}{\bar{x}}\right) + \pi, \text{ se } \bar{x} < 0 \\ \left(\arctan\frac{\bar{y}}{\bar{x}}\right) + 2\pi, \text{ se } \bar{x} \geq 0, \bar{y} < 0 \\ \text{Indefinido, se } \bar{x} = 0 \bar{y} = 0 \end{cases} \quad \text{Eq. (9)}$$

A função inversa da tangente ( $\arctan$ ) é definida no intervalo  $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$  pois, assim, tem-se um único valor para qualquer ângulo  $\theta$ .

Para um conjunto de dados circulares agrupados tem-se que as coordenadas cartesianas  $(\bar{x}, \bar{y})$  são definidas como  $\bar{x} = \frac{1}{n} \sum_{i=1}^n f_i \cos \theta_i$  e  $\bar{y} = \frac{1}{n} \sum_{i=1}^n f_i \sin \theta_i$ , com  $f_i$  sendo a frequência angular observada,  $n$  é o tamanho da amostra dos dados circular,  $\theta_i$  é o ângulo angular, em que  $i = 1, \dots, n$  (ZAR, 2010). No caso em que  $\bar{x} = 0$ ,  $\bar{y} = 0$  e  $R = 0$ , então o vetor média é igual ao vetor nulo (BATSCHELET, 1981).

Lee, Kwon e Kim (2011) desenvolveram uma pesquisa com a finalidade de avaliar a variabilidade temporal do evento de precipitação extrema na Coreia do Sul, considerando uma série temporal de 30 anos. Nesta pesquisa, verificaram a ocorrência de máximos diários e mensais, por meio da estatística circular. Como exemplo, estes autores identificaram que na região Sul houve um pico mensal de ocorrência da precipitação no início do mês de agosto, e na região central, este pico mensal ocorreu em meados do mês de agosto. Isto ocorreu, pelo fato do deslocamento da chuva de maneira transitória do Sul para o norte da Coreia do Sul durante as monções. Por meio desta análise de ocorrência de picos, os autores concluíram que é possível estabelecer planos futuros de política para a gestão de recurso hídrico para toda a Coreia do Sul.

### 3.3. Análise de Dados Espaciais: Estatística espacial de área

Entender a distribuição espacial de dados proveniente de fenômenos ocorridos no espaço é um grande desafio hoje para a elucidação de questões centrais, em diversas áreas do conhecimento (DRUCK et al., 2004).

Um dos pioneiros a utilizar a análise espacial em um determinado evento foi o médico John Snow (em 1854): no intuito de elucidar o caso da epidemia da cólera em Londres, foi construído um mapeamento dos casos de óbitos da região da localização da bomba Broad Street. Nesse estudo, notou-se que 58% das pessoas que utilizavam a água dessa bomba eram contaminadas e desenvolviam a doença e que apenas 7% dos que não bebiam a água

da bomba desenvolviam a doença, ocorrendo 700 óbitos em um raio de 250 metros em torno da bomba de Broad Street (NEWSON, 2006).

A dependência espacial segue o teorema fundamental da ciência regional, definido como: observações próximas uma da outra tem maior grau de dependência espacial do que aquelas mais distantes, ou seja, a dependência espacial entre observações diminui à medida que a distância entre as observações aumenta (LESAGE, 1999).

De acordo com Câmara (2004), a classificação dos dados espaciais pode ser considerada de três tipos:

- **Eventos ou Padrões Pontuais:** São os fenômenos que expressam a sua ocorrência por meio de pontos localizados no espaço, sendo denominados de processos pontuais, como, por exemplo: eventos relacionados à localização de crimes, ocorrências de doenças e localização de espécies vegetais;
- **Superfícies Contínuas:** São os tipos de dados estimados a partir de um conjunto de amostras de campo, que podem estar distribuídas de maneira regular ou irregular. Este tipo de dados é resultado de levantamento de recursos naturais, como mapa geológicos, topográficos, ecológicos, entre outros;
- **Áreas com Contagens e Taxas Agregadas:** Dados espaciais relacionados a levantamentos populacionais, como censos e estatísticas de saúde, que se referem a indivíduos localizados em pontos específicos do espaço. Por razões de confidencialidade, estes dados são agrupados em unidades de análise, por meio de polígonos fechado como, por exemplo, os setores censitários e os municípios.

Em relação ao desenvolvimento de modelagem da estatística espacial, algumas técnicas podem ser aplicadas, como a representação dos dados representados espacialmente por meio de mapas.

Outra etapa da análise exploratória dos dados espaciais que melhor descreve os dados em estudo é a análise de autocorrelação espacial, podendo ser utilizado o Índice Moran Global ( $I$ ) e o Local (LISA) (DRUCK et al., 2004).

Para análise da autocorrelação espacial das variáveis, estando estas localizadas em unidade amostrais, os autores Wartenberg (1985) e Câmara et al. (2004) mencionam que deve ser utilizado o Índice de Moran Global ( $I$ ) dado pela Equação 10:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n z_i z_j w_{ij}}{S_0 \sum_{i=1}^n z_i^2}, \quad \text{Eq. (10)}$$

em que  $n$  é o tamanho da amostra,  $z_i = (x_i - \bar{x})$  e  $z_j = (x_j - \bar{x})$  para  $i \neq j$  com  $i, j = 1, \dots, n$ , são os valores observados nas populações de  $i$  e  $j$  centradas na média  $\bar{x}$  da variável  $x$  em estudo;  $w_{ij}$  são os elementos de  $\mathbf{W}$  que é uma matriz simétrica de peso espacial, na qual seus elementos representam a proximidade entre as regiões  $i$  e  $j$ . Se as observações apresentam fronteira comum, então  $w_{ij} = 1$ ; e quando não possuem fronteira comum,

então  $w_{ij} = 0$ . Além disso,  $S_0$  é o somatório dos elementos de  $w_{ij}$ .

Para a utilização da matriz de proximidade, deve-se primeiramente definir a sua estrutura de contiguidade binária, ou seja, as relações existentes entre as regiões  $i$  e  $j$  que estão localizadas na  $i$ -ésima linha (unidade de área analisada) e na  $j$ -ésima coluna (unidade vizinha) da matriz  $\mathbf{W}$  (MACEDO; SIMÕES, 1998).

De acordo com Anselin (2005), existem alguns critérios de contiguidade para a definição da matriz  $\mathbf{W}$ , tais como:

- i) Contiguidade Torre (Rook):  $w_{ij} = 1$  para locais que compartilham um lado com a região de interesse e, caso contrário, tem-se que  $w_{ij} = 0$ , com  $i, j = 1, \dots, n$  e  $i \neq j$ ;
- ii) Contiguidade Bispo (Bishop):  $w_{ij} = 1$  para locais que compartilham o mesmo vértice, e, caso contrário, tem-se que  $w_{ij} = 0$ , com  $i, j = 1, \dots, n$  e  $i \neq j$ ;
- iii) Contiguidade da Rainha (Queen):  $w_{ij} = 1$  para locais  $i$  e  $j$  que compartilham um lado ou um vértice, e caso contrário, tem-se  $w_{ij} = 0$ , com  $i, j = 1, \dots, n$  e  $i \neq j$ .

Outro conceito importante em análise espacial de área é o de matriz de pesos espacial padronizada por linha, definida pela seguinte equação (ALMEIDA, 2004):

$$w_{ij}^p = \frac{w_{ij}}{\sum_{j=1}^n w_{ij}}, \quad \text{Eq. (11)}$$

em que:  $w_{ij}^p$  é o valor padronizado de  $w_{ij}$ , que é um elemento da matriz  $\mathbf{W}$  padronizada por linha. Assim, nesta matriz, a soma dos elementos das suas linhas é igual a um, facilitando vários cálculos de índices de autocorrelação espacial (CÂMARA et al., 2004).

De acordo com Câmara et al. (2004) e Anselin, Tita e Gorr (2000), os valores positivos do Índice de Moran (entre 0 e +1) indicam a existência de autocorrelação espacial direta e clustering, enquanto que os valores negativos (entre 0 e -1) indicam a autocorrelação espacial inversa e outliers.

Outra maneira de visualização da dependência espacial é o diagrama de espalhamento do Índice de Moran Global (Figura 8), que proporciona os elementos necessários para comparação dos valores normalizados (eixo  $x$ ) dos atributos em cada área com a média dos seus vizinhos (eixo  $y$ ), sendo equivalente ao coeficiente de regressão linear (CÂMARA et al., 2004).

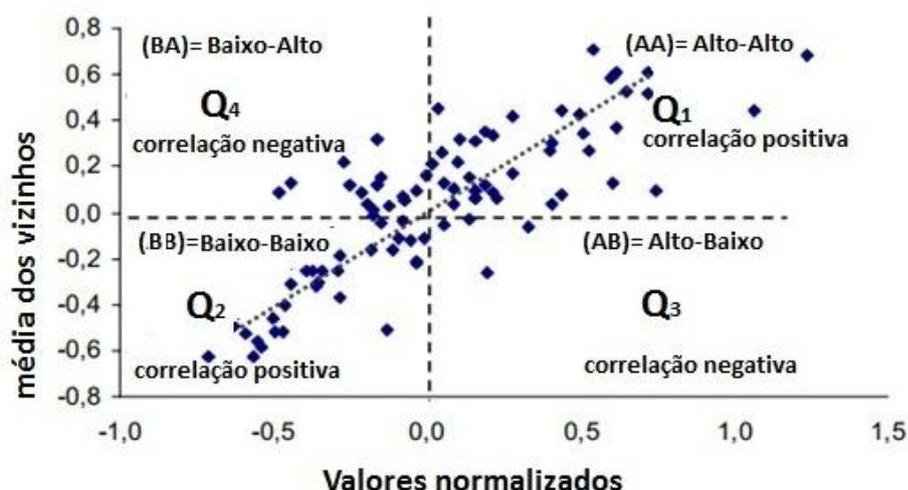


Figura 8 Representação Gráfica do Diagrama do Índice de Moran  
 Fonte: Adaptado de Câmara et al. (2004)

De acordo com Luzardo, Castañeda Filho e Rubin (2017) cada quadrante do gráfico bidimensional ( $Q_1$ ,  $Q_2$ ,  $Q_3$  e  $Q_4$ ) é definido como:

- $Q_1$  : É o quadrante que contém os casos onde o valor do atributo normalizado de cada polígono e o valor médio nos polígonos vizinhos são maiores do que a média global (denominado como AA ou Alto-Alto);
- $Q_2$  : Como os casos em que o valor do atributo normalizado de cada polígono e o valor médio nos polígonos vizinhos são menores do que a média global (denominado como BB ou Baixo-Baixo);
- $Q_3$  : É o quadrante que contém os casos em que o valor do atributo normalizado de cada polígono está acima da média geral e o valor médio dos polígonos vizinhos estão abaixo da média global (denominado como AB ou Alto-Baixo);
- $Q_4$  : Contém os casos onde o valor do atributo normalizado de cada polígono está abaixo da média global e o valor médio do atributo nos polígonos vizinhos está acima da média global (denominado como BA ou Baixo-Alto).

Quando a maioria dos pontos estão concentrados nos quadrantes  $Q_1$  e  $Q_2$  tem-se uma autocorrelação espacial positiva, indicando um agrupamento espacial de valores semelhantes, e se os quadrantes  $Q_3$  e  $Q_4$  tiverem a maior quantidade de pontos, isso indica uma autocorrelação espacial negativa, que corresponde a uma aglomeração espacial de valores dissimilares. Já no caso que os pontos estiverem igualmente distribuídos pelos quatro quadrantes, isto indica uma ausência de autocorrelação (ANSELIN, 1996).

Os indicadores globais de autocorrelação espacial como, por exemplo, o Índice de Moran, fornecem um valor único com medida de associação espacial para todo o conjunto de dados em estudo, sendo útil para a caracterização da região como um todo. Mas quando analisam-se várias áreas, é muito provável que ocorram diferentes regimes de associação espacial e que apareçam máximos locais de autocorrelação espacial, onde a dependência

espacial é mais apresentada (CÂMARA et al., 2004). Desta maneira, muitas vezes é desejável examinar padrões de forma mais detalhada, utilizando indicadores de associação espacial que possam relacionar as diferentes localizações de uma variável distribuída espacialmente, podendo utilizar o Índice de Moran Local (LISA), que produz um valor específico para cada área (LUZARDO; CASTAÑEDA FILHO; RUBIM, 2017).

O Índice Local de Moran ( $I_i$ ) é expresso pela Equação 12 (ANSELIN, 1996; ARAÚJO URIBE-OPAZO; JOHANN, 2014).

$$I_i = \frac{x_i - \mu}{\sigma_0^2} \sum_{j=1}^n w_{ij} (x_j - \mu), \quad \text{Eq. (12)}$$

em cada área  $i$ , com  $i = 1, \dots, n$ ,  $x_i$  e  $x_j$  são as observações de uma determinada variável nos locais de  $i$  e  $j$ , respectivamente,  $\mu$  é a média populacional e  $\sigma_0^2$  a variância populacional da variável da área em estudo e  $w_{ij}$  são os elementos de  $\mathbf{W}$  sendo uma matriz simétrica de peso espacial, na qual seus elementos representam a proximidade entre as regiões  $i$  e  $j$ .

Os valores positivos encontrados no índice de Moran Local significam a existência de agrupamentos espaciais com valores similares (alto ou baixo) e, para valores negativos, estes significam que existem agrupamentos espaciais com valores diferentes entre as regiões (ARAÚJO; URIBE-OPAZO; JOHANN, 2014).

Para a análise de correlação espacial entre duas variáveis aleatórias ( $X$  e  $Y$ ) em estudo, cada uma obtida nas mesmas  $n$  regiões, utiliza-se o Índice de Moran Global bivariado ( $I_{xy}$ ), definido na Equação 13 (ANSELIN, 1996; ARAÚJO; URIBE-OPAZO; JOHANN, 2014).

$$I_{xy} = \frac{\sum_{i=1}^n \sum_{j=1}^n u_i z_j w_{ij}}{S_0 \sqrt{S_u^2} \sqrt{S_z^2}}, \quad \text{Eq. (13)}$$

em que  $n$  é o número de observações,  $u_i = (x_i - \bar{x})$  e  $z_j = (y_j - \bar{y})$ , com  $i \neq j$  e  $i, j = 1, \dots, n$  são os valores observados centrados nas médias das variáveis  $X$  e  $Y$ , respectivamente;  $w_{ij}$  são os elementos de  $\mathbf{W}$ ;  $S_0$  é o somatório dos elementos de  $w_{ij}$  e  $S_u^2$ ,  $S_z^2$  são as variâncias de  $X$  e  $Y$ , respectivamente.

O objetivo desse índice é verificar se os valores de uma variável observada em uma região têm relação com os valores de outra variável observada em outra região vizinha, e sua interpretação é análoga à estatística do Índice de Moran univariada (CÂMARA et al., 2004).

Estes autores ressaltam que, após o cálculo do Índice de Moran, é importante estabelecer sua validade estatística, ou seja, se os valores medidos representam correlação espacial significativa. Para isso, pode-se analisar associando o Índice à distribuição de probabilidade ou pelo teste de pseudo-significância, que não necessita da pressuposição da distribuição.

A estatística do Índice de Moran é uma medida de dependência espacial, que possui um valor esperado  $E(I) = [1/(n - 1)]$  (VIDIGAL; VIDIGAL; PARRÉ, 2018). Se os valores do índice de I de Moran excederem ao valor esperado, tem-se a indicação da existência de uma correlação positiva. Por outro lado, se os valores do índice de Moran forem menores que o valor esperado, tem-se que há uma autocorrelação espacial negativa (BETARELLI JUNIOR; ALMEIDA, 2009).

Prudente et al. (2014) realizaram uma análise da distribuição espacial dos valores da produção municipal da soja no estado do Paraná, para sete anos-safras, de 2003/2004 a 2009/2010, por meio também dos índices de Moran Global. Os autores demonstraram como os municípios estão relacionados espacialmente quanto à produção da soja, permitindo a indicação de clusters de alta e baixa produção, identificando os principais produtores de soja do estado.

### 3.4. Comparação de mapas temáticos

De acordo com Uribe-Opazo e Rojas (2015), existem diversos procedimentos de comparação entre dois mapas temáticos, sendo um deste mapas, denominado como mapa referência (relacionado aos pixels das colunas) e o outro como mapa modelo (relacionado aos pixels das linhas). Os autores definem também a matriz de erros (Tabela 3), em cada elemento da matriz, representa a quantidade dos pixels pertencente à classe  $C_i$  ( $i = 1, \dots, k$ ) do mapa modelo e  $C_j$  ( $j = 1, \dots, j$ ) do mapa referência. A diagonal principal (quando  $i = j$ ) são os casos em que o pixel foi classificado na mesma classe em ambos os mapas, e os elementos fora da diagonal principal desta matriz representam o pixel cujas as classificações foram diferentes entre os dois mapas.

Tabela 3 Matriz de erros de ordem k x k

		Mapa de referência				
classes		$C_1$	$C_2$	...	$C_k$	Total
Mapa modelo	$C_1$	$n_{11}$	$n_{12}$	...	$n_{1k}$	$n_{1+}$
	$C_2$	$n_{21}$	$n_{22}$	...	$n_{2k}$	$n_{2+}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$C_k$	$n_{k1}$	$n_{k2}$	...	$n_{kk}$	$n_{k+}$
Total		$n_{+1}$	$n_{+2}$	...	$n_{+k}$	N

k: número de classes;  $C_i$ : classe  $i$ ;  $n_{i+}$ : total de pixels da classe  $i$  do mapa modelo;  $C_j$  do mapa classe  $j$ ;  $n_{j+}$ : total de pixels da classe  $j$  do mapa de referência; N: número total de pixels

A partir da matriz de erros, é possível calcular as seguintes medidas de similaridade entre os mapas temáticos: Exatidão Global e índices de concordância Kappa e Tau (CONGALTON, 1991).

A Exatidão Global (EG) (Equação 14) é uma medida de avaliação a similaridade entre o mapa de referência e o mapa modelo, em que, de acordo com Anderson et al. (1976), o nível mínimo de precisão entre os mapas o valor encontrado deve ser igual ou maior que 0,85.

$$EG = \sum_{i=1}^k \frac{n_{ii}}{N} \quad \text{Eq. (14)}$$

em que  $n_{ii}$  é a diagonal principal da matriz de erro e  $N$  o número total de pixels.

De acordo com Congalton (1991), o índice de concordância Kappa ( $K$ ) é utilizado para mensurar a exatidão das classificações temáticas, sendo uma medida de concordância ou precisão entre dois mapas (mapa modelo e de referência), expresso pela seguinte Equação 15:

$$K = \frac{N \sum_{i=1}^k n_{ii} - \sum_{i=1}^r (n_{i+} * n_{+i})}{N^2 - \sum_{i=1}^r (n_{i+} * n_{+i})} \quad \text{Eq. (15)}$$

em que  $k$  é o número de classes expressas da matriz de erros;  $n_{ii}$  é o número de pixels da classe  $i$  e da classe  $j$ ;  $n_{i+}$  e  $n_{+i}$  são as quantidades totais da linha  $i$  e da coluna  $j$ , respectivamente, e  $N$  é o total de pixels. Segundo Landis e Koch (1977), a qualidade da classificação do índice Kappa e Tau é definida como mostra a Tabela 4:

Tabela 4 Qualidade da classificação do Índice Kappa ( $K$ ) e Tau ( $\tau$ )

Índice Kappa e Tau	Qualidade da classificação
0,00	Péssimo
0,01 a 0,20	Ruim
0,21 a 0,40	Razoável
0,41 a 0,60	Boa
0,61 a 0,80	Muito Boa
0,81 a 1,00	Excelente

Fonte: Landis e Koch (1977)

Por fim, o índice Tau ( $\tau$ ) é similar ao índice Kappa, seguindo a mesma classificação (Tabela 5), sendo calculado pela Equação 16 (MA; REDMOND, 1995).

$$\tau = \frac{\sum_{i=1}^k n_{ii} - p_i}{1 - p_i} \quad \text{Eq. (16)}$$

em que  $n_{ii}$  é o número de pixels na linha  $i$  e na coluna  $j$  e  $p_i = \frac{1}{k}$  onde  $k$  é o número de classe da matriz de erro.

### 3.5. Análise Estatística Multivariada

Dentre as diversas análises estatísticas multivariadas estão: a análise de agrupamento, a análise de componente principal (PCA) e a análise de discriminante.

A análise de agrupamento tem como objetivo dividir os elementos da amostra ou da população em grupos, sendo que cada grupo tem elementos similares entre si de acordo com as variáveis (características) que neles foram medidos (MANLY, 2008). Este autor menciona também que existem duas técnicas de agrupamento: a técnica não hierárquica e a hierárquica, sendo que na técnica não hierárquica pode-se determinar inicialmente o número de clusters  $k$  desejado.

