

ISSN 1678-9644

Novembro, 2013

*Empresa Brasileira de Pesquisa Agropecuária
Embrapa Arroz e Feijão
Ministério da Agricultura, Pecuária e Abastecimento*

Documentos 293

Subdivisão da Região Produtora de Arroz de Terras Altas Baseada em Variáveis Climáticas

*Felipe Franco Mendes
Alexandre Bryan Heinemann
Alessandra da Cunha Moraes
Silvando Carlos da Silva*

Embrapa Arroz e Feijão
Santo Antônio de Goiás, GO
2013

Exemplares desta publicação podem ser adquiridos na:

Embrapa Arroz e Feijão

Rod. GO 462, Km 12
Caixa Postal 179
75375-000 Santo Antônio de Goiás, GO
Fone: (0xx62) 3533 2110
Fax: (0xx62) 3533 2123
www.cnpaf.embrapa.br
cnpaf.sac@embrapa.br

Comitê de Publicações

Presidente: *Roselene de Queiroz Chaves*
Secretário-Executivo: *Luiz Roberto Rocha da Silva*
Membros: *Ana Lúcia Delalibera de Faria*
Flávia Aparecida de Alcântara
Heloísa Célis Breseghello
Fábio Fernandes Nolêto
Luís Fernando Stone
Márcia Gonzaga de Castro Oliveira
Camilla Souza de Oliveira

Supervisor editorial: *Camilla Souza de Oliveira*
Revisão de texto: *Camilla Souza de Oliveira*
Normalização bibliográfica: *Ana Lúcia D. de Faria*
Tratamento de ilustrações: *Fabiano Severino*
Editoração eletrônica: *Fabiano Severino*

1ª edição

Versão online (2013)

Todos os direitos reservados

A reprodução não-autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)
Embrapa Arroz e Feijão

Subdivisão de região produtora de arroz de terras altas baseada em variáveis climáticas / Felipe Franco Mendes... [et al.]. - Santo Antônio de Goiás : Embrapa Arroz e Feijão, 2013.
28 p. - (Documentos / Embrapa Arroz e Feijão, ISSN 1678-9644 ; 293)

1. Arroz – Variação sazonal. 2. Arroz – Climatologia. I. Mendes, Felipe Franco. II. Embrapa Arroz e Feijão. III. Série.

CDD 633.18 (21. ed.)

© Embrapa 2013

Autores

Felipe Franco Mendes

Graduando em Estatística pela Universidade Federal de Goiás – UFG, bolsista da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, felipe.mendes951@hotmail.com

Alexandre Bryan Heinemann

Engenheiro agrônomo, Doutor em Irrigação e Drenagem, pesquisador da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, alexandre.heinemann@embrapa.br

Alessandra da Cunha Moraes

Bacharel em Tecnologia em Geoprocessamento, analista da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, alessandra.moraes@embrapa.br

Silvando Carlos da Silva

Engenheiro agrícola, Mestre em Agrometeorologia, pesquisador da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, silvando.silva@embrapa.br

Apresentação

As variáveis climáticas regem o crescimento e o desenvolvimento de plantas. Conhecendo-se esses elementos de uma região, poder-se-á caracterizá-la climaticamente, possibilitando definir estratégias de manejo de cultivos que minimizem os riscos oriundos das alterações na disponibilidade hídrica, energética, na variação térmica, e de eventos climáticos extremos, como também, a sua influência na incidência de pragas e doenças.

Em função disso, torna-se possível reduzir os efeitos negativos decorrentes da complexa interação entre planta e variáveis meteorológicas. Portanto, é de suma importância a utilização de métodos para a definição de áreas climaticamente homogêneas para subsidiar, principalmente, as atividades relacionadas ao melhoramento de plantas.

Assim, esse documento objetiva verificar se existe diferença no agrupamento climático utilizando-se variáveis meteorológicas anuais, mensais e relativas aos meses de safra (novembro-abril), considerando-se o cultivo do arroz de terras altas.

Os autores

Sumário

Introdução	9
Materiais e Métodos	10
Descrição dos dados	10
Análise de componentes principais (PCA)	12
Agrupamentos	13
Análise de concordância	13
Espacialização dos agrupamentos.....	14
Análise estatística	14
Resultados	14
Resultado da análise de PCA	14
Agrupamentos	15
Caracterização dos conjuntos de dados.....	19
Dados resumidos (conjunto de dados 1).....	19
Dados dos meses da safra (conjunto de dados 2)	22
Dados dos meses completos (conjunto de dados 3).....	24
Conclusões	27
Referências	27

Subdivisão da Região Produtora de Arroz de Terras Altas Baseada em Variáveis Climáticas

Felipe Franco Mendes

Alexandre Bryan Heinemann

Alessandra da Cunha Moraes

Silvando Carlos da Silva

Introdução

O programa de melhoramento de planta tem como objetivo desenvolver cultivares de alta produtividade para uma vasta região. Entretanto, nesse processo os melhoristas necessitam decidir por uma seleção para toda a região produtora ou explorar os ganhos de produtividade que podem ser obtidos por meio da subdivisão da área de produção, de maneira a explorar a adaptação específica (ATLIN et al., 2000). A estratificação da área de produção em regiões homogêneas, caso haja diversidade agroecológica, permite uma seleção com acurácia de variedades candidatas na região alvo, possibilitando um ganho na seleção devido à redução na interação complexa entre genótipo e ambiente. A subdivisão da área de produção também permite a otimização da alocação dos ensaios de valor de cultivo e uso (VCU), de forma a explorar a diversidade agroecológica da região. No Brasil, para a cultura do arroz de terras altas, a estratificação da região produtora de arroz de terras altas foi realizada por Heinemann e Sentelhas (2011), que utilizaram dados de produtividade do arroz provenientes do IBGE (Instituto Brasileiro de Geografia e Estatística) para 54 microrregiões. Nesse estudo foram obtidos três grupos ambientais denominados altamente favorável, favorável e pouco favorável. Esse estudo explora a utilização de variáveis climáticas para subdividir a região produtora de arroz nos estados de RO, GO, TO e MT. Assim, o objetivo específico

desse estudo é verificar se há diferença no agrupamento climático utilizando-se variáveis climáticas anuais, variáveis climáticas mensais e variáveis climáticas relativas aos meses de safra (novembro-abril).

Materiais e