Segundo Mingoti (2013), o k-Means é um dos métodos de agrupamento não hierárquicos e consiste em alocar cada elemento amostral no cluster cujo centroide (vetor de médias amostral) é o mais próximo do vetor de valores observado do respectivo elemento. Este método é originalmente composto pelos seguintes passos:

- 1) Primeiramente escolhem-se o  $k$  centroides, denominado de protótipo, iniciando assim o processo de partição;
- 2) Cada elemento do conjunto de dados é comparado com cada centroide inicial, por meio de uma medida de distância (geralmente a distância Euclidiana), sendo que cada elemento será alocado ao grupo cuja distância for menor;
- 3) Depois disso, recalculam-se os valores dos centroides para os novos grupos formados e repete-se o passo 2, considerando os centroides destes novos grupos;
- 4) Deve-se repetir o passo 2 e 3 até que nenhuma realocação de elementos seja necessária.

A análise de componentes principais (PCA) é uma técnica da estatística multivariada que tem como objetivo determinar combinações lineares de  $p$ -variáveis ( $X_1, X_2, \dots, X_p$ ), produzindo, assim, índices  $Z_i$  (com  $i = 1, \dots, p$ ) não correlacionados na ordem de sua importância, denominados como componentes principais, tal que  $Var(Z_1) \geq Var(Z_2) \geq \dots \geq Var(Z_p)$ , em que  $Var(Z_i)$  é a variância de  $Z_i$  (MANLY, 2008).

De modo geral, deseja-se reduzir o número de variáveis originais a serem avaliadas, ou seja, a informação contida nas  $p$ -variáveis originais é substituída pela informação contida nos  $k$  ( $k < p$ ) componentes principais não correlacionados. A determinação dos números de componentes principais pode ser feita por meio da avaliação da proporção da variação total (MINGOTI, 2013).

Por sua vez, a análise de discriminante é uma técnica utilizada para a classificação de elementos de uma amostra ou de uma população, sendo necessária a determinação dos grupos a priori, baseado nas características gerais dos elementos, tendo como objetivo principal minimizar a classificação incorreta de um elemento amostral pertencer a uma população quando na realidade pertence a outra população (MINGOTI, 2013).

Nos casos que as matrizes covariâncias ( $\Sigma$ ) são diferentes, utiliza-se para a classificação dos elementos a função discriminante quadrática (Equação 17) (MINGOTI, 2013; DUTRA; BIAZI, 2008):

$$fd(x) = [(x - \mu_1)' \Sigma_1^{-1} (x - \mu_1)] - [(x - \mu_2)' \Sigma_2^{-1} (x - \mu_2)] + [\ln|\Sigma_1| - \ln|\Sigma_2|] \quad \text{Eq. (17)}$$

em que  $fd(x)$  é a função discriminante quadrática,  $x$  é o vetor aleatório ou de observações fixo;  $\mu_1$ ,  $\Sigma_1$  e  $\mu_2$ ,  $\Sigma_2$  vetor médio e matriz de covariância das populações  $\pi_1$  e  $\pi_2$ , respectivamente.

A regra de classificação consiste em que um elemento pertence à população  $\pi_1$  se  $fd(x)$  for menor que zero; caso contrário, pertence à população  $\pi_2$  (MANLY, 2008).

De acordo com Mardia, Kent e Biley (1995), existem vários procedimentos para a verificação da qualidade de uma função discriminante, dentre eles está o método da ressubstituição, que é baseado nos dados da Tabela 5. Nesta tabela tem-se que  $n_{ij}$  é o número de elementos pertencentes à população de origem  $i$  classificados pela função discriminante pertencentes à população  $j$  (com  $i, j = 1, 2$ ). Quando  $i = j$ , tem-se que a classificação do elemento na qual pertence a população é correto; caso contrário, é incorreta.

Tabela 5 Tabela genérica das frequências dos erros de classificação

		População classificação pela regra	
		1	2
População de origem	1	$n_{11}$	$n_{12}$
	2	$n_{21}$	$n_{22}$

Fonte: Mingoti (2013)

$n_{ij}$ : número de elementos da população  $\pi_i$  classificados na população  $\pi_j$ ;

Desta maneira é possível calcular as estimativas da probabilidade de ocorrência dos erros da população  $\pi_1$  e  $\pi_2$  pelas seguintes Equações:

$$\hat{P}_{\pi_1}(2/1) = \frac{n_{12}}{n_{11} + n_{12}} \quad \text{e} \quad \hat{P}_{\pi_2}(1/2) = \frac{n_{21}}{n_{21} + n_{22}} \quad \text{Eq. (18)}$$

Para a classificação da função de discriminante, tem-se que quanto menor é o valor obtidos nas estimativas, melhor é a função de discriminação (MINGOTI, 2013).

#### 4. MATERIAL E MÉTODOS

A metodologia da pesquisa foi dividida nas seguintes etapas: determinação da área de estudo, aquisição dos dados, distribuição espacial das variáveis no estado do Paraná e análise espacial (utilizando dos Índices de Moran univariado e bivariado), análise da ocorrência de sazonalidade e do cálculo do decêndio médio das medidas associadas às variáveis agrometeorológicas (por meio da estatística circular) (Figura 9).



Figura 9 Etapas do desenvolvimento da pesquisa  
Fonte: Elaborado pelo autor (2019)

##### 4.1 Área de estudo

A área de estudo compreende o estado do Paraná, estando situado na região sul do Brasil, limitado entre os paralelos 22°29'S e 26°43'S e os meridianos 48°2'W e 54°38'W (Figura 10). A escolha dos anos-safras em estudo foi feita a partir da análise da série histórica dos anos-safras, compreendidos entre 2000/2001 e 2015/2016 com relação à precipitação e à produtividade média. Nesta série histórica, observou-se que o ano-safra 2011/2012 que apresentou o menor valor médio de precipitação e um valor baixo de produtividade média, enquanto que o ano-safra 2013/2014 apresentou um valor médio de precipitação e valor de produtividade média próxima da média do período analisado. Desta forma, foram escolhidos nessa pesquisa estes dois anos-safra, pois correspondem a cenários diferenciados, um ano-safra de estiagem e outro com valores médios de precipitações.

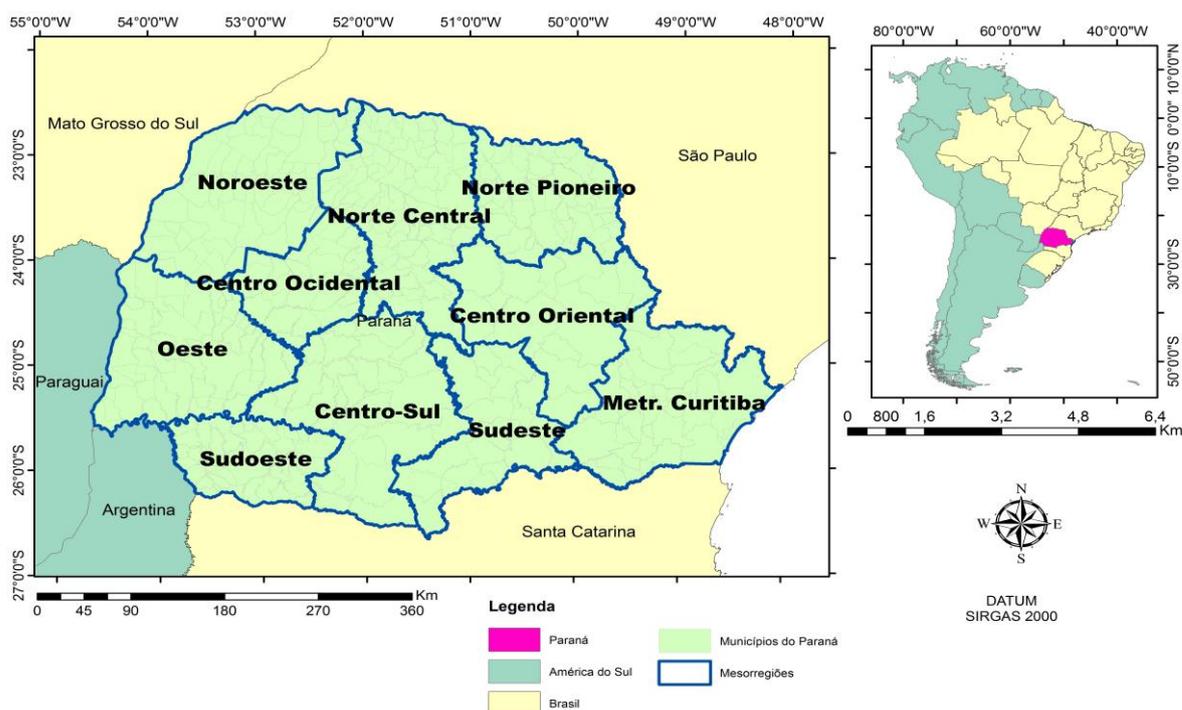


Figura 10 Área de estudo com as 10 mesorregiões do estado do Paraná  
Fonte: Elaborado pelo autor (2019)

## 4.2 Aquisição dos dados

### 4.2.1 Variáveis agrometeorológicas

Os dados das variáveis agrometeorológicas foram obtidos das estações virtuais (EV), por meio do modelo European Centre for Medium-Range Weather Forecasts (ECMWF), que é obtido de maneira gratuita no website do JRC (ECMWF, 2018), no formato shape, tendo uma resolução espacial em pixel com uma área aproximada de 25 x 25 km (Figura 12a) e uma resolução temporal decendial (cada 10 dias).

Para a utilização das variáveis agrometeorológicas, foi considerada a data do máximo desenvolvimento vegetativo (DMDV), de acordo com Becker et al. (2017), obtido pelo mapeamento da soja do ano-safra 2011/2012, realizado por Souza et al. (2015), e do ano-safra de 2013/2014, realizado por Grzegozewsk et al. (2017).

Dessa maneira, definiram-se duas variações decendiais em torno do DMDV, sendo elas: um decêndio antes até um depois do DMDV (1aDMDV1d) e dois decêndios antes até dois depois do DMDV (2aDMDV2d).

Para cada intervalo e para todo o ciclo da cultura, foi calculado o valor médio temperatura do ar [TMed](C) e o valor acumulado para as variáveis: precipitação pluvial [Prec](mm), radiação solar global [Rs] ( $\text{MJ m}^{-2} \text{dia}^{-1}$ ) e evapotranspiração potencial [ETp](mm) (Tabela 6).

Calcularam-se também os referidos valores destas variáveis mencionadas anteriormente, no decêndio do DMDV.

Tabela 6 Medidas associadas às variáveis agrometeorológicas

Sigla	Significado	Estádios Vegetativos <sup>(a)</sup>
Rs_DMDV	Rs na DMDV	R <sub>1</sub> à R <sub>3</sub>
ETp_DMDV	ETp na DMDV	R <sub>1</sub> à R <sub>3</sub>
DMDV_TMed	TMed na DMDV	R <sub>1</sub> à R <sub>3</sub>
DMDV_Prec	Prec na DMDV	R <sub>1</sub> à R <sub>3</sub>
Rs_1aDMDV1d	Rs acumulada no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
Rs_2aDMDV2d	Rs acumulada no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>
R_SRs_1aDMDV1d	Comprimento do vetor médio da Rs no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
R_SRsR_2aDMDV2d	Comprimento do vetor médio da Rs no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>
ETp_1aDMDV1d	ETp acumulada no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
ETp_2aDMDV2d	ETp acumulado no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>
R_ETp_1aDMDV1d	R da ETp no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
R_ETp_2aDMDV2d	R da ETp no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>
TMed_1aDMDV1d	Média da TMed no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
TMed_2aDMDV2d	Média da TMed no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>
R_TMed_1aDMDV1d	Comprimento do vetor médio da TMed no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
R_TMed_2aDMDV2d	Comprimento do vetor médio da TMed no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>
Prec_1aDMDV1d	Prec acumulada no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
Prec_2aDMDV2d	Prec acumulada no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>
R_Prec_1aDMDV1d	Comprimento do vetor médio da Prec no 1aDMDV1d	V <sub>4</sub> à R <sub>4</sub>
R_Prec_2aDMDV2d	Comprimento do vetor médio da Prec no 2aDMDV2d	V <sub>2</sub> à R <sub>5</sub>

Rs: radiação solar; DMDV: data do máximo de desenvolvimento vegetativo; ETp: evapotranspiração potencial; TMed: temperatura média; Prec: precipitação; 1aDMDV1d: um decêndio antes até um depois do DMDV; 2aDMDV2d: dois decêndios antes até dois depois do DMDV; R: comprimento do vetor média; V<sub>2</sub>: Segundo nó do estádios vegetativo; V<sub>4</sub>: Quarto nó do estádios vegetativo; R<sub>1</sub>: Início do florescimento; R<sub>3</sub>: início da formação de vagem; R<sub>4</sub>: Florescimento Pleno; R<sub>5</sub>: Vagem completamente desenvolvida.

<sup>(a)</sup>Fonte: Fehr e Caviness (1977); Embrapa (2007)

#### 4.2.2 Produtividade da soja

Os dados da produtividade média da soja foram obtidos para cada município do estado do Paraná, por meio de informações provenientes do Instituto Brasileiro de Geografia e Estatística (IBGE). Dos 399 municípios, não há informação sobre a produtividade para 38 municípios no ano-safra 2011/2012 e para 29 municípios para o ano-safra 2013/2014.

Para a determinação do valor da produtividade da soja nas mesmas localizações, ou seja, nas estações virtuais (EV) do ECMWF, calculou-se em cada EV a produtividade média ponderada (Figura 11), utilizando a Equação 19:

$$PMP = \frac{\sum_{j=1}^n a_j p_j}{\sum_{j=1}^n a_j}, \quad \text{Eq. (19)}$$

em que  $PMP$  é a produtividade média ponderada,  $a_j$  é a área do município dentro da EV do ECMWF e  $p_j$  é a produtividade do município, com  $j = 1, \dots, n$ , sendo  $n$  o número de municípios dentro da EV.

Além disso, os municípios que não apresentaram produtividade ( $p_j = 0$ ) foram desconsiderados do cálculo da  $PMP$ .

Na Figura 11c tem-se a representação de uma EV, como exemplo, a qual contém um percentual de área de três municípios.

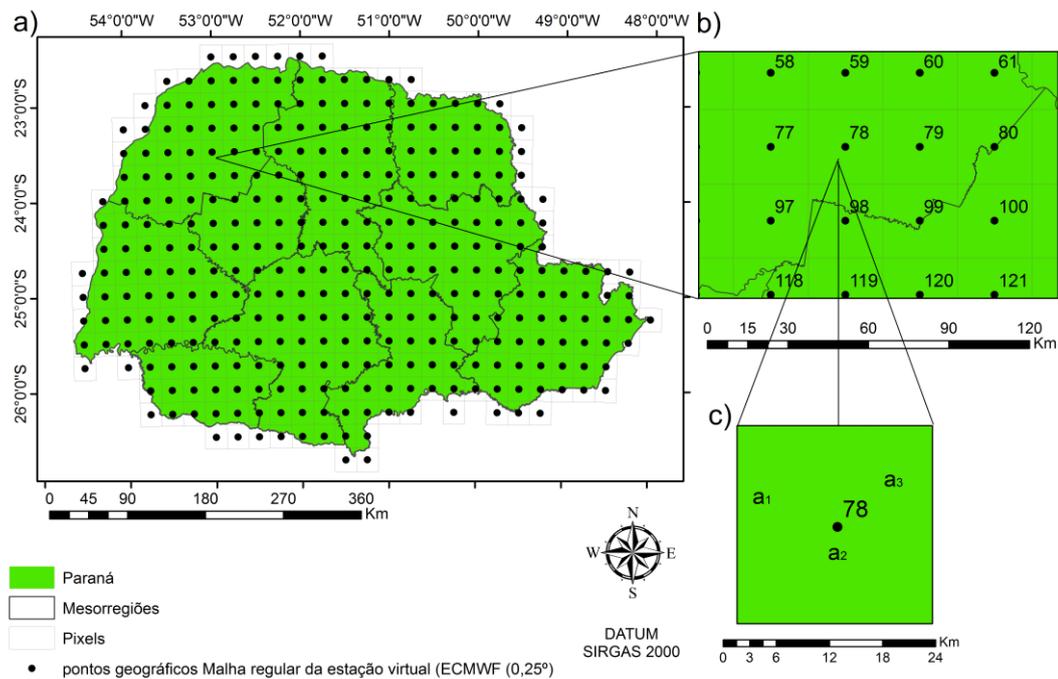


Figura 11 (a) e (b) Representações dos pontos geográficos da malha regular das estações virtuais do ECMWF; (c) Exemplo de uma EV selecionada com o percentual da área dos 3 municípios ( $a_1$ ,  $a_2$ ,  $a_3$ ) que a compõe.

Fonte: Elaborado pelo autor (2019)

### 4.3 Análise descritiva das variáveis em estudos

Realizou-se a análise da estatística descritiva linear, considerando as seguintes medidas: média, desvio padrão, coeficiente de variação, para as variáveis em estudos dos dois anos-safras. Realizou-se a distribuição espacial das variáveis utilizando o software ArcGis 10.0 (ESRI, 2015).

Na sequência, comparou-se os mapas da produtividade dos dois anos-safras, assim como os mapas dos decêndios do máximo desenvolvimento vegetativo (DMDV), por meio das medidas de similaridade entre os mapas temáticos: Exatidão Global (Equação 14), índice de concordância Kappa (Equação 15) e Tau (Equação 16).

### 4.3 Estatística Circular

Para as variáveis agrometeorológicas, nos intervalos decendiais dos anos-safras em estudo, calculou-se para cada EV as medidas da estatística circular: a média circular (Equação 8), que determinou o decêndio médio da maior ocorrência da variável, e o comprimento do vetor médio circular ( $R$ ), que determina a uniformidade da ocorrência das variáveis (Equação 6) (Tabela 6).

Além disso, realizou-se a representação gráfica, por meio do histograma circular, de algumas EV do ECMWF para a variável precipitação. Para a análise estatística descritiva e as representações gráficas dos dados circulares, utilizou-se o pacote circular do software estatístico gratuito R (R DEVELOPMENT CORE TEAM, 2018).

### 4.4 Análise Espacial de Área

Realizou-se a distribuição espacial da produtividade média ponderada nas EV do ECMWF, da cultura soja, com as medidas associadas às variáveis agrometeorológicas nos intervalos decendiais e das medidas das estatísticas circular (decêndio médio e  $R$ ) (Tabela 6). Para representar esta distribuição, utilizou-se o software ArcGis 10.0 (ESRI, 2015) para a construção de mapas em que suas legendas foram representadas por divisão interquartilica, considerando os dados dos anos-safras em estudo.

Em seguida, para realizar a análise espacial das variáveis em estudo, utilizou-se o Índice de Moran univariado e bivariado. Primeiramente, verificou-se a normalidade dos dados por meio do teste de Jones (1969), o qual avalia a normalidade pela distribuição dos estimadores de assimetria e curtose.

Para as variáveis que não apresentaram normalidade, utilizou-se a transformação dos dados, por meio da transformação de Johnson (Equação 20) (YOE; JOHNSON, 2000) e de Box Cox (Equação 21) (BOX; COX, 1964), realizado no software gratuito Action versão 3.3.2 (EQUIPE ESTATCAMP, 2018).

$$\Psi(\lambda, x) = \begin{cases} (x+1)^\lambda - 1/\lambda & (x \geq 0, \lambda \neq 0) \\ \log(x+1) & (x \geq 0, \lambda = 0) \\ -(x+1)^{2-\lambda} + \frac{1}{2-\lambda} & (x < 0, \lambda \neq 2) \\ -\log(-x+1) & (x < 0, \lambda = 2) \end{cases} \quad \text{Eq. (20)}$$

em que  $\lambda$  é o parâmetro de transformação,  $x$  é a variável observada e  $\Psi(\lambda, x)$  é variável transformada.

$$Y^{(\lambda)} = \begin{cases} \frac{y^\lambda}{\lambda}, & \text{se } \lambda \neq 0 \\ \log(y), & \text{se } \lambda = 0 \end{cases}, \quad \text{Eq. (21)}$$

em que  $\lambda$  é o parâmetro de transformação,  $y$  é a variável observada e  $Y^{(\lambda)}$  é a variável transformada.

Após verificar a normalidade das variáveis em estudo e realizar as transformações necessárias, foi definida a matriz de proximidade espacial **W** variando os seus critérios da matriz de peso: Rainha (Queen) e Torre (Rook), sendo escolhida a matriz de proximidade, de acordo com Colusso, Parré e Almeida (2013) e Montenegro e Betarelli Junior (2008).

Posteriormente, calculou-se a autocorrelação espacial da variável produtividade e as medidas associadas às variáveis agrometeorológicas, por meio do Índice de Moran Global Univariado (Equação 10).

Para visualizar a autocorrelação espacial das variáveis, em cada estação foram criados mapas de agrupamentos para apresentar os padrões espaciais, sendo suas classes definidas pelo diagrama de espalhamento do Índice de Moran Global, como: Alto-Alto, Baixo-Baixo, Baixo-Alto e Alto-Baixo.

Além disso, foi analisada a correlação espacial entre pares de variáveis, ou seja, entre a produtividade de cada ano-safra com as medidas associadas às variáveis agrometeorológicas (Tabela 6), por meio do Índice de Moran Global Bivariado (Equação 12). Para a análise espacial dos dados foram utilizados os softwares gratuitos Geoda 1.12 (ANSELIN; SYABRI; KHO, 2006) e ArcGIS 10.0 (ESRI, 2015).

#### **4.4 Análise Estatística Multivariada**

Realizou-se uma análise de agrupamento das estações virtuais do modelo ECMWF, considerando as variáveis: produtividade e as medidas associadas à temperatura, radiação solar, evapotranspiração potencial e precipitação para cada ano-safra (como também seus respectivos valores de  $R$ ). Para o agrupamento, utilizou-se o método K-means e o número de grupo foi definido pelo critério Calínski (CALÍNSKI; HARABASZ, 1974). Utilizou-se o software estatístico gratuito R (R DEVELOPMENT CORE TEAM, 2018).

Aplicou-se também a análise de componente principal (PCA) e a análise discriminante nas medidas associadas das variáveis agrometeorológicas e do valor de  $R$  apenas da variável precipitação, dos intervalos decendiais considerados, sendo nesta análise definida a função discriminante quadrática (Equação 17), verificando a qualidade desta por meio da Equação 19. Para a análise da estatística multivariada utilizou-se o software gratuito R (R DEVELOPMENT CORE TEAM, 2018).

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Análise descritiva da produtividade da soja e das variáveis agrometeorológicas

A distribuição espacial da produtividade da soja no estado do Paraná no ano-safra 2011/2012 (Figura 12a) apresentou um valor acima da média do estado ( $2,45 \text{ t ha}^{-1}$ ) para a maior parte das seguintes mesorregiões: Centro Oriental, Centro Sul, Sudeste e dos Campos Gerais (IBGE, 2018).

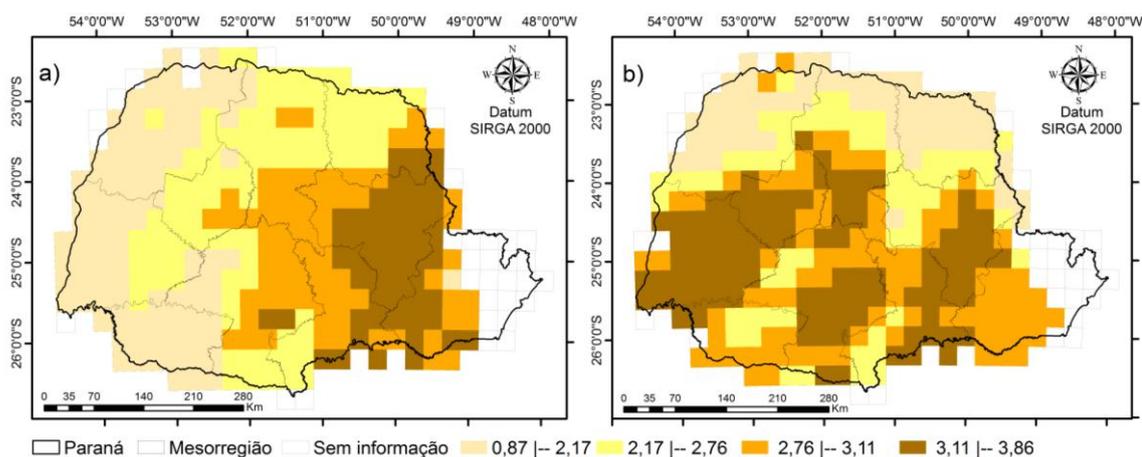


Figura 12 Mapa da distribuição de produtividade ( $\text{t ha}^{-1}$ ) dos anos-safra: (a) 2011/2012 e (b) 2013/2014

Os menores valores da produtividade estão localizados em grande parte das mesorregiões Noroeste, Oeste e Sudoeste. Essa tendência da distribuição espacial de valores de produtividade da soja está de acordo com os resultados obtidos pela EMBRAPA (2016).

Para o ano-safra 2013/2014 (Figura 12b), constatou-se que os menores valores de produtividade estão nas seguintes mesorregiões: Noroeste, Norte Pioneiro e parte do Norte Central. As demais mesorregiões apresentaram valores próximos ou acima da média do estado, que foi de  $2,86 \text{ t ha}^{-1}$  (IBGE, 2018).

Ao comparar as produtividades médias entre os anos-safra em estudos, por meio do teste Z, rejeitou-se a hipótese nula de que as médias são iguais, ou seja, as médias são estatisticamente diferentes a 5% de significância. Isto foi confirmado pelos valores das medidas: Exatidão Global ( $EG= 0,27$ ) e índices de concordância Kappa ( $K= 0,05$ ) e índice Tau ( $\tau = 0,05$ ), calculadas considerando o mapa de referência (produtividade do ano-safra de 2013/2014) e o mapa modelo (produtividade do ano-safra de 2011/2012). Segundo Anderson et al. (1976) e Krippendorff (1980), estas medidas evidenciam que os mapas que expressam a produtividade nos dois anos safra analisados apresentam uma relevante diferença.

Constatou-se que no ano-safra o 2013/2014, obtiveram-se maiores valores de produtividade, em relação ao ano-safra 2011/2012. Essa diferença na produtividade entre os anos-safra pode ter ocorrido pela falta da disponibilidade hídrica no ciclo da cultura no ano-

safras 2011/2012, afinal, essa é uma das principais causas de variabilidade da produtividade de um ano para outro (EMBRAPA SOJA, 1999, FERREIRA; COSTA; COSTA, 2007).

Realizou-se a estatística descritiva linear: a média, desvio padrão e coeficiente de variação (CV) para a média ponderada da produtividade (*PMP*) ( $t\ ha^{-1}$ ); e as medidas associadas às variáveis agrometeorológicas: temperatura média ( $^{\circ}C$ ), radiação solar ( $MJ\ m^{-2}\ dia^{-1}$ ) precipitação (mm) e evapotranspiração potencial (mm) dos anos-safras em estudos (Tabela 7). Quando comparada à *PMP* entre os anos-safras no estado, verificou-se que o ano-safra 2013/2014 teve um maior valor ( $2,92\ t\ ha^{-1}$ ) em relação ao ano-safra de 2011/2012 ( $2,58\ t\ ha^{-1}$ ), e para o coeficiente de variação, a *PMP* apresentou alta e moderada dispersão, respectivamente, para os dois anos (GOMES, 1990).

Tabela 7 Estatística descritivas para média ponderada da produtividade (PMP) da soja e as medidas associadas às variáveis agrometeorológicas

Variáveis	2011/2012			2013/2014		
	Média	Desvio padrão	CV (%)	Média	Desvio padrão	CV (%)
<i>PMP</i>	2,58 $t\ ha^{-1}$	0,69 $t\ ha^{-1}$	26,74	2,92 $t\ ha^{-1}$	0,56 $t\ ha^{-1}$	19,18
TMed_DMDV	22,23 $^{\circ}C$	2,07 $^{\circ}C$	9,31	24,01 $^{\circ}C$	2,39 $^{\circ}C$	9,95
TMed_1aDMDV1d	21,16 $^{\circ}C$	2,35 $^{\circ}C$	11,11	23,69 $^{\circ}C$	2,14 $^{\circ}C$	9,03
TMed_2aDMDV2d	21,90 $^{\circ}C$	2,24 $^{\circ}C$	10,23	23,75 $^{\circ}C$	2,03 $^{\circ}C$	8,55
Rs_DMDV	0,24 MJ	0,04 MJ	16,67	0,24 MJ	0,02 MJ	8,33
Rs_1aDMDV1d	0,68 MJ	0,06 MJ	8,82	0,73 MJ	0,04 MJ	5,48
Rs_2aDMDV2d	1,16 MJ	0,07 MJ	6,03	1,22 MJ	0,05 MJ	4,10
ETp_DMDV	41,00 mm	16,08 MJ	39,22	53,00 mm	7,16 mm	13,51
ETp_1aDMDV1d	129,38 mm	30,54 mm	23,60	160,00 mm	20,53 mm	12,83
ETp_2aDMDV2d	221,69 mm	46,13 mm	20,81	271,00 mm	34,41 mm	12,70
Prec_DMDV	64,00 mm	49,00 mm	76,56	54,50 mm	30,72 mm	56,37
Prec_1aDMDV1d	152,00 mm	70,42 mm	46,33	153,00 mm	54,42 mm	35,57
Prec_2aDMDV2d	237,00 mm	68,75 mm	29,01	246,00 mm	50,13 mm	20,38

DMDV= decêndio do máximo vegetativo; 1aDMDV1d: um decêndio antes até um depois do DMDV; 2aDMDV2d: dois decêndios antes até dois depois do DMDV; TMed\_DMDV: Temperatura média no DMDV; TMed\_1aDMDV1d: Temperatura média do 1aDMDV1d; TMed\_2aDMDV2d: Temperatura média do 2aDMDV2d; Rs\_DMDV: Radiação solar no DMDV; Rs\_1aDMDV1d: Radiação acumulada no 1aDMDV1d; Rs\_2aDMDV2d: Radiação acumulada no 2aDMDV2d; ETp\_DMDV: Evapotranspiração potencial no DMDV; S\_ETp\_1aDMDV1d: Evapotranspiração potencial acumulada no 1aDMDV1d; ETp\_2aDMDV2d: Evapotranspiração potencial acumulada no 2aDMDV2d; Prec\_DMDV: precipitação no DMDV; Prec\_1aDMDV1d: precipitação acumulada no 1aDMDV1d; Prec\_2aDMDV2d: precipitação acumulada no 2aDMDV2d; CV: coeficiente de variação.

Além disso, com exceção da precipitação no DMDV, as demais variáveis apresentaram menores valores médios no ano-safra de 2011/2012, quando comparado com o ano-safra de 2013/2014 (Tabela 7). Considerando o coeficiente de variação das medidas associadas às variáveis temperatura média, radiação solar nos dois anos-safras e evapotranspiração potencial no ano-safra 2013/2014, estas apresentaram baixa moderada dispersão. Por outro lado, para as medidas associadas às variáveis evapotranspiração

potencial no ano-safra 2011/2012 e precipitação nos dois anos-safras, estas apresentaram respectivamente, alta e muito alta dispersão, de acordo com a classificação de Gomes (1990).

## 5.2 Análise dos decêndios do máximo desenvolvimento vegetativo (DMDV) da soja no estado do Paraná

No ano-safra 2011/2012, o DMDV da soja no estado do Paraná variou entre o 2º decêndio de novembro até o 3º decêndio de janeiro (Figura 12a), mas com maior ocorrência nos 1º e 2º decêndios de janeiro (27,93 e 32,43%, respectivamente), principalmente, nas mesorregiões com maiores rendimentos de produtividade (Figura 12a).

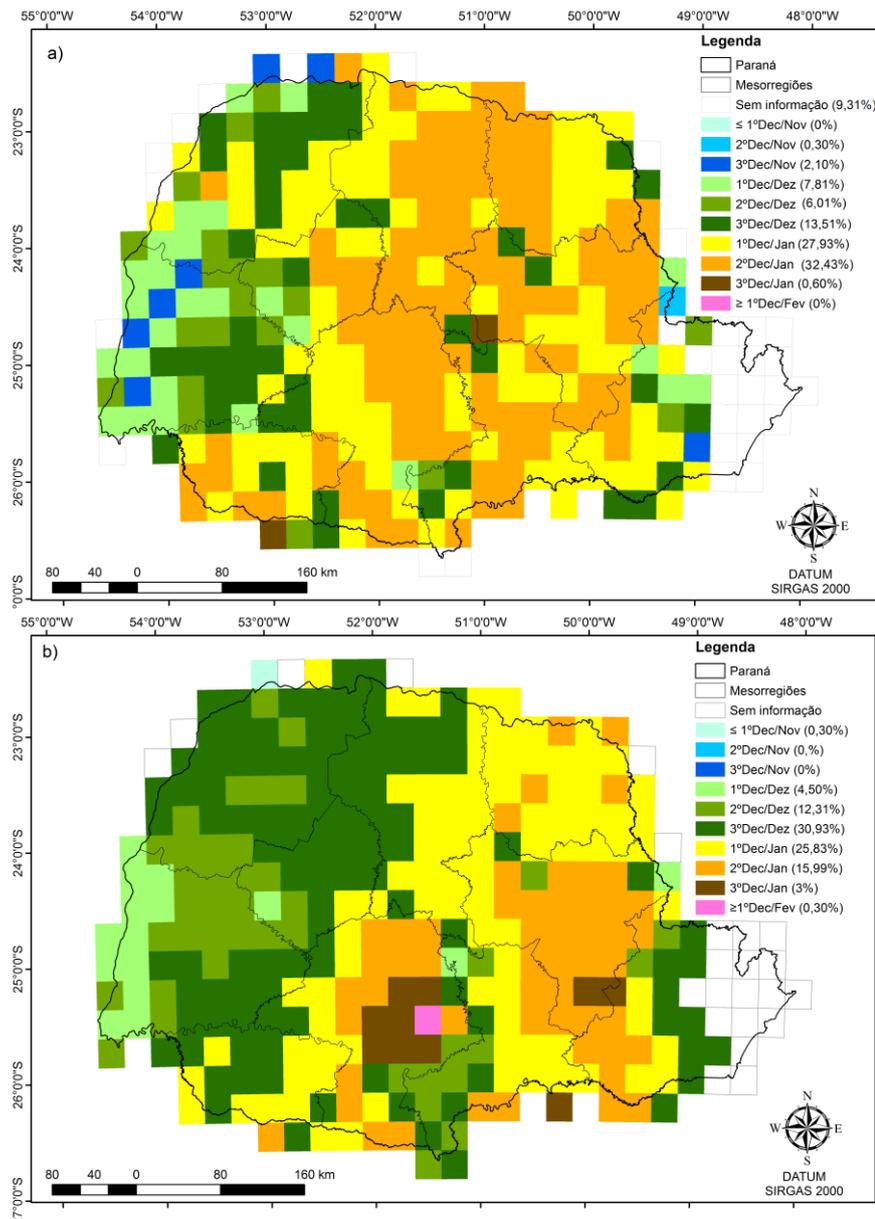


Figura 13 Mapa da distribuição do decêndio de máximo desenvolvimento vegetativo no estado do Paraná: a) do ano-safra de 2011/2012 e b) do ano-safra de 2013/2014.

Além disso, nas regiões com menores rendimentos de produtividade (Figura 12a), o DMDV ocorreu nos decêndios de dezembro. Esses resultados foram similares ao encontrados

por Johann et al. (2016), ao considerar o DMDV da cultura para esse mesmo ano-safra, ao considerar os municípios paranaenses.

No ano-safra 2013/2014, para o decêndio do DMDV da soja no estado, obtiveram-se os decêndios entre o 1º decêndio de dezembro até 3º decêndio de janeiro e no até 2º decêndio de março (Figura 13b), estando o DMDV concentrado no 3º decêndio de dezembro de (30,93%) e 1º decêndio de janeiro de (25,83%). Na grande maioria das mesorregiões com maiores valores de produtividade, o DMDV ocorreu no 1º decêndio de janeiro e no 2º decêndio de janeiro.

Verificou-se a similaridade entre os mapas do decêndio no DMDV, considerando as medidas de dissimilaridade, podendo constatar as medidas obtidos foram valores de baixos (Exatidão Global  $EG = 0,35$ , Kappa ( $K = 0,28$ ) e Tau ( $\tau = 0,18$ )), classificado como uma baixa exatidão de acordo com Anderson et al. (1976) e Krippendorff (1980).

### **5.3. Dados agrometeorológicos no decêndio de máximo desenvolvimento vegetativo (DMDV)**

De forma geral, ao analisar as variáveis agrometeorológicas no momento do DMDV nos anos-safra em estudo, constatou-se maior variabilidade no ano 2011/2012, quando comparado ao ano-safra de 2013/2014 (Figura 14).

As maiores temperaturas média do estado no momento do DMDV nos anos-safra ocorreram nas regiões Noroeste e Norte Central. A partir dessas regiões, verificou-se, uma diminuição gradativa da temperatura em direção à região Sudeste do estado (Figuras 14a e 14b), com os menores valores de temperatura média nas mesorregiões Sudeste e Campos Gerais.

A radiação solar no ano-safra 2011/2012 no momento do DMDV apresentou uma amplitude entre 0,14 a 0,30 MJ m<sup>-2</sup> dia<sup>-1</sup>, com os maiores valores (Figura 14c) ocorrendo em algumas mesorregiões com os menores valores de produtividade, como no Oeste, Noroeste e Sudoeste (Figura 12a). A menor radiação esteve concentrada nas mesorregiões com maiores produtividades, como Norte Central, Norte Pioneiro e Centro Oriental (Figura 12a). Em relação à radiação solar no ano-safra 2013/2014 no momento do DMDV, apresenta-se que em quase todo o estado os valores variaram entre 0,22 e 0,29 MJ m<sup>-2</sup> (65% das estações virtuais) (Figura 14f).

Para a variável evapotranspiração potencial no momento do DMDV, no ano-safra 2011/2012 (Figura 14e), tem-se uma variação entre 14 e 75 mm e uma tendência da distribuição dos valores da evapotranspiração com a produtividade, similar ao ocorrido com a radiação solar.

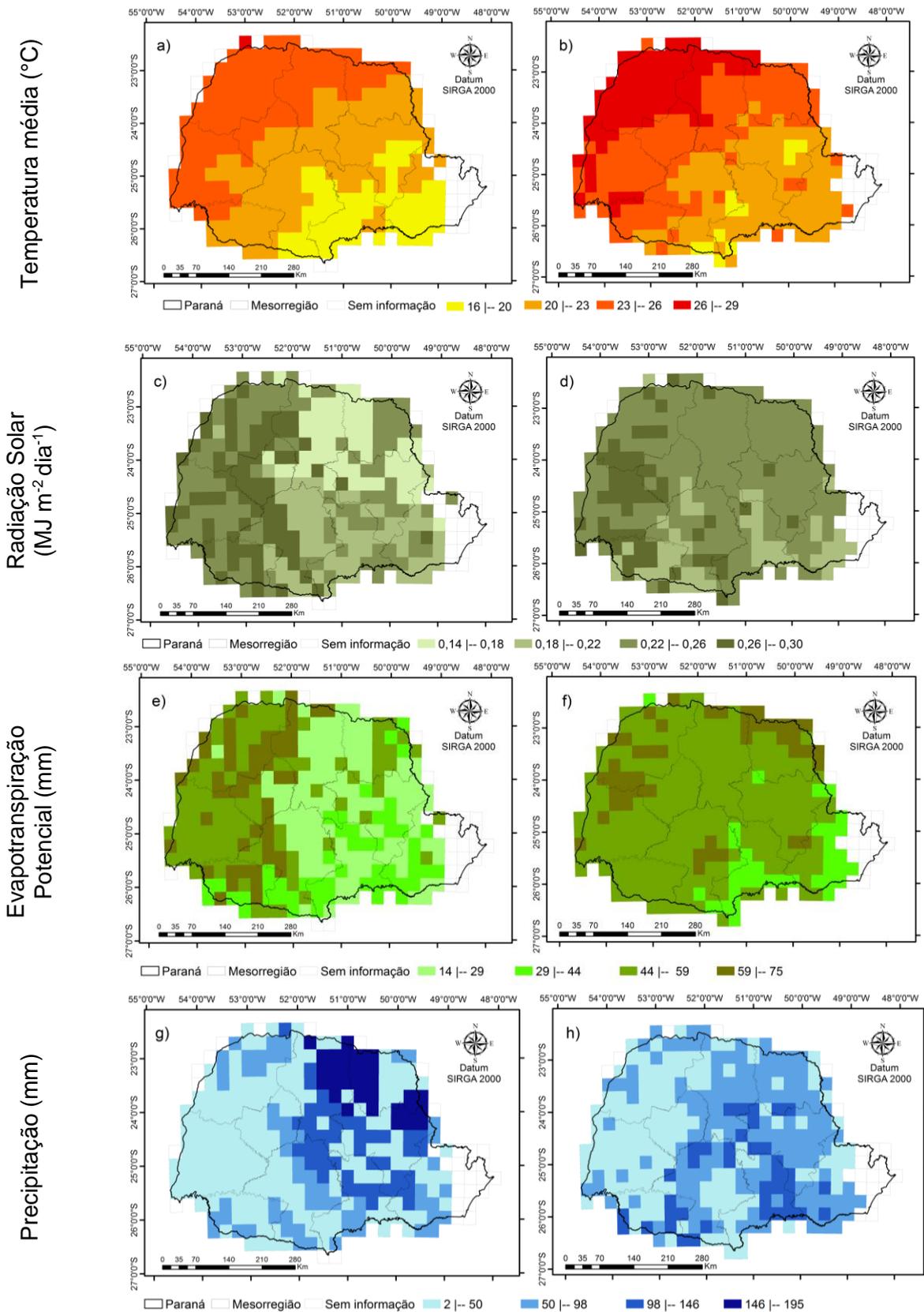


Figura 14 Mapa dos valores das seguintes variáveis no momento do DMDV: a) e b) temperatura média; c) e d) radiação solar; e) e f) evapotranspiração potencial; g) e h) precipitação; para os anos-safra do 2011/2012 e 2013/2014, respectivamente.

No ano-safra de 2013/2014 (Figura 14f), a evapotranspiração apresentou uma variação entre 29 e 75 mm, resultando em uma maior concentração dos valores da evapotranspiração entre 44 a 59 mm (76,77% das estações virtuais).

Em relação à distribuição espacial da quantidade total de precipitação ocorrida nos anos-safra de 2011/2012 e 2013/2014, no momento do DMDV (Figuras 14g e 14h), observou-se uma relação direta dos valores da precipitação com os valores da produtividade (Figuras 12a e 12b). Para o ano-safra 2011/2012, observou-se que a maior quantidade de precipitação (entre 98 a 195 mm) aconteceu em várias localidades das mesorregiões com maior produtividade. Em contrapartida, na maior parte das regiões com menor produtividade da soja, ocorreram também os menores volumes de precipitação, entre 2 e 97 mm.

Além disso, em 2013/2014 (Figuras 14g e 14h) constatou-se que em algumas mesorregiões com alta produtividade, como a Oeste e Centro Ocidental, houve uma quantidade de precipitação entre 50 a 98 mm. Já nas demais mesorregiões, com valores altos de produtividade, o volume de precipitação variou, de forma geral, entre 98 a 195 mm. Para as regiões com menores valores de produtividade, o volume de precipitação variou entre 2 a 97 mm – ou seja, de modo geral, as regiões com maior produtividade apresentaram um volume acumulado de precipitação adequado (50 a 195 mm/decêndio), como recomendado pela Embrapa (2007), de 7 a 8 mm dia<sup>-1</sup>.

Também constatou-se que, no ano-safra de 2013/2014, houve maiores volumes de chuva no DMDV no estado do Paraná (Figura 14h), quando comparado com o ano-safra de 2011/2012 (Figura 14g). Estudos realizados por Radin, Schönhofen e Tazzo (2017) mostraram que a variação do rendimento de grãos de soja entre anos-safras pode ser explicada pela quantidade de precipitação ocorrida no estágio de floração e enchimento dos grãos.

#### **5.4 Dados agrometeorológicos em decêndios próximos ao máximo desenvolvimento vegetativo (DMDV)**

Ao considerar os dados agrometeorológicos em torno do DMDV, ou seja, um decêndio antes e um depois da DMDV (1aDMDV1d) e dois decêndios antes e dois depois da DMDV (2aDMDV2d), observou-se que a temperatura média dos anos-safras em estudos (Figuras 15) apresentou padrão similares, com a mesma tendência da distribuição espacial ocorrida no momento do DMDV (Figuras 14a e 14b).

Resultados semelhantes na distribuição espacial da temperatura média foram observados por Aparecido et al. (2016), ao considerar a temperatura média do ar anual do estado do Paraná de uma série histórica desde 1989 até 2014.

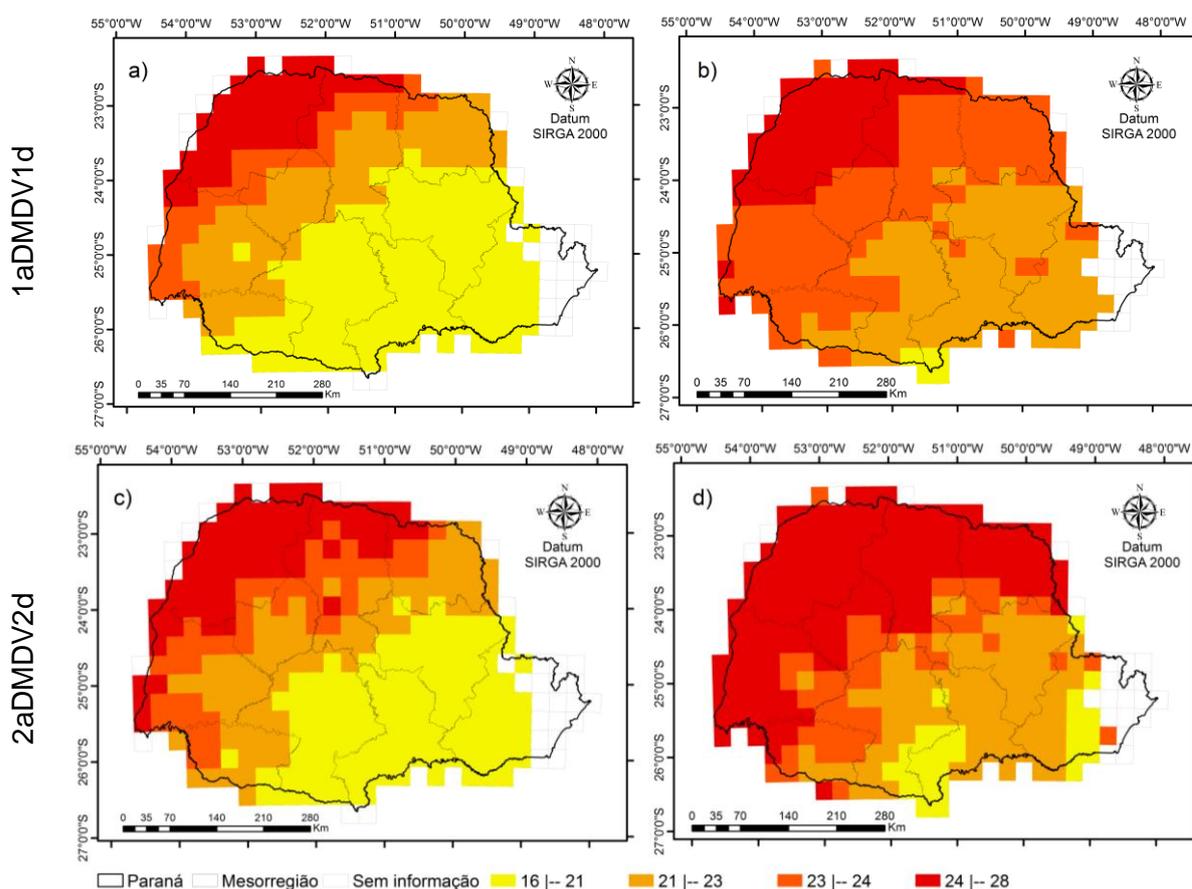


Figura 15 Distribuição espacial das medidas associadas à temperatura média (TMed) considerando as variações decenciais: a) e c) do anos-safra do 2011/2012 e, b) e d) do ano-safra 2013/2014.

Constatou-se que no ano-safra 2013/2014 (Figuras 15b e 15c) ocorreu temperatura média mais adequada ao cultivo da soja, nos dois intervalos decenciais considerados, estando dentro da variação de 20 a 30 °C, recomendada pela Embrapa (2016), o que pode justificar uma maior produtividade (Figura 12b) neste ano-safra.

A variável radiação solar ( $R_s$ ) apresentou, nos dois anos-safra, uma quantidade acumulada entre 0,49 a 0,88 MJ m<sup>-2</sup> dia<sup>-1</sup> para o intervalo decencial 1aDMDV1d (Figuras 17a e 17b). Já para o intervalo decencial 2aDMDV2d, a radiação acumulada esteve entre 0,88 e 1,33 MJ m<sup>-2</sup> dia<sup>-1</sup>, para os anos-safra em estudo (Figuras 16c e 16d). Resultados semelhantes foram encontrados por Heiffig et al. (2006), que observaram um aumento no valor da radiação solar na fase reprodutiva da soja em uma área experimental.

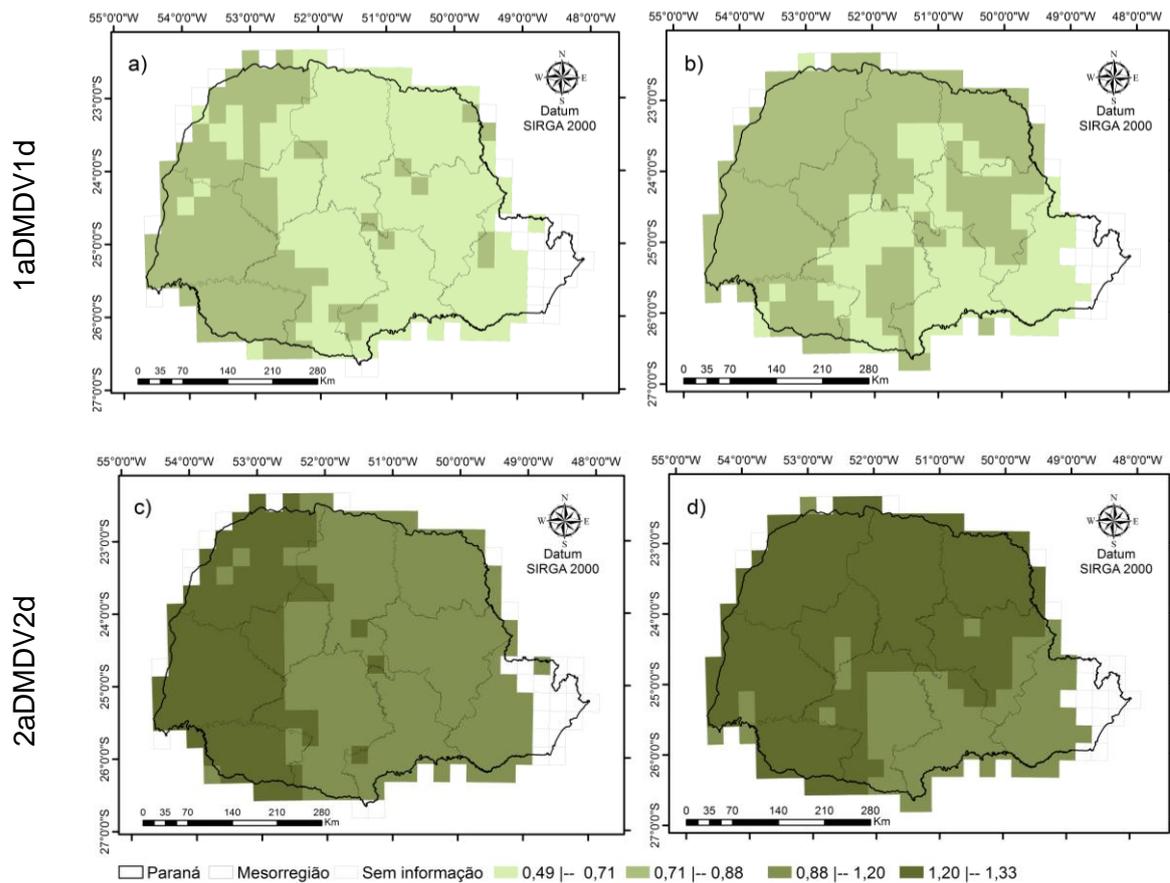


Figura 16 Distribuição espacial das medidas associadas a radiação solar ( $R_s$ ) considerando as variações decenciais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014.

A evapotranspiração potencial (ETp), no ano-safra de 2011/2012 no intervalo decencial 1aDMDV1d (Figura 17a), variou entre 56 e 152 mm nas mesorregiões com maiores valores de produtividade (Figura 12a). Nas mesorregiões com menores produtividades (Oeste e Noroeste), a ETp variou entre 152 e 181 mm, mostrando que, claramente, a cultura teve maior evapotranspiração e, em casos de ausência de precipitação regular, inevitavelmente haverá uma redução de produtividade de soja. Para o período 2aDMDV2d (Figura 17b), para as mesorregiões com maiores produtividades, a ETp variou entre 56 e 225 mm. Nas regiões com menores produtividades verificou-se que a ETp variou entre 225 a 325 mm. Dessa forma, nas regiões com menores valores de produtividades ocorrem maiores valores de ETp nos intervalos considerados, em relação às regiões com maiores produtividades.

Além disso, para o ano-safra 2013/2014 ao considerar intervalo decencial 1aDMDV1d (Figura 17b), apresentou-se a mesma tendência de distribuição da quantidade de ETp do ano-safra 2011/2012 em relação às regiões com maiores ou menores produtividades. Por outro lado, no intervalo decencial 2aDMDV2d (Figura 17d), a ETp (225 a 324 mm) apresentou valores mais elevados, em geral, no estado paranaense (89% das estações virtuais).

Ano-safra 2011/2012

Ano-safra 2013/2014

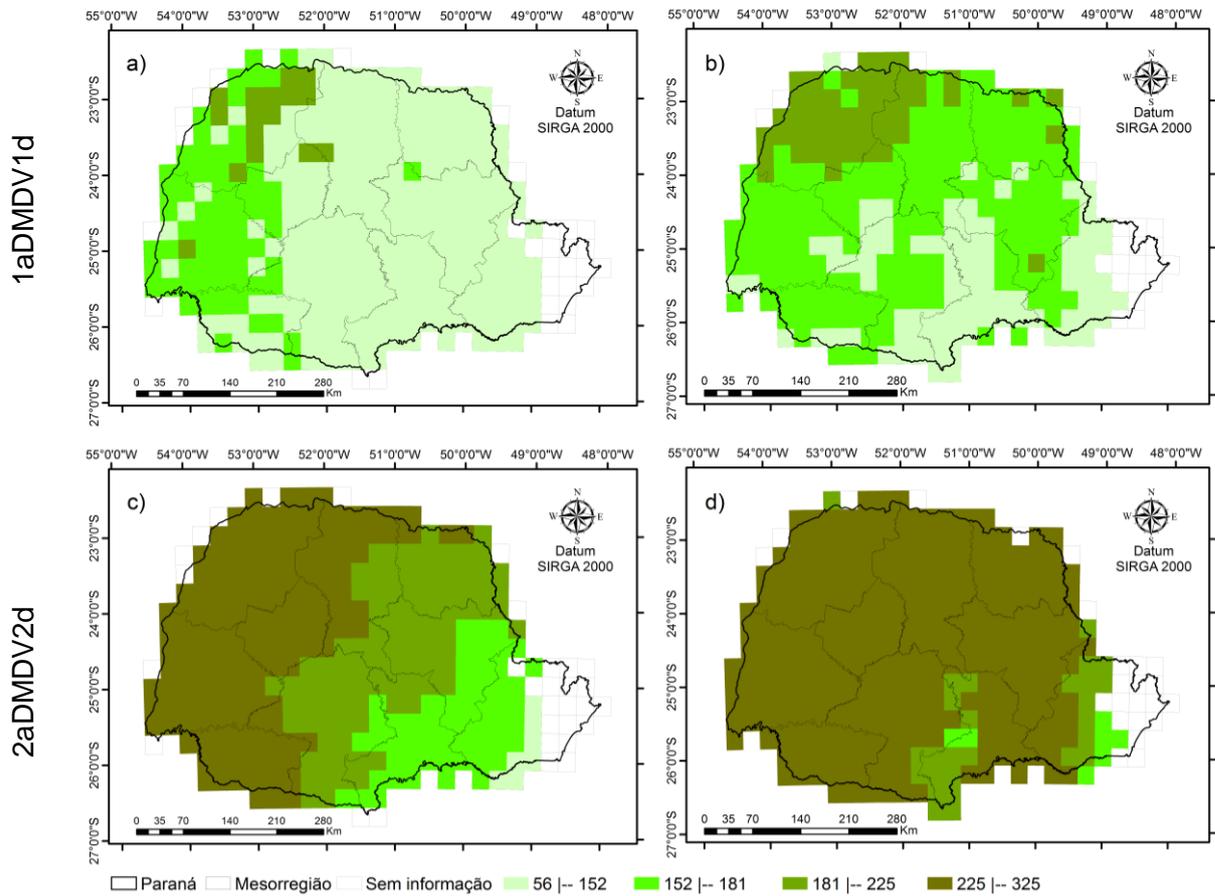


Figura 17 Distribuição espacial das medidas associadas à evapotranspiração potencial (ET<sub>p</sub>) considerando os intervalos decendiais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014.

Nas regiões em que ocorreu menor quantidade de precipitação, houve maior radiação solar (Figura 16). Isto pode ser justificado pois em dias sem a presença de nuvens há maior radiação solar (ALBERTO et al., 2006) e, conseqüentemente, há um aumento na quantidade de evapotranspiração das culturas (GRZEGOZEWSKI et al., 2017).

Constatou-se que no ano-safra 2011/2012 (Figuras 18a e 18c) as mesorregiões com maiores precipitações são também as mesorregiões nas quais ocorreram as maiores produtividades do estado (Figura 12a). Já para as regiões com menores produtividades, para ambos os intervalos decendiais, ocorreram os menores volume de precipitação acumulada (Figura 12a).

Observou-se que a quantidade de precipitação acumulada, nos intervalos decendiais considerados, demonstram a mesma tendência, principalmente no ano-safra 2011/2012.

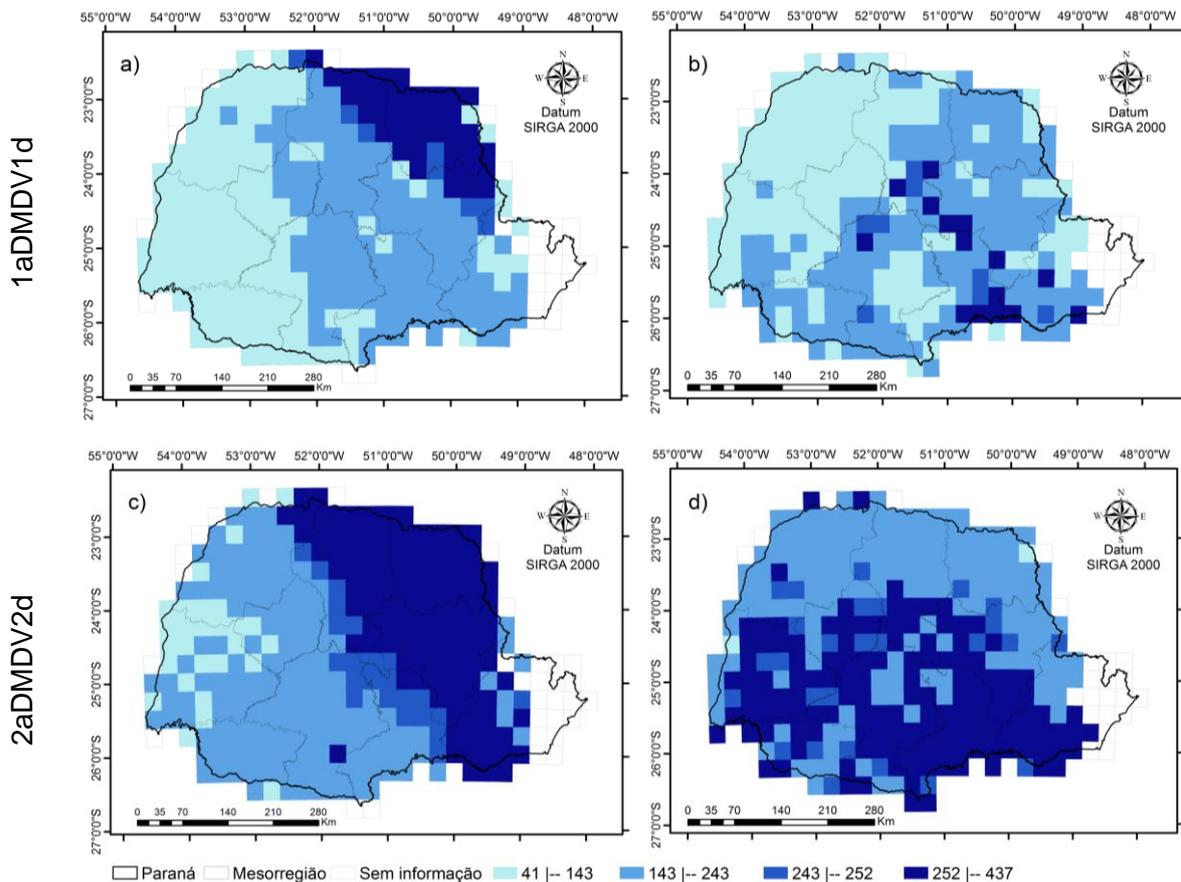


Figura 18 Distribuição espacial das medidas associadas à variável precipitação acumulada (Prec) considerando as variações decendiais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014.

No ano-safra de 2013/2014, no intervalo decendial 1aDMDV1d (Figura 18b), observa-se, de modo geral, menores volumes de precipitação acumulada em todo o estado ao comparar com os outros intervalos decendiais, deste mesmo ano e do ano-safra 2011/2012 (Figuras 18a, 18c e 18d). Neste intervalo decendial, nas mesorregiões com maiores valores de produtividade, ocorreram maiores volume de precipitação e, para as mesorregiões com menores produtividades, o volume de precipitação foi menor.

O intervalo decendial 2aDMDV2d (Figura 18d) do ano-safra 2013/2014 apresentou um maior volume de chuva em todo o estado, principalmente nas mesorregiões com maiores valores de produtividade (Figura 12b).

Ao considerar todo o ciclo da soja, a distribuição espacial das variáveis agrometeorológicas temperatura média, radiação solar acumulada e evapotranspiração potencial acumulada apresentam comportamento similar ao momento do DMDV e nos intervalos decendiais próximos ao DMDV (Figuras 19a a 19f).

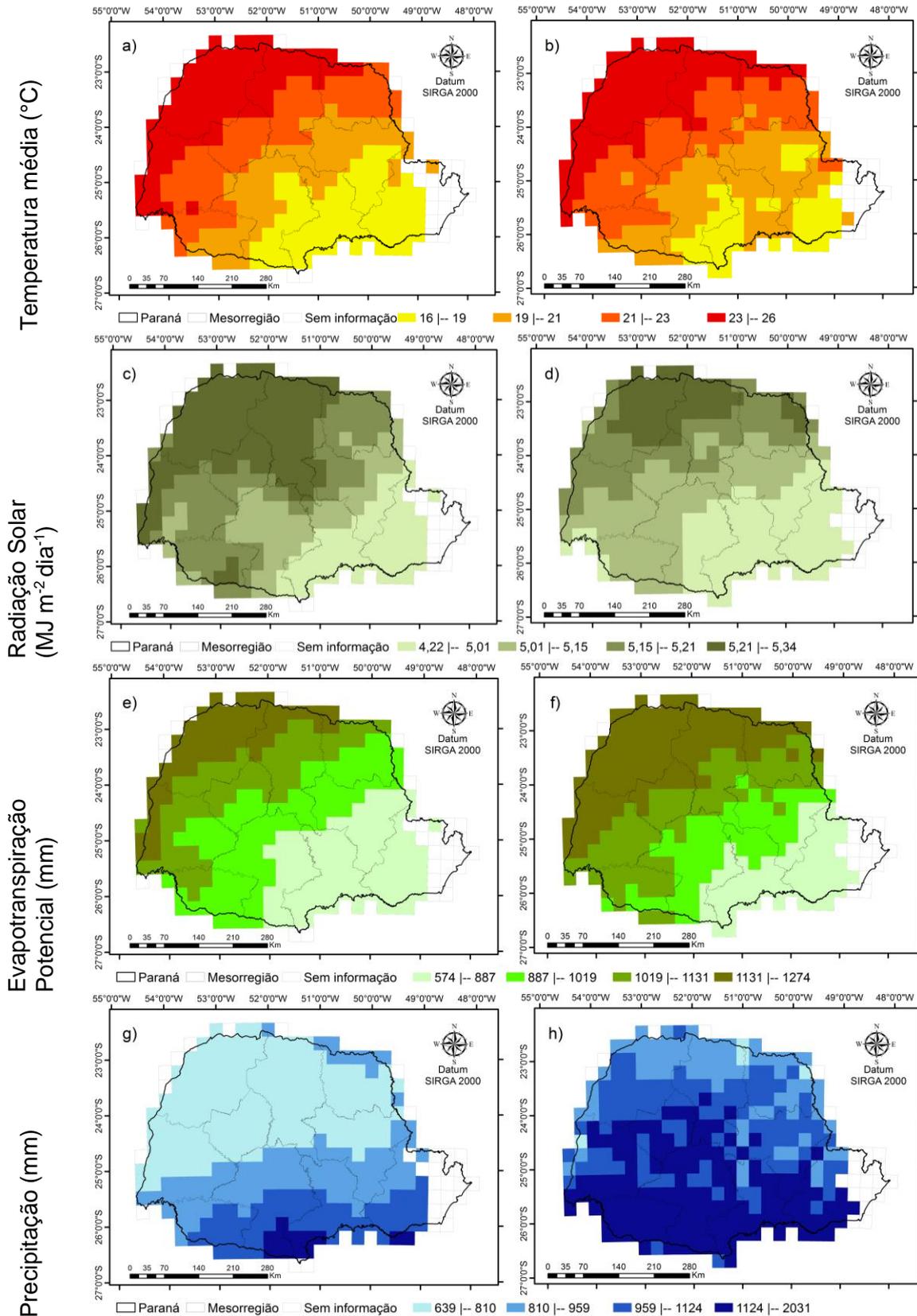


Figura 19 Distribuição espacial das variáveis em todo ciclo vegetativo da soja das seguintes variáveis agrometeorológicas: a) e b) temperatura média; c) e d) radiação solar acumulada; e) e f) evapotranspiração potencial acumulada; g) e h) precipitação acumulada; para os anos-safra do 2011/2012 e 2013/2014.

Em relação à precipitação acumulada no intervalo do ciclo vegetativo, este não apresentou comportamento de similaridade com os outros intervalos de variações decendiais

considerados, pois no ano-safra 2011/2012 (Figura 19g) houve maior concentração do volume nas mesorregiões do sul do estado, diminuindo gradativamente até o norte do estado. Por outro lado, no ano-safra 2013/2014 a distribuição da precipitação ocorreu em maior quantidade em praticamente todo o estado do Paraná (810 a 2031 mm), estando acima do valor recomendado pela Embrapa (2007), que é de 650 a 700 mm, devendo estar bem distribuído em todo o ciclo, principalmente na fases mais críticas, que são a emergência, reprodutiva e o enchimento do grãos (Figura 19h).

Além disso, de uma maneira geral, nas regiões com menor volume de precipitação, houve um aumento da quantidade da radiação. Isto pode ser explicado pelo fato que dias límpidos aumentam a disponibilidade da radiação solar (ALBERTO et al., 2006) e, conseqüentemente, há um aumento na quantidade de evapotranspiração (GRZEGOZEWSKI et al., 2017).

Resultados semelhantes foram encontrados por Silva, Sartori e Wollmann (2014), que concluíram que a precipitação influencia na produtividade da soja, principalmente a distribuição da precipitação no momento das fases emergência, reprodutiva e enchimento de grãos.

## **5. 5 Sazonalidade e Decêndio médio da ocorrência das variáveis**

Por meio da estatística circular, realizou-se o cálculo da média circular e do comprimento do vetor média ( $R$ ), o qual representa o decêndio médio da maior ocorrência de cada variável em estudos (TING; HARTLEY; BURNS, 2008).

Com isso, obtiveram-se valores próximos de zero para as seguintes variáveis: temperatura média, radiação solar e evapotranspiração potencial, em todas as intervalos decenciais (1aDMDV1d e 2aDMDV2d) e em todo o ciclo total da cultura da soja (Figura 20); Este resultado indica a não existência de sazonalidade segundo a classificação de Beskow (2014) e Neto et al. (2018) e, conseqüentemente, uma distribuição uniforme destas variáveis ao longo dos períodos considerados.

Para a variável precipitação (Figuras 20d e 20h) no ano-safra 2011/2012, nos intervalos decenciais 1aDMDV1d e 2aDMDV2d, encontraram-se valores moderados a altos de  $R$  de acordo com Beskow (2014), indicando uma distribuição uniforme do volume de chuva nos períodos analisados (ou seja, a presença de sazonalidade) na maior parte do estado. Destaca-se o intervalo decencial 2aDMDV2d e, em todo o ciclo vegetativo da cultura (Figuras 20h e 20l), a ocorrência de uma maior sazonalidade nas mesorregiões do Norte Pioneiro, Norte Central e parte do Centro Oriental. No entanto, ao considerar todo o ciclo vegetativo da cultura, observa-se a ocorrência de uniformidade nas outras regiões do estado.

No ano-safra 2013/2014 no intervalo decencial 1aDMDV1d, a precipitação apresentou uma distribuição uniforme em algumas regiões do Norte Pioneiro e Norte Central, ao passo que nas demais regiões do estado ocorreu sazonalidade de moderada a alta (Figura 20p).

O intervalo decendial 2aDMDV2d (Figura 20t) do ano-safra 2013/2014 apresentou uma sazonalidade de valores moderados a altos em todo o estado, enquanto que, ao considerar todo o ciclo vegetativo, a distribuição da chuva deu-se de maneira uniforme (Figura 20x).

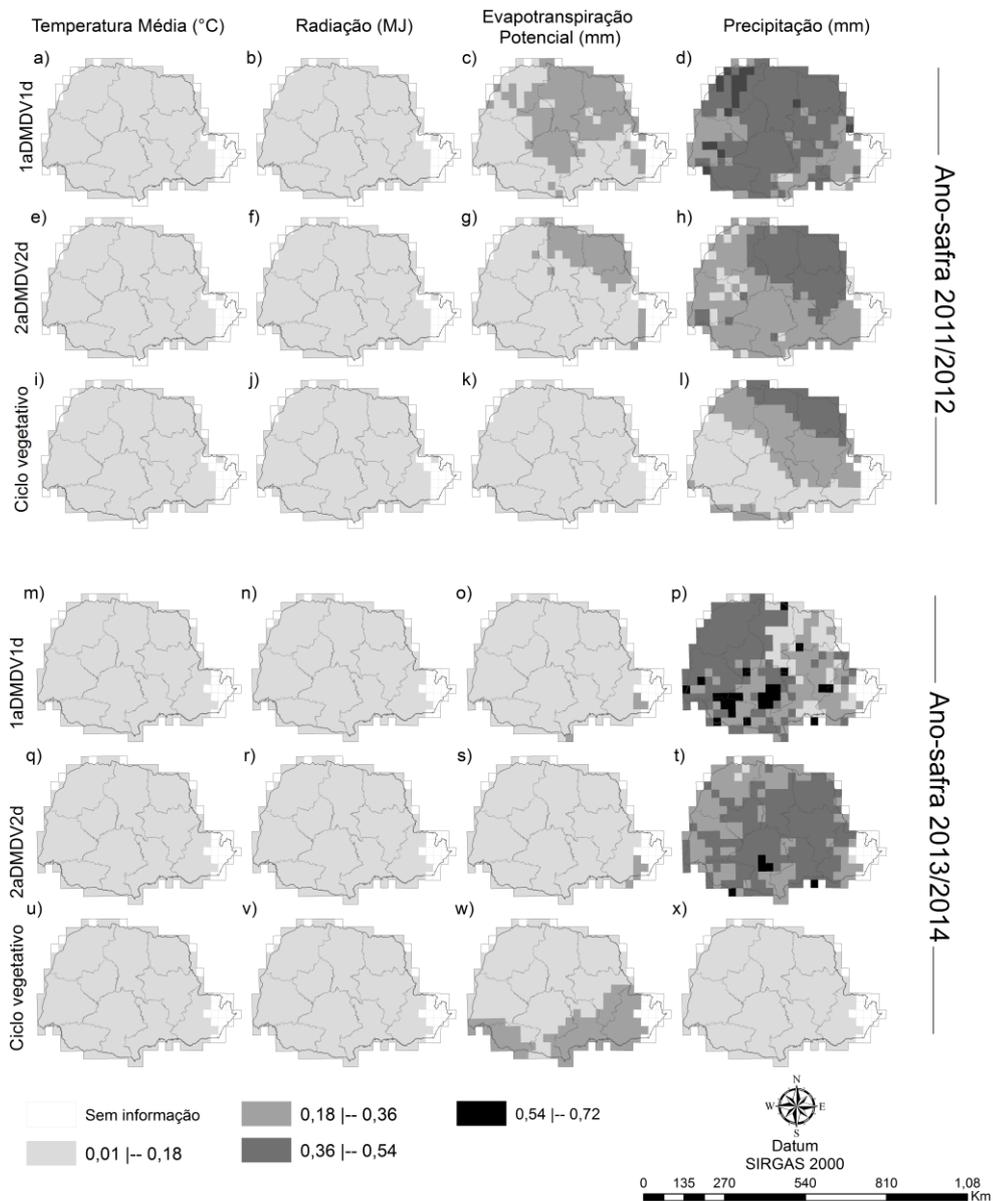


Figura 20 Distribuição espacial dos valores do  $R$  (comprimento do vetor média) e das variáveis: temperatura média, radiação solar acumulada, evapotranspiração potencial acumulada e precipitação acumulada do ano-safra do 2011/2012 e 2013/2014.

Outra maneira de mostrar a sazonalidade de determinado evento é por meio do histograma circular, sendo verificado o comprimento do vetor média (Figuras 21a a 21d).

Têm-se nas Figuras 21a a 21d as representações de histogramas circulares das estações virtuais, localizadas nas regiões com menores e maiores valores de produtividade, considerando os intervalos decendiais 1aDMDV1d e 2aDMDV2d dos dois anos-safra em estudos. O comprimento do vetor média (seta) indica a ocorrência da sazonalidade da variável precipitação acumulada nos intervalos decendiais considerados.

Na Figura 21a (histograma lado esquerdo), localizada na mesorregião Noroeste, região com menor produtividade (Figura 12a), e na Figura 21a (histograma lado direito), localizada no Centro Oriental, sendo região com maior valor de produtividade, tem-se o comprimento do vetor médio indicando uma moderada sazonalidade na distribuição do volume de precipitação no variação decendial 1aDMDV1d no ano-safra 2011/2012. Isto também ocorreu no ano-safra de 2013/2014 (Figura 21b) ao considerar o mesmo intervalo decendial, nas mesmas estações virtuais, localizadas nos histogramas circulares.

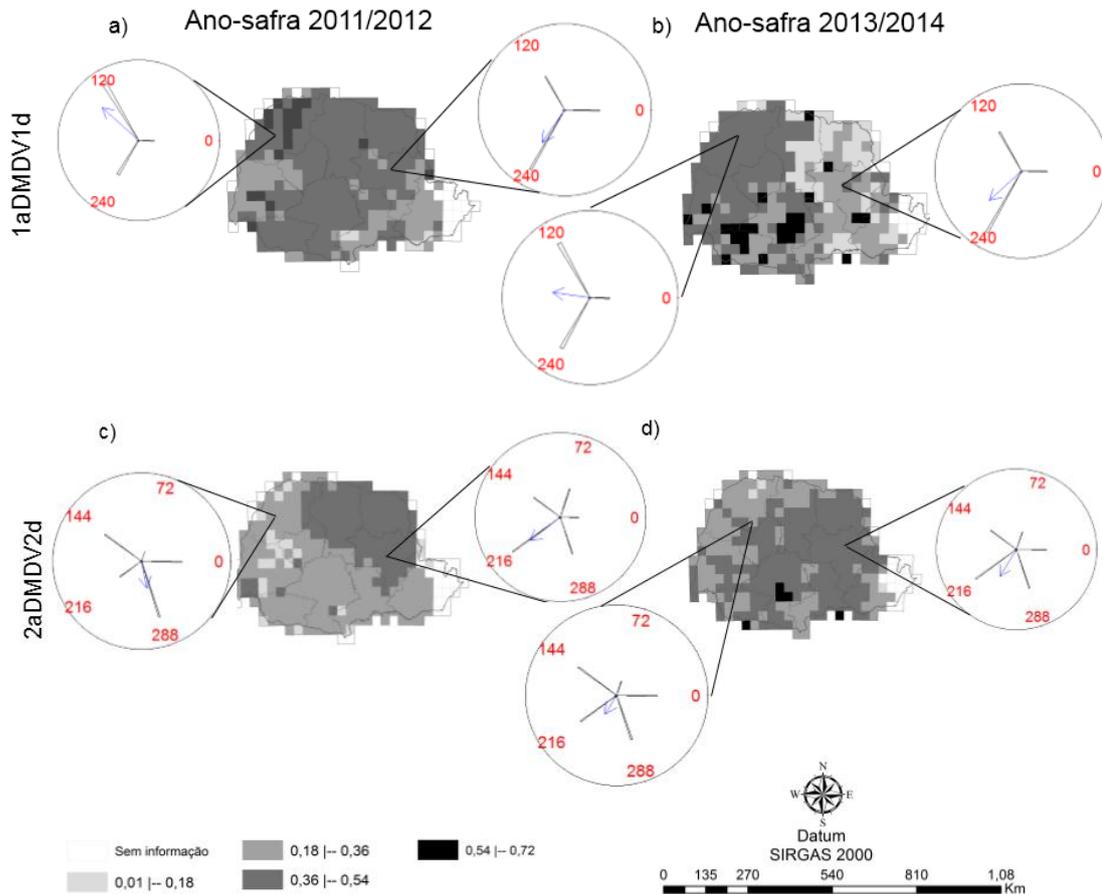


Figura 21 Distribuição espacial do volume de chuva do ano-safra 2011/2012 no intervalo decendial 1aDMDV1d. Histograma circular (unidade de medida em graus) do lado direito representa o pixel com a maior produtividade no estado do Paraná. Histograma circular (unidade de medida em graus) do lado esquerdo representa o pixel com a menor produtividade no estado do Paraná, a seta aponta para a data média e o comprimento da seta é o valor de R (varia de 0 a 1).

Já no intervalo decendial 2aDMDV2d (Figuras 21c e 21d, histograma lado esquerdo), estando os histogramas circulares localizados na mesorregião Noroeste (região com menores valores de produtividade), o comprimento do vetor médio (seta) indica uma baixa sazonalidade, ou seja, uma maior distribuição do volume da precipitação no intervalo considerado nos dois anos-safra. Ainda neste mesmo intervalo decendial, nas Figuras 21c e 21d, no histograma lado direito, sendo estes localizados na mesorregião Centro Oriental, o

comprimento do vetor médio (seta) indica uma moderada sazonalidade da ocorrência do volume da precipitação no intervalo decendial no dois anos-safras.

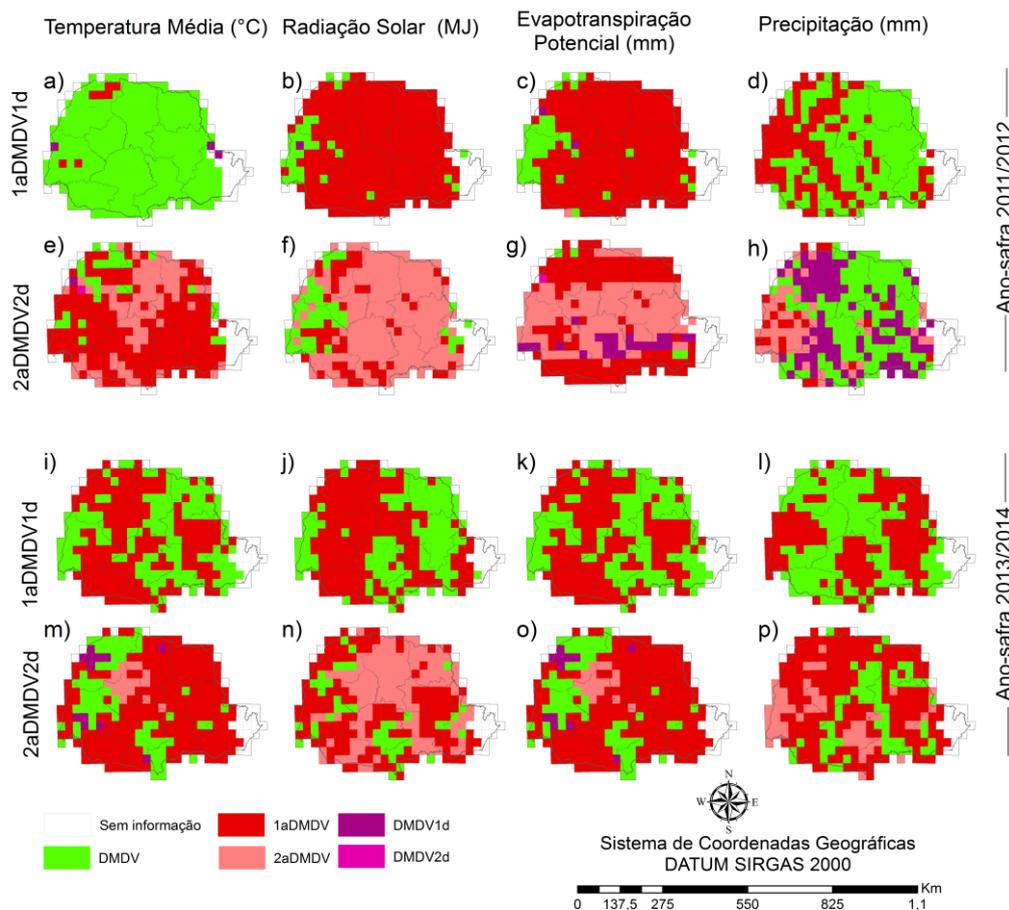


Figura 22 Distribuição espacial da média circular das variáveis: temperatura média, radiação solar acumulada, evapotranspiração potencial acumulada e precipitação acumulada do ano-safra do 2011/2012 e 2013/2014.

Mesmo com a distribuição uniforme das variáveis temperatura média, radiação solar e evapotranspiração potencial nos períodos considerados, as médias circulares (Figura 22) ocorreram, na maioria, no momento de um decêndio antes do DMDV (1aDMDV) e no momento do DMDV para os intervalos decenciais considerados nos dois anos-safras. Isto foi constatado pela Embrapa Soja (2007), em relação à maior quantidade da evapotranspiração potencial que ocorre no DMDV da soja.

Já o decêndio médio de maior ocorrência da variável precipitação acumulada (Figuras 22d, 22h, 22l e 22p) incidiu nos decêndios anteriores ao momento do DMDV e no decêndio do momento do DMDV para os intervalos decenciais (1aDMDV1d e 2aDMDV2d) dos anos-safras em estudo, ou seja, na fase necessária para atender às necessidades da cultura (EMBRAPA, 2007).

## 5.6 Análise Espacial de Área

Na análise espacial dos dados da produtividade dos anos-safras em estudo,

constatou-se a evidência de associação espacial entre as estações virtuais. Isto foi confirmado pelo valor estimado da estatística de Moran Global Univariado, que indicou uma autocorrelação espacial positiva e significativa ( $I = 0,81$ ) (Tabela 8). Logo, nas regiões paranaenses com alta ou baixa produtividades, as suas regiões vizinhas seguem o mesmo padrão de alta ou baixa produtividades, respectivamente.

O Lisa Cluster Map apresentou, para os dois anos-safras (Figuras 23a e 23b), a formação de dois clusters, ambos significativos (5% de significância), sendo um do tipo Baixo-Baixo, que são as mesorregiões com valores menores de produtividade (Nordeste e Oeste no ano-safra 2011/2012, e parte do Norte Pioneiro, Noroeste e Norte Central em 2013/2014) e o outro Alto-Alto, que corresponde àqueles com maiores produtividades da soja (Centro Oriental e parte do Sudeste e região Campos Gerais para o ano-safra 2011/2012 e Oeste parte das regiões Centro Ocidental, Centro Sul e Centro Oriental no ano-safra 2013/2014).

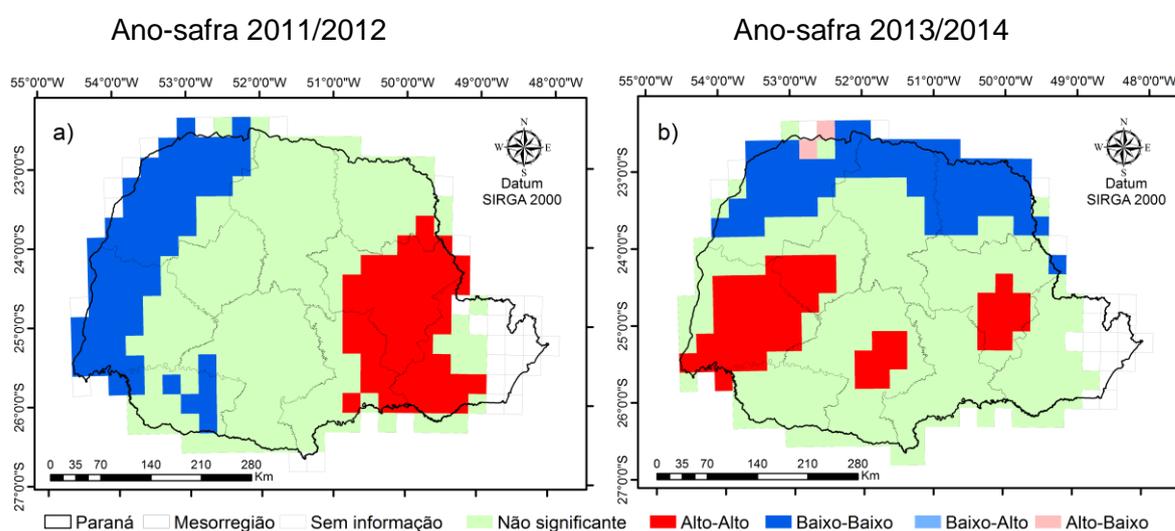


Figura 23 Lisa Cluster Map da variável produtividade nos anos-safras: a) 2011/2012; b) 2013/2014.

Felema et al. (2016) analisaram a produtividade de soja no estado do Paraná, entre os anos 2000 e 2010 e encontraram autocorrelação existente entre as regiões, sendo regiões de alta produtividade, vizinhas de regiões com características semelhantes, sendo que o mesmo se verificou em regiões de baixa produtividade.

Foi analisado também o Índice de Moran Global Univariado para as medidas associadas às variáveis agrometeorológicas e obtiveram-se, para todas elas, valores de autocorrelação espacial positivos e significativos (Tabelas 8 a 11).

Na Tabela 8 observa-se que, para os valores do Índice de Moran Global Bivariado, entre a produtividade e as medidas:  $ETp\_DMDV$ ,  $ETp\_1aDMDV1d$  para os dois anos-safras,  $ETp\_2aDMDV2d$  e  $R\_ETp\_1aDMDV1d$  apenas no ano-safra 2013/2014 obtiveram-se valores negativos (correlação espacial inversa). Ou seja, para as regiões com maiores valores de produtividade, ocorreram menores valores de  $ETp$  (vice-versa). Em contrapartida, as demais medidas associadas a  $ETp$  apresentaram uma correlação espacial bivariada positiva e

significativa com a produtividade. Isto é, nas regiões com maiores valores de produtividade, têm-se também os maiores valores de ETp, em todos os intervalos considerados e nos dois anos-safras (vice-versa).

Tabela 8 Índice de Moran Global Univariado ( $I$ ) e Bivariado ( $I_{xy}$ ) entre produtividade da soja e as medidas associadas à evapotranspiração potencial (ETp)

Variáveis	Ano-safra 2011/2012		Ano-safra 2013/2014	
	$I$	$I_{xy}$	$I$	$I_{xy}^{(a)}$
Produtividade	0,81 <sup>*(b)</sup>	-	0,81 <sup>*(b)</sup>	-
ETo_DMDV	0,55 <sup>*(b)</sup>	-0,54 <sup>*(b)</sup>	0,63*	-0,02*
ETo_1aDMDV1d	0,53 <sup>*(b)</sup>	-0,34 <sup>*(b)</sup>	0,69 <sup>*(b)</sup>	-0,18 <sup>*(b)</sup>
ETo_2aDMDV2d	0,81 <sup>*(b)</sup>	0,70 <sup>*(b)</sup>	0,70 <sup>*(b)</sup>	-0,24 <sup>*(b)</sup>
R_ETo_1aDMDV1d	0,57 <sup>*(b)</sup>	0,31 <sup>*(b)</sup>	0,42 <sup>*(b)</sup>	-0,17 <sup>*(b)</sup>
R_ETo_2aDMDV2d	0,76 <sup>*(b)</sup>	0,34 <sup>*(b)</sup>	0,68 <sup>*(b)</sup>	0,11 <sup>*(b)</sup>

ETp\_DMDV: dados de ETp no decêndio da data do DMDV; ETp\_1aDMDV\_1d: soma da ETp no 1aDMDV1d; ETp\_2aDMDV2d: soma da ETp no 2aDMDV2d; R\_ETp\_1aDMDV1d: valor do comprimento do vetor médio da ETp do 1aDMDV1d; R\_ETp\_2aDMDV2d: valor do comprimento do vetor médio da ETp no 2aDMDV2d.

\*Índice de teste significativo com 5% de significância (p-valor < 0,05).

<sup>(a)</sup>Índice de Moran Bivariado Global entre a variável Produtividade e as demais variáveis.

<sup>(b)</sup> dados transformados para apresentarem comportamento da distribuição normal.

Na Tabela 9 tem-se o  $I_{xy}$  entre produtividade e as medidas associadas à radiação solar (Rs). Por meio destes valores, observa-se que, com exceção da Rs\_1aDMDV1d do ano-safra 2011/2012, os valores de  $I_{xy}$ s são significativos.

As variáveis Rs\_DMDV do ano-safra 2011/2012 e Rs\_1aDMDV1d e Rs\_2aDMDV2d do ano-safra 2013/2014 têm associação espacial negativa, indicando que regiões com maior valor destas variáveis associadas à Rs têm valores com menor produtividade (vice-versa). Entretanto, as demais medidas associadas à Rs, nos dois anos-safras, têm valores de correlação espacial bivariada positiva, ou seja, nas regiões com maiores valores de produtividade ocorreram os maiores valores de Rs (vice-versa).

Tabela 9 Índice de Moran Global Univariado ( $I$ ) e Bivariado ( $I_{xy}$ ) entre produtividade da soja e as medidas associadas à radiação solar (Rs)

Variáveis	Ano-safra 2011/2012		Ano-safra 2013/2014	
	$I$	$I_{xy}$	$I$	$I_{xy}^{(a)}$
Produtividade	0,81 <sup>*(b)</sup>	-	0,81 <sup>*(b)</sup>	-
Rs_DMDV	0,48 <sup>*(b)</sup>	-0,43 <sup>*(b)</sup>	0,59*	0,06*
Rs_1aDMDV1d	0,53 <sup>*(b)</sup>	-0,03 <sup>ns(b)</sup>	0,72 <sup>*(b)</sup>	-0,20 <sup>*(b)</sup>
Rs_2aDMDV2d	0,55*	0,70*	0,77 <sup>*(b)</sup>	-0,40 <sup>*(b)</sup>
R_Rs_1aDMDV1d	0,59*	0,23*	0,64*	0,20*
R_Rs_2aDMDV2d	0,62*	0,15*	0,74*	0,31*

Rs\_DMDV: dados de Rs no decêndio da data do DMDV; Rs\_1aDMDV1d: soma da Rs no 1aDMDV1d; Rs\_2aDMDV2d: soma da Rs no 2aDMDV2d; R\_Rs\_1aDMDV1d: valor do comprimento do vetor médio da Rs do 1aDMDV1d; R\_Rs\_2aDMDV2d: valor do comprimento do vetor médio da Rs no 2aDMDV2d. \*Índice de teste significativo com 5% de significância (p-valor < 0,05). <sup>(ns)</sup> Não significativo com 5% de significância (p-valor < 0,05). <sup>(a)</sup>Índice de Moran Bivariado Global entre a variável Produtividade e as demais variáveis. <sup>(b)</sup> dados transformados para apresentarem comportamento da distribuição normal.

A correlação espacial bivariada (Tabela 10) entre a produtividade e as medidas associadas à temperatura média (TMed): R\_TMed\_1aDMDV1d e R\_TMed\_2aDMDV2d dos dois anos-safras apresentou valores positivos. Isto é, as regiões com maiores valores de

produtividade têm os maiores valores das medidas da temperatura média e vice-versa. Já as outras medidas associadas à TMed mostraram uma correlação espacial bivariada negativa, indicando que regiões com maior valor destas variáveis associadas à TMed têm valores com menor produtividade (vice-versa).

Tabela 10 Índice de Moran Global Bivariado entre produtividade da soja e as medidas associadas à Temperatura média (TMed)

Variáveis	Ano-safra 2011/2012		Ano-safra 2013/2014	
	$I$	$I_{xy}$	$I$	$I_{xy}^{(a)}$
Produtividade	0,81 <sup>*(b)</sup>	-	0,81 <sup>*(b)</sup>	-
TMed_DMDV	0,85 <sup>*(b)</sup>	-0,62 <sup>*(b)</sup>	0,81 <sup>*(b)</sup>	-0,28 <sup>*(b)</sup>
TMed_1aDMDV1d	0,84 <sup>*(b)</sup>	-0,64 <sup>*(b)</sup>	0,87 <sup>*(b)</sup>	-0,30 <sup>*(b)</sup>
TMed_2aDMDV2d	0,82 <sup>*(b)</sup>	-0,64 <sup>*(b)</sup>	0,88 <sup>*(b)</sup>	-0,30 <sup>*(b)</sup>
R_TMed_1aDMDV1d	0,77 <sup>*(b)</sup>	0,75 <sup>*(b)</sup>	0,53 <sup>*(b)</sup>	0,21 <sup>*(b)</sup>
R_TMed_2aDMDV2d	0,50 <sup>*(b)</sup>	0,42 <sup>*(b)</sup>	0,44 <sup>*(b)</sup>	0,31 <sup>*(b)</sup>

TMed\_DMDV: dados de TMed no decêndio da data do DMDV; TMed\_1aDMDV1d: soma da TMed no 1aDMDV1d; TMed\_2aDMDV2d: soma da TMed no 2aDMDV2d; R\_TMed1aDMDV1d: valor do comprimento do vetor médio da TMed do 1aDMDV1d; R\_TMed2aDMDV2d: valor do comprimento do vetor médio da Rs no 2aDMDV2d. \*Índice de teste significativo com 5% de significância ( $p$ -valor < 0,05).<sup>(ns)</sup> Não significativo com 5% de significância ( $p$ -valor < 0,05).<sup>(a)</sup> Índice de Moran Bivariado Global a variável Produtividade e as demais variáveis.<sup>(b)</sup> dados transformados para apresentarem comportamento da distribuição normal.

Considerando o Índice de Moran Global Bivariado entre a variável produtividade e as medidas associadas com a precipitação (Tabela 11), tem-se que, com exceção da medida Prec\_DMDV do ano-safra 2013/2014, os valores encontrados para correlação espacial foram positivos e significativos, ou seja, nas regiões com maiores valores de produtividade ocorrem os maiores valores das medidas associadas à precipitação.

Observa-se, pela Figura 24, a apresentação do Lisa Cluster Map das medidas associadas a cada variável agrometeorológica nos anos-safra em estudo. Apresentou-se a formação de dois clusters maiores e significativos (com 5% de significância), sendo um do tipo Baixo-Baixo e o outro de Alto-Alto.

Para as medidas Prec\_1aDMDV1d e Prec\_2aDMDV2d, associadas à variável precipitação do ano-safra 2011/2012, foi possível verificar a formação do cluster Baixo-Baixo, nas mesorregiões: Oeste, e em parte das seguintes mesorregiões: Noroeste, Sudoeste, Centro Sul e Centro Ocidental, sendo estas as regiões com baixa produtividade (Figura 23a). Por outro lado, os clusters Alto-Alto estão nas mesorregiões: Norte Pioneiro, em parte do Centro Oriental e Norte Central. Constatou-se que para a medida Prec\_DMDV não ocorreu a formação de cluster.

Já na formação de cluster para a medida R\_Prec\_1aDMDV1d ocorreu o tipo Alto-Alto na região Norte do estado e o tipo Baixo-Baixo em alguns pixels espalhados pelo estado, principalmente na região Campos Gerais.

Por fim, tem-se a medida R\_Prec\_2aDMDV2d, que apresentou apenas o cluster Alto-Alto nas mesorregiões: Centro Oriental, Norte Central e Norte Pioneiro, e o cluster.

Tabela 11 Índice de Moran Global Univariado e Bivariado entre produtividade da soja e as

## medidas associadas a precipitação (Prec)

Variáveis	Ano-safra 2011/2012		Ano-safra 2013/2014	
	$I$	$I_{xy}$	$I$	$I_{xy}^{(a)}$
Produtividade	0,81 <sup>*(b)</sup>	-	0,81 <sup>*(b)</sup>	-
Prec_DMDV	0,46 <sup>*(b)</sup>	0,42 <sup>*(b)</sup>	0,45 <sup>*(b)</sup>	-0,03 <sup>ns(b)</sup>
Prec_1aDMDV1d	0,75*	0,48*	0,54 <sup>*(b)</sup>	0,11 <sup>*(b)</sup>
Prec_2aDMDV2d	0,83*	0,51*	0,58*	0,32*
R_Prec_1aDMDV1d	0,52 <sup>*(b)</sup>	0,23 <sup>*(b)</sup>	0,55*	0,20*
R_Prec_2aDMDV2d	0,69*	0,21*	0,67*	0,22*

Prec\_DMDV: dados de Prec no decêndio da data do DMDV; Prec\_1aDMDV1d: soma da Prec no 1aDMDV1d; Prec\_2aDMDV2d: soma da Prec no 2aDMDV2d; R\_Prec\_1aDMDV1d: valor do comprimento do vetor médio da Prec do 1aDMDV1d; R\_Prec\_2aDMDV2d: valor do comprimento do vetor médio da Prec no 2aDMDV2d. \*Índice de teste significativo com 5% de significância (p-valor < 0,05).<sup>(a)</sup>Índice de Moran Bivariado Global entre a variável Produtividade e as demais variáveis.<sup>(b)</sup> dados transformados para apresentarem comportamento da distribuição normal.

No ano-safra 2013/2014, as medidas Prec\_DMDV e Prec\_1aDMDV1d (Figura 24) apresentaram a formação do cluster Alto-Alto nas mesorregiões: Sudeste, Centro Sul, Sudoeste, Norte Central, Centro Ocidental e nos Campos Gerais. Além disso, apresentou o cluster Baixo-Alto em alguns pixels dessas mesorregiões. Por outro lado, o cluster Baixo-Baixo está concentrado nas mesorregiões: Oeste, Noroeste e Centro Ocidental

Para a medida Prec\_2aDMDV2d, tem-se a formação do cluster Alto-Alto em uma parte das mesorregiões Sudeste e Campos Gerais. Já o cluster Baixo-Baixo ocorreu em alguns pixels nas mesorregiões Noroeste, Norte Central e Norte Pioneiro.

A medida R\_Prec\_1aDMDV1d apresentou cluster Alto-Alto em alguns pixels das seguintes mesorregiões: Oeste, Sudoeste, Centro Sul, Centro Ocidental, Noroeste e Norte Central e o cluster Baixo-Baixo concentrou-se no Norte Pioneiro, Norte Central. Por fim, a medida R\_Prec\_2aDMDV2d, tem a formação do cluster Alto-Alto nas mesorregiões do Centro do estado.

Em relação às medidas Tmed\_DMDV, TMed\_1aDMDV1d e TMed\_2aDMDV2d dos dois anos-safra em estudo (Figura 24), estas apresentaram o cluster Alto-Alto nas mesorregiões: Oeste, Noroeste e uma pequena região do Norte; enquanto que o cluster Baixo-Baixo concentrou-se nas mesorregiões: Centro Sul Sudeste, Centro Oriental e Campos Gerais, regiões consideradas com maiores valores de produtividade (Figuras 23a e 23b).

Na medida R\_TMed\_1aDMDV1d do ano-safra 2011/2012, notou-se a formação do cluster Alto-Alto nas mesorregiões: Centro Sul, Sudeste, Centro Oriental e Campos Gerais; o cluster Baixo-Baixo nas mesorregiões Noroeste e Oeste.

Para a medida R\_TMed\_2aDMDV2d do mesmo ano-safra anterior, apresentou-se o cluster Alto-Alto com uma maior concentração nas mesorregiões: Sudeste e Centro Oeste; e o cluster Baixo-Baixo na mesorregião Noroeste (Figura 24).

Em relação aos R\_TMed\_1aDMDV1d e R\_TMed\_2aDMDV2d do ano-safra 2013/2014, apresentaram o cluster Alto-Alto nas mesorregiões: Oriental, Centro Sul e Campos Gerais. Por outro lado, o cluster Baixo-Baixo da medida R\_TMed\_1aDMDV1d se concentrou

no Noroeste, Centro Ocidental e Norte Central; e para R\_TMed\_2aDMDV2d nas mesorregiões Norte Pioneiro e Norte Central, sendo que ambas as medidas apresentaram cluster Alto-Baixo em alguns pixels espalhado pelo estado.

A medida Rs\_DMDV do ano-safra 2011/2012 apresentou o cluster Alto-Alto nas mesorregiões: Sudoeste, Noroeste e Centro Ocidental; e o clusters Baixo-Baixo no Norte Pioneiro, Centro Oriental e Norte Central.

Nas medidas Rs\_1aDMDV1d e Rs\_2aDMDV2d do ano-safra mencionado anteriormente, destaca-se o cluster Alto-Alto nas mesorregiões Oeste, Noroeste e Sudoeste, sendo as regiões com menor produtividade (Figura 23a).

Na Figura 24, o cluster Alto-Alto para R\_Rs\_1aDMDV1d aparece localizado no Sudoeste e Centro Sul; e o cluster Baixo-Baixo no Norte Pioneiro e Norte Central. Tem-se também a representação dos clusters da medida R\_Rs\_2aDMDV2d, e constatou-se a concentração do Alto-Alto no Norte Pioneiro e Norte Central, apresentando o Baixo-Baixo em alguns pixels no Oeste e Sudeste.

No ano-safra 2013/2014, a medida Rs\_DMDV apresentou uma maior concentração do cluster Alto-Alto nas mesorregiões Oeste, sendo esta mesorregião com maiores valores de produtividade (Figura 23b).

Para a medida Rs\_1aDMDV1d deste ano-safra mencionado anteriormente, notou-se a concentração do cluster Alto-Alto na mesorregião Noroeste, apresentando o cluster Baixo-Baixo no Sudeste na região Campos Gerais.

Já a medida Rs\_2aDMDV2d apresentou o cluster Alto-Alto nas mesorregiões Noroeste e Norte Pioneiro e o cluster Baixo-Baixo na região Sudeste e Campos Gerais. Em ambas as medidas mencionadas, o cluster Alto-Alto ocorreu na mesorregião (Noroeste), sendo uma das mesorregiões com menores valores de produtividade (Figura 23b).

Ao considerar as medidas Rs\_1aDMDV1d e Rs\_2aDMDV2d do ano-safra 2013/2014, notou-se a formação do cluster Alto-Alto no Sudoeste e em parte do Centro Sul. Entretanto, o cluster Baixo-Baixo ocorreu em regiões distintas, no Norte Pioneiro e leste do Norte Central para Rs\_1aDMDV1d e noroeste e oeste do Norte Central para Rs\_2aDMDV2d.

Ao analisar a medida ETp\_DMDV no ano-safra de 2011/2012 (Figura 24), verificou-se a formação do cluster Alto-Alto estando concentrado nas mesorregiões Noroeste, e o clusters Baixo-Baixo no Norte Pioneiro e Centro Oriental.

Para as medidas ETp\_1aDMDV1d e ETp\_2aDMDV2d do mesmo ano-safra anterior, tem-se a concentração do cluster Alto-Alto nas mesorregiões Oeste, Noroeste e Sudeste (Figura 24). Por outro lado, o cluster Baixo-Baixo da medida ETp\_2aDMDV2d, ocorreu na região Campos Gerais, em parte das mesorregiões do Centro Oriental e Sudeste, sendo regiões com altas produtividades (Figura 23a).

Em relação ao R\_ETp\_1aDMDV1d e R\_ETp\_2aDMDV2d do ano-safra 2011/2012 verificou-se uma maior concentração do cluster Alto-Alto, no Norte Pioneiro e Norte Central. Já o cluster Baixo-Baixo, na mesorregião Oeste.

Para o ano-safra 2013/2014, notou para as medidas ETp\_DMDV, ETp\_1aDMDV1d e ETp\_2aDMDV2d uma mesma tendência das regiões nas localizações dos clusters, para o Alto-Alto nas mesorregiões Noroeste e Norte Pioneiro (regiões com valores menores de produtividade, Figura 23b) e o clusters Baixo-Baixo em uma parte da região Campos Gerais.

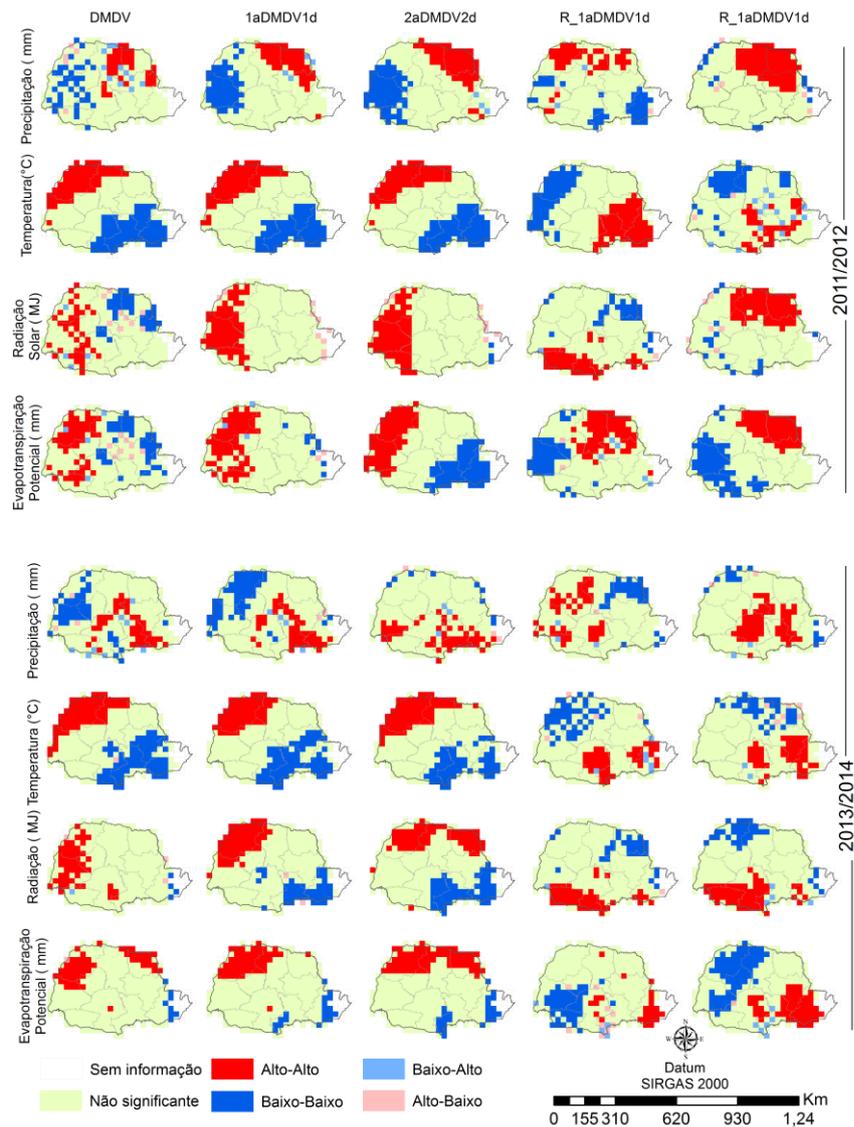


Figura 24 Lisa Cluster Map das medidas das variáveis agrometeorológicas dos anos-safra 2011/2012 e 2013/2014

Já para a medida R\_ETp\_1aDMDV1d do ano-safra 2013/2014 tem-se a concentração do cluster Alto-Alto no Oeste, Sudoeste e Centro Sul, e o cluster Baixo-Baixo na região dos Campos Gerais. Além disso, verificou-se a concentração do cluster Alto-Alto nas regiões Campos Gerais e Centro Oriental, e do cluster Baixo-Baixo de R\_ETp\_2aDMDV2d no Oeste, Noroeste, Norte Central e Centro Ocidental.

Estudos semelhantes realizados por Araújo, Uribe-Opazo e Johann (2014) encontraram uma similaridade nos resultados, em relação à correlação inversa entre produtividade e algumas medidas associadas às variáveis agrometeorológicas consideradas e detectaram áreas de agrupamentos por meio do Lisa Cluster Map.

Grzegozewski et al. (2017) analisaram correlação espacial da produtividade da soja e algumas variáveis agrometeorológicas (balanço hídrico, radiação global e temperatura média) no estado do Paraná, em escala decendial, e verificaram que a falta de precipitação causada pelo balanço hídrico negativo, juntamente com a alta radiação global, evapotranspiração e temperaturas na semeadura nas fases de desenvolvimento da cultura afetaram a sua produtividade.

## 5.7 Análise Multivariada

Conforme o agrupamento realizado pelo método k-means, para o ano-safra 2011/2012 (Figuras 25a e 25c), constatou-se um mesmo padrão da distribuição espacial destes grupos, quando comparados com as medidas associadas às variáveis radiação solar e precipitação no 2aDMDV2d (Figuras 16c e 18c).

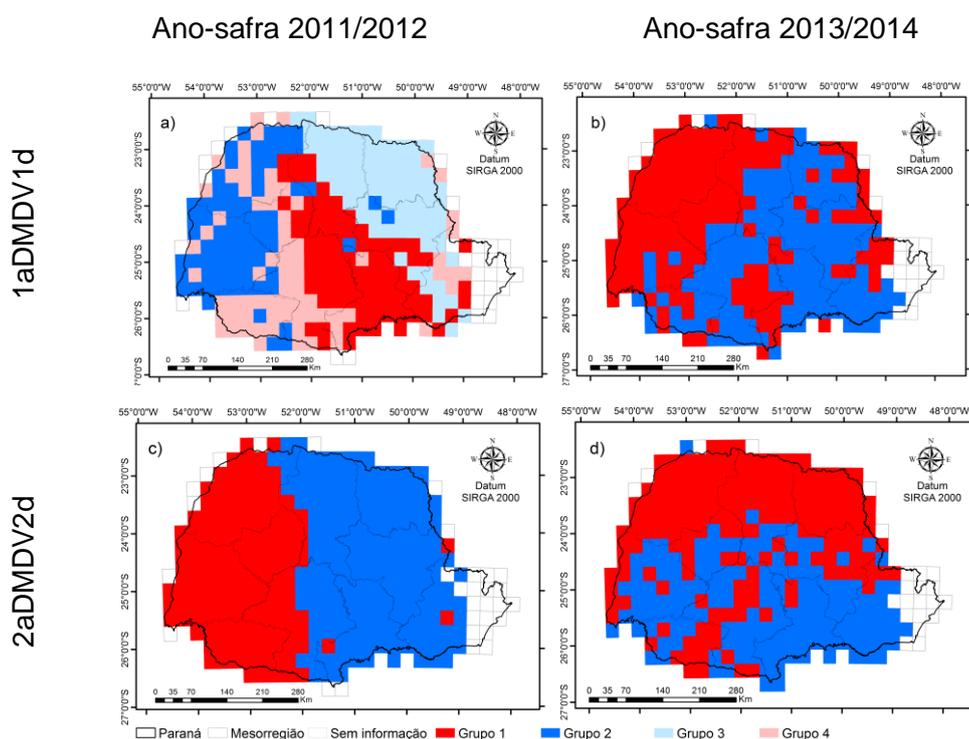


Figura 25 Distribuição espacial do agrupamento nas variações decendiais: a) e c) do ano-safra 2011/2012, b) e d) do ano-safra 2013/2014

Por outro lado, em ambos os intervalos decendiais no ano-safra 2013/2014 (Figuras 25a e 25c), quase todas as variáveis (Figuras 15b, 15d, 16b, 16d 18b e 18d) apresentaram semelhança na distribuição espacial, quando comparado com os grupos obtidos – exceto a

evapotranspiração, que apresentou apenas uma semelhança na distribuição espacial em relação ao agrupamento no intervalo decendial 1aDMDV1d (Figura 17d).

Aplicou-se também a Análise Estatística Multivariada (análise de componentes Principais e análise discriminante) nas medidas associadas às variáveis agrometeorológicas dos intervalos decendiais considerados em cada ano-safra.

Inicialmente selecionou-se o número de componentes principais (Tabela 12), com base no critério proposto por Ferreira (2008), que considera como ideal para a proporção acumulada da variância total valores iguais ou superiores a 70%, sem perda significativa de informação, reduzindo-se a dimensão dos dados originais. Desta maneira, tem-se a representação na Tabela 12 dos componentes principais, que explicam 81 e 86% dos seus respectivos intervalos de variações decendiais da variância total das variáveis originais.

No intervalo decendial 1aDMDV1d, o CP1 foi influenciado pelas variáveis precipitação (relação direta), evapotranspiração potencial, temperatura média e radiação solar, sendo esta com relação inversa. Por outro lado, o CP2 foi influenciado pela medida R da variável precipitação (relação direta).

Para o intervalo decendial 2aDMDV2d (Tabela 12), o CP1 foi influenciado principalmente pelas variáveis evapotranspiração potencial, temperatura média e radiação solar, de maneira negativa. No CP2 as variáveis que exercem influência são a precipitação e a medida do R, desta mesma variável

Tabela 12 Componentes Principais; Proporção da variância explicada; Proporção acumulada da variância total para as variáveis agrometeorológicas dos anos-safra em estudos

Variável	1aDMDV1d		2aDMDV2d	
	CP1	CP2	CP1	CP2
Precipitação acumulada	0,42	-0,28	0,35	-0,56
Medida do R da precipitação	-0,19	0,86	0,03	-0,76
Evapotranspiração potencial	-0,53	-0,25	-0,56	-0,20
Temperatura Média	-0,45	-0,31	-0,50	-0,24
Radiação solar	-0,54	-0,01	-0,55	0,01
Desvio padrão	1,75	0,98	1,70	1,19
Proporção da variância explicada	0,61	0,19	0,58	0,28
Proporção acumulada da variância total	0,61	0,81	0,58	0,86

1aDMDV1d: variação decendial de um decênio antes e um depois do DMDV; 2aDMDV2d: variação decendial de um decênio antes e um depois do DMDV; CP: Componente Principal

Na sequência, utilizou-se a análise discriminante nos CPs na variação decendial 1aDMDV1d (Equação 22) e 2aDMDV2d (Equação 23), gerando a função quadrática, utilizando uma amostra de treino de 500 pixels, em que esta determina a alocação dos indivíduos (variáveis) em grupos (ano-safra).

$$D^2(CP1, CP2) = 0,1375CP1^2 - 1,163CP2^2 + 0,324CP1CP2 + 0,675CP1 + 2,25CP2 - 0,173 \quad \text{Eq. (22)}$$

$$D^2(CP1, CP2) = 0,083CP1^2 + 0,50CP2^2 + 0,043CP1CP2 + 0,732CP1 + 1,364CP2 - 0,484 \quad \text{Eq. (23)}$$

em que:

$$\begin{cases} \text{Se } D^2(\text{CP1}, \text{CP2}) \geq 0, \text{ então o indivíduo pertence ao ano de 2011/2012} \\ \text{Caso contrário o indivíduo pertence ao ano de 2013/2014} \end{cases}$$

Em ambas as Equações 22 e 23, os valores dos coeficientes dos CP2 são maiores que os valores dos coeficientes dos CP1, ou seja, as variáveis que fazem parte do CP2 influenciam mais na classificação dos grupos.

Após determinar a função quadrática verificou-se a qualidade desta, por meio do Índice da Exatidão Global e a proporção da má classificação, sendo realizado em uma amostra de teste de 112 pixels.

Constata-se, na Tabela 13, que todos os índices são classificados para uma boa apresentação da função de discriminação (ANDERSON et al., 1976; MINGOTI, 2013).

Tabela 13 Índices de classificação de qualidade da função quadrática

Índices	1aDMDV1d	2aDMDV2d
Exatidão Global	0,72	0,77
Proporção de más classificações ano-safra 2011/2012	0,21	0,14
Proporção de más classificações ano-safra 2013/2014	0,06	0,08

1aDMDV1d: variação decenal de um decênio antes e um depois do DMDV;  
2aDMDV2d: variação decenal de um decênio antes e um depois do DMDV.

O gráfico obtido a partir do CP1 e CP2 (Figura 26a) do intervalo decenal 1aDMDV1d mostra que o ano-safra de 2011/2012, quando comparado com 2013/2014, apresentou menor uniformidade na distribuição de chuva (valores altos de R) (Figura 20d). Já o ano de 2013/2014 é composto por dois grupos de pixels que se diferenciam dos resultados observados em 2011/2012.

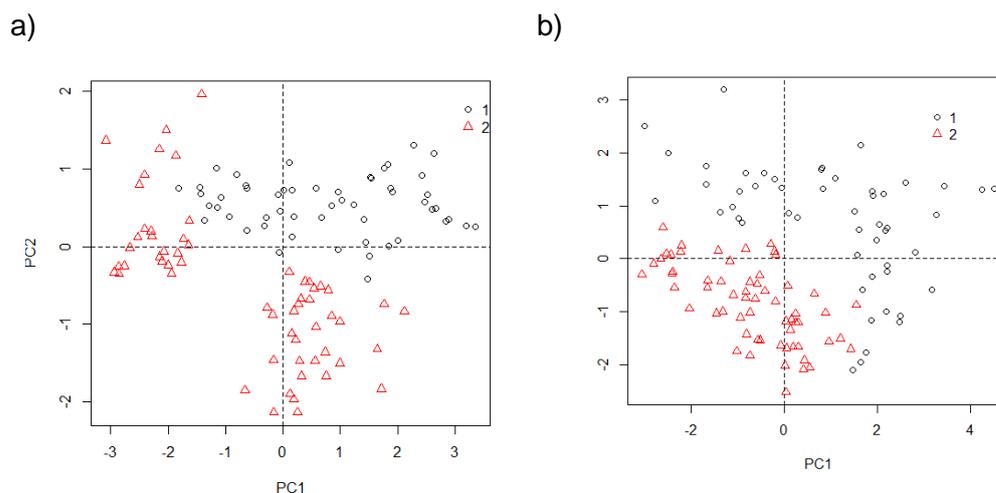


Figura 26 Representação gráfica da classificação obtida para o ano-safra 2011/2012 (o 1) e 2013/2014 ( $\Delta$  2): a) intervalo decenal 1aDMDV1d e b) 2aDMDV2d

O primeiro grupo composto por pixels classificados no ano-safra de maior produtividade (2013/2014) tem valores altos de PC2 e os valores mais baixos de PC1, maiores valores de temperatura (Figura 15b), radiação solar (Figura 16b) e evapotranspiração (Figura

17b), assim como menor precipitação acumulada (Figura 18b), o que significa que este grupo apresenta menor uniformidade na distribuição da precipitação (valores altos de R) (Figura 20p).

O segundo grupo composto por pixels classificados no ano-safra de maior produtividade (2013/2014) tem os menores valores de PC2 e uma grande variação dos valores de PC1 (valores intermediários a altos), o que significa que este grupo apresenta os menores valores de R (Figura 20p) e, conseqüentemente, uma maior uniformidade na distribuição da precipitação, sendo que a informação que está exercendo maior influência em distinguir entre os dois anos é a uniformidade da precipitação.

Na Figura 22b, tem-se a representação do intervalo decendial 2aDMDV2d do ano-safra de 2013/2014, que teve maior temperatura (Figura 15d), radiação solar (Figura 16d), evapotranspiração (Figura 17d), precipitação (Figura 18d) e menor uniformidade (maiores valores de R) (Figura 20t) na distribuição da precipitação no intervalo considerado quando comparado com 2011/2012, sendo que a informação que mais influencia para distinguir os dois anos-safra neste intervalo foram os valores da precipitação e a uniformidade da distribuição de chuva no intervalo considerado.

Por outro lado, no ano-safra de 2011/2012, houve menores valores de pelo menos um dos seguintes grupos de variáveis, denominado como grupo 1: temperatura (Figura 15c), radiação solar (Figura 16c) e evapotranspiração (Figura 17c); e como grupo 2: precipitação (Figura 18c) e valores de R (Figura 20h), ou seja, maior uniformidade na distribuição da precipitação do que no intervalo considerado anteriormente.

## 6. CONCLUSÃO

Verificou-se que as regiões com maiores e menores valores de produtividade da soja no estado do Paraná estão correlacionadas com os valores das variáveis agrometeorológicas no DMDV e em intervalos próximos.

Observou-se uma distribuição uniforme para as variáveis: temperatura média, radiação solar e evapotranspiração potencial; e para a precipitação uma sazonalidade moderada, com média decendial no DMDV e decêndios anteriores nos intervalos decendiais considerados nos dois anos-safras.

Por meio do Índice de Moran univariado e bivariado, verificou-se uma forte e significativa autocorrelação espacial das variáveis em estudo e uma correlação espacial entre a variável produtividade e as agrometeorológicas, exceto a precipitação foi negativa e significativa.

Em relação à associação espacial localizada, constatou-se para a variável produtividade e as medidas associadas às variáveis agrometeorológicas dois clusters, sendo um Alto-Alto e outro Baixo-Baixo.

No ano-safra 2011/2012, houve, em média, uma menor produtividade e uma maior amplitude dos valores das medidas associadas às variáveis agrometeorológicas quando comparado ao ano-safra 2013/2014.

## 7. CONSIDERAÇÕES FINAIS

Foram observadas algumas limitações na realização deste trabalho, como, por exemplo, a presença de dados faltantes das variáveis agrometeorológicas e das estações do modelo ECMWF.

Para dar continuidade a esta pesquisa, são apresentados a seguir algumas sugestões para trabalhos futuros:

- Considerar outras variáveis agrometeorológicas na análise, tais que: balanço hídrico, temperatura mínima e temperatura máxima;
- Utilizar os dados das variáveis agrometeorológicas e outras variações decendiais até contemplar todo o ciclo fenológico da planta e realizar uma análise no momento da semeadura e na colheita;
- Utilizar os dados provenientes de mapeamento da produtividade soja e realizar comparativo com os dados do IBGE;
- Aplicar na análise outras medidas da estatística circular como: mediana e interquartil e aplicar o teste Watson-Williams, que considera duas médias circulares amostrais para verificar se as amostras estimam a mesma média populacional entre os diferentes anos-safras.

## REFERÊNCIAS

- ADAPAR – AGÊNCIA DE DEFESA AGROPECUÁRIA DO PARANÁ. **Portaria ADAPAR nº 202, de 19 de julho de 2017**. Estabelece o período do vazio sanitário, as datas limites para a semeadura e colheita da soja e outras medidas de controle da ferrugem asiática *Phakopsora pachyrhizi* no Estado do Paraná. Curitiba, 2017.
- ALBERTO, C. M.; STRECK, N. A.; HELDWEIN, A. B; BURIOL, G. A.; MEDEIROS, S. L. P. Água no solo e rendimento do trigo, soja e milho associados ao El Niño Oscilação Sul. **Pesquisa agropecuária brasileira**, Brasília, v.4 1, n. 7, p. 1067-1075, 2006.
- ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. **Crop Evapotranspiration: guidelines for computing crop requirements**. Irrigation and Drainage Paper 56. Rome: FAO, 1998, 301 p.
- ALMEIDA, E. S. **Curso de econometria espacial aplicada**. Piracicaba: ESALQ-USP. 2004, 128 p.
- ANDERSON, J.F.; HARDY, E.E.; ROACH, J.T.; WITMER, R.E. A land use and land cover classification system for use with remote sensor data. U.S. **Geological Survey Professional Paper**, n. 964. USGS, Washington, D.C, 1976. 28p.
- ANJOS, I.B.; NERY, T. J. Variáveis meteorológicas associadas ao rendimento de grãos no Estado do Paraná. **Acta Scientiarum. Biological Sciences**, Maringá, v. 27, n. 1, p. 133-144, 2005.
- ANSELIN, D. The Moran Scatter Plot as an ESDA Tool to Assess Local Instability in Spatial Association. In: FISCHER, M.; SCHOTEN, H.; UNWIN, D. **Spatial Analytical Perspectives on GIS in Environmental and SocioEconomic Sciences**, 1996, p. 111-125.
- ANSELIN, L. **Exploring Spatial Data with Exploring Spatial Data with GeoDaT: a workbook**. Urbana: University of Illinois, p. 106-115, 2005.
- ANSELIN, L.; SYABRI, I.; KHO, Y. GeoDa: an introduction to spatial data analysis. **Geographical Analysis**, Columbus, OH, v. 38, n. 1, p. 5-22, 2006.
- ANSELIN, L.; TITA, G.; GORR; W. Spatial analyses of crime. In: DUFFE, D. (Ed.). **Measurement and Analysis of Crime and Justice**. Vol. 4. Criminal Justice, Washington, DC, 2000. p. 213-262.
- APARECIDO, L. E. O.; ROLIM, G. S.; RICHETTI, J.; SOUZA, P. S.; JOHANN, J. A. Koppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil. **Ciência e Agrotecnologia**, Lavras, v. 40, n. 4, p. 405-417, 2016.
- ARAÚJO, E. C.; URIBE-OPAZO, M. A. C.; JOHANN, J. A. Modelo de regressão espacial para estimativa da produtividade da soja associada a variáveis agrometeorológicas na região do oeste do Estado do Paraná. **Revista Engenharia Agrícola**, Jaboticabal, v. 34, n. 2, p. 286-299, 2014.
- ASSIS, J. P.; NETO, D. D.; MANFRON, P. A.; MARTINS, T. N.; SPAROVEK, G.; TIMM, L. C. Ajuste de séries históricas de temperatura e radiação solar global diária às funções densidade de probabilidade normal e log-normal, em Piracicaba, SP. **Revista Brasileira de Agrometeorologia**, v. 12, n. 1, p. 113-121, 2004.
- BATSCHLET, E. **Circular Statistics in Biology Data**. Cambridge: Academic Press, 1981, p. 03-61.

BECKER, W. R.; JOHANN, A.; RICHETTI, J. J.; SILVA, L. C. A. Data mining techniques for separation of summer crop based on satellite images. **Engenharia Agrícola**, Jaboticabal, n. 4, v. 4, p.750-759, 2017.

BESKOW, S.; MELLO, C. R.; FARIA, A. C.; SIMOES, M. C.; CALDEIRA, T. L.; NUNES, G. S. Índices de sazonalidade para regionalização hidrológica de vazões de estiagem no Rio Grande do Sul. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande-PB, v. 18, n. 7, p.748-754, 2014.

BETARELLI JUNIOR, A. A.; ALMEIDA, E. Os principais fatores internos e as exportações microrregionais brasileiras. **Revista de Economia Contemporânea**, Rio de Janeiro, v. 13, n. 2, p. 201-227, 2009.

BONATO, E. R. **Estress Soja**. Passo Fundo: Embrapa Trigo, 2000, p. 19-55.

BOX, G.; COX, D. An analysis of transformations. **Journal of the Royal Statistical Society**, Londres, v. 26, n. 2, p. 211- 252, 1964.

BRAIDO, L. M. H.; TOMMASELLI, J. T. G. Caracterização climática e dos anos extremos (chuvoso e seco): seus efeitos na produção de cana-de-açúcar, milho e soja para a região do Pontal do Paranapanema - SP. **Revista Formação**, n. 17, v. 1, p. 13-34, 2010.

CALÍNSKI, T.; HARABASZ, J. A dendrite method for cluster analysis. **Communications in Statistics**, v. 3, p. 1-27, 1974.

CÂMARA, G.; CARVALHO, M. S.; CRUZ, O. G.; CORREA, V. Análise Espacial de Área. In: FUCKS, S. D.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. **Análise Espacial de Dados Geográficos**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais – Divisão de Processamento de Imagens, p. 19-29, 2004.

CAMPOS, M. C. Modernização da agricultura, expansão da soja no Brasil e as transformações socioespaciais no Paraná. **Revista Geografar**, v.6, n.1, p.161-191, 2011.

COLUSSO, M. V. S.; PARRÉ, J. L.; ALMEIDA, E. S. Degradação ambiental e crescimento econômico: a curva de Kuznets ambiental para o Cerrado. **Revista de Economia e Agronegócio**, Viçosa, v.10, n. 3, p.335-358, 2013.

CONAB – COMPANHIA NACIONAL DE ABASTECIMENTO. ACOMPANHAMENTO DA SAFRA BRASILEIRA. **Grãos**, SAFRA 2017/18, 7º Levantamento, Vol. 5. p. 15, 2018.

CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v.37, n.1, p.35-46, 1991.

DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. V. M. **Análise Espacial de Dados Geográficos**. Brasília, EMBRAPA, 2004. 209 p.

DUTRA, M. S.; BIAZI, E. Combinação de técnicas para construção de um modelo de credit scoring, usando análise discriminante. **Movendo Ideias**, Belém, v. 13, n. 2, p. 23-32, 2008.

ECMWF – EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS. **Webpage da European Centre for Medium-Range Weather Forecasts**. 2018. Disponível em: <http://spirits.jrc.ec.europa.eu/download/downloadaddata/downloadmeteodata/>. Acesso em: 10 jan. 2018.

- EMBRAPA SOJA. **Resultados de pesquisa de soja 1991/92**. p. 816. doc.138. Londrina, Paraná, 1999.
- EMBRAPA SOJA. **Resultados do Manejo Integrado de Pragas da Soja na safra 2016/2017 no Paraná**. p.9, ed. 21, Londrina- PR, 2017.
- EMBRAPA SOJA. **Tecnologias de produção de soja – Região Central do Brasil 2014**. p.98. Sistemas de Produção n.16 Londrina, Paraná: Embrapa Soja, 2013.
- EMBRAPA SOJA. **Tecnologias de produção de soja Paraná 2003**. p. 11-29, Londrina, Paraná, 2003.
- EMBRAPA. **Ecofisiologia da Soja**. Circular Técnica 48, p. 01-08, Londrina, Paraná, 2007.
- EMBRAPA. **Irrigação de Hortaliças no Estado do Amapá**. Macapá, PA, p. 1-13, 2004.
- EMBRAPA. **Tecnologias de produção de soja região central do Brasil 2011**. p. 225. n. 14, Londrina- PR, 2010.
- EMBRAPA. **Variabilidade espacial e temporal da produção de soja no Paraná e definição de ambientes de produção**. p. 19, Londrina, Paraná: Embrapa Soja, 2016.
- EQUIPE ESTATCAMP. **Software Action**. Estatcamp- Consultoria em estatística e qualidade, São Carlos, SP, Brasil, 2018.
- ESRI. **ArcGIS for Windows Version 10.3**. License type ArcInfo. ESRI - Environmental Systems Research Institute, 2015.
- FARIAS, J. R. B.; NEUMAIER, N.; NEPOMUCENO, A. L. **Agrometeorologia dos cultivos**. Brasília, DF: INMET, 2009. p. 263-277.
- FEHR, W. R.; CAVINESS, C. E. **Stages of soybean development**. Special Report 80, Ames: Iowa State University of Science and Technology, p. 4-9, 1977.
- FELEMA, J.; MEDEIROS, E. R.; FERREIRA, C. R.; CAMARA, M. R. G.; NASCIMENTO, S. P. Um estudo da produtividade de feijão, do milho e da soja na agricultura paranaense, nos anos de 2000 e 2010: uma análise espacial. **Ensaio FEE**, Porto Alegre, v. 36, n.4, p. 817-842, 2016.
- FERREIRA, D. F. **Análise Multivariada**. Ministério da Educação e do Desporto Universidade Federal de Lavras, Departamento de Ciências Exatas. Lavras, MG, 2008.
- FERREIRA, W. P. M.; COSTA, L. C.; SOUZA, C. F. Teste de um modelo agrometeorológico para estudo da influência da variabilidade climática na cultura da soja. **Revista Ceres**, Viçosa, v. 54, p. 207-214, 2007.
- FONTANA, D. C.; PINTO, D. G.; JUNGES, A. H.; BREMM, C. Inferências sobre o calendário agrícola a partir de perfis temporais de NDVI/MODIS. **Bragantia**, Campinas, v. 74, n. 3, p.350-358, 2015.
- GOMES, F. P. **Curso de estatística experimental**. 15. ed. Piracicaba: Nobel, 1990.
- GRZEGOZEWSKI, D. M.; JOHANN, A. J.; URIBE-OPAZO, M. A.; MERCANTE, E.; COUTINHO, A. C. Mapping soya bean and corn crops in the State of Paraná, Brazil, using EVI

- images from the MODIS sensor. **International Journal of Remote Sensing**, n. 6, v. 37, p. 1257-1275, 2016.
- GRZEGOZEWSKI, D. M.; URIBE-OPAZO, M. A.; JOHANN, A. J.; GUEDES, L. P. C. Spatial correlation of soybean productivity, enhanced v vegetation index (EVI) and agrometeorological variables. **Engenharia Agrícola**, Jaboticabal, v.37, n.3, p.541-555, 2017.
- HEIFFIG, L. S.; CÂMARA, G. M. S.; MARQUES, L. A.; PEDROSO, D. B.; PIEDADE, S. M. S. Fechamento e índice de área foliar da cultura da soja em diferentes arranjos espaciais. **Bragantia**, Campinas, v. 65, n. 2, p. 285-295, 2006.
- IBGE – INSTITUTO BRASILEIRO DE GEOGRAFICA E ESTATÍSTICA. **Relatórios metodológicos - Pesquisas Agropecuárias - Departamento de Agropecuária**. 2. ed. Rio de Janeiro: IBGE, 2018.
- INMET. **Agrometeorologia dos cultivos: o fator Meteorológico na produção agrícola**. Brasília: Ministério da Agricultura e do Abastecimento, 2009.
- JAMMALAMADAKA, S.R.; SENGUPTA, A. **Topics in Circular Statistics**. River Edge, NJ: World Scientific, 2001. p. 1-131.
- JOHANN, J. A.; BECKER, W. R.; URIBE-OPAZO, M. A; MERCANTE, E. Uso de imagens do sensor orbital MODIS na estimação de datas do ciclo de desenvolvimento da cultura da soja para o estado do Paraná – Brasil. **Engenharia Agrícola**, Jaboticabal, v. 3, n. 1, p. 126-142, 2016.
- JONES, T. A. Skewness and kurtosis as criteria of normality in observed frequency distributions. **Journal of Sedimentary Research**. Colorado, v. 39, n. 4, p. 1622-1627, 1969.
- KLOSOWSKI, E. S. Estimativa da Produtividade de Cultivares de Soja por meio do Modelo Soygro para Londrina, Estado do Paraná. **Revista UNIMAR**, Maringá, v 19, n. 3, p. 751-765, 1997.
- KOLLER, A. A.; TAYLOR, K.; RAUN, W. B.; WECKLER, P. R.; BUSER, M. D. Validation of maize seed orientation by pushing. **Biosystems Engineering**, v. 151, p. 338-349, 2016.
- KRIPPENDORFF, K. **Content Analysis: an introduction to its methodology**. Newbury Park, CA: Sage, 1980. 189p.
- LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, USA, v. 33, p. 159-174,1977.
- LEE, J.; KWON, H.; KIM, T. Spatio-temporal analysis of extreme precipitation regimes across South Korea and its application to regionalization. **Journal of Hydro-Environment Research**, v. 6, n. 2, p. 101-110, 2012.
- LESAGE, J. P. **Spatial Econometrics**. Toledo, OH: University of Toledo, 1999. p. 1-20.
- LUZARDO, A. J. R.; CASTAÑEDA FILHO, R. M. C.; RUBIM, I. B. Análise Espacial Exploratória com o emprego do Índice de Moran. **Revista Geographia**, Niterói, v.19, n. 40, p. 61-179, 2017.
- MACEDO, P. B. R.; SIMÕES, R. Amenidades Urbanas e Correlação Espacial: uma Análise Intra-urbana para BH (MG). **Revista Brasileira de Educação (RBE)**, Rio de Janeiro, v. 52, n. 4, p. 525-539, 1998.

- MANLY, B. F. J. **Métodos Estatísticos Multivariada - uma Introdução**. 3a. ed. Porto Alegre, 2008.
- MARDIA, K. V. **Statistics of Directional Data**. Londres: Academic Press, 1972. p. 13-18.
- MARDIA, K. V.; JUPP, P. E. **Directional Statistics**. West Sussex: John Wiley & Sons, 1999. 451 p.
- MARDIA, K. V.; KENT, J. T.; BILEY, J. M. **Multivariate Analysis**. London: Academic Press, 1995. 321 p.
- MINGOTI, S.A. **Análise de dados através de métodos de estatística multivariada uma abordagem aplicada**. Belo Horizonte: UFMG. 2013.
- MONTENEGRO, R. L. G.; BETARELLI JUNIOR, A. A. Análise e investigação dos fatores determinantes da inovação nos municípios de São Paulo. **Revista Associação Brasileira de Estudos Regionais e Urbanos**, Curitiba, v. 2, n. 2, p. 3-31. 2008.
- NETO, R. L. B.; ARAÚJO, E. I. P.; MACIEL, C. M. S.; PAULA, A.; TAGLIAFERRE, C. Fenologia de *Astronium graveolens* Jacq. em floresta estacional decidual em Vitória da Conquista, Bahia. **Ciência Florestal**, Santa Maria, v. 28, n.1, p. 641-650, 2018.
- NEWSON, S. W. B. Pioneers in infection control: John Snow, Henry Whitehead, the Broad Street pump, and the beginnings of geographical epidemiology. **Journal of Hospital Infection**, Londres, v. 64, n. 3, p. 210-216, 2006.
- PEWSEY, A.; NEUHÄUSER, M.; RUXTON, G. D. **Circular Statistics in R**. Oxford: Oxford University Press, 2013. p.11-56.
- PINTO, N. L.; HOLTS, A. C. T; MARTINS, J. A.; GOMIDE, F. L. S. **Hidrologia Básica**. 2. ed. São Paulo: Edgard Blucher, 1976, p. 56-59.
- PRUDENTE, V. H. R.; SOUZA, C. H. W.; MERCANTE, E.; JOHANN, J. A.; OPAZO, M. A. U. Spatial statistics applied to soybean production data from Paraná state from 2003-04 to 2009-10 crop-years. **Engenharia Agrícola**, Jaboticabal, v. 34, n. 4, p. 755-769, 2014.
- R DEVELOPMENT CORE TEAM. **R: a Language and Environment for Statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- RADIN, B.; MATZENAUER, R. Uso das informações meteorológicas na agricultura do Rio Grande do Sul. **Agrometeoros**, Passo Fundo, v.24, n.1, p.41-54, 2016.
- RADIN, B.; SHONHOFEN, A.; TAZZO, I. F. Impacto da quantidade e frequência de chuva no rendimento. **Revista Agrometeoros**, Passo Fundo – RS, v. 25, n.1, p.19-26, 2017.
- RODRIGUES, O.; TEIXEIRA, M. C. C.; LHAMBY, J. C. B.; BONATO, E. R.; BERTAGNOLLI, P. F. Sistema tardio de semeadura de soja (STS). **Circular Técnica Online**, 11. Passo Fundo: Embrapa Trigo, 2002, 16 p.
- SILVA, R. R.; SANTORI, M. G. B.; WOLLMANN, C. A. Relação entre precipitação pluviométrica e produtividade da cultura de soja, no município de Ibirubá- RS. **Revista do Departamento de Geografia – USP**, v. 27, p. 296-314, 2014.
- SOUZA, C. H. W. de; MERCANTE, E.; JOHANN, J. A.; LAMPARELLI, R. A. C.; URIBE-OPAZO, M. A. Mapping and discrimination of soybean and corn crops using spectrum temporal

profiles of vegetation indices. **International Journal of Remote Sensing**, v. 36, p. 1809-1824, 2015.

THOMAS, B. **Photoperiodism**. Vol.1. Coventry, UK: University of Warwick 2017. p. 448-454.

TING, S.; HARTLEY, S.; BURNS, K. C. Global patterns in fruiting seasons. **Global Ecology and Biogeography**, New Zealand, v. 17, p. 648-657, 2008.

TORRE, G.; VOSSENKEMPER, J.; RAUN, W.; TAULOR, R. Maize (Zea Mays) leaf angle and emergence as affected by seed orientation at planting. **Explore Agriculture**, London, v. 47, n. 4, p. 579-592, 2011.

URIBE-OPAZO, M.A.; ROJAS, M. G. **Geoestadística aplicada a la Ingeniería Agrícola**. Cascavel: Edunioeste, 2015.

VAREJÃO-SILVA, M. D. **Meteorologia e Climatologia**. Recife, Pernambuco: online, 2005. p. 17-22.

VIDIGAL, V. G.; VIDIGAL, C. B. R.; PARRÉ, J. L. Distribuição espacial da produtividade da soja no Rio Grande do Sul: um estudo exploratório. **Acta Scientiarum. Human and Social Sciences**, Maringá, v. 40, p. 2-9, 2018.

VIVAN, G. A.; ROBAINA, A. D.; PEITER, M. X.; PARIZI, A. R. C.; BARBOZA, F. S.; SOARES, F. C. Rendimento e rentabilidade das culturas da soja, milho e feijão cultivados sob condições de sequeiro. **Semina: Ciências Agrárias**, Londrina, v. 36, n. 5, p. 2943-2950, 2015.

WARTENBERG, D. Multivariate spatial correlation: a method for exploratory geographical. **Analysis Geographical**, Columbus, Analysis Geographical, OH, v. 7, n. 4, p. 263-265, 1985.

YEO, K.; JOHNSON, R. A. A new family of power transformations to improve normality or symmetry. **Journal Biometrika**, Oxford, v. 87, p. 954-959, 2000.

ZAR, J. H. **Biostatistical Analysis**. 5a. ed. New Jersey: Prentice Hall. 2010. 609 p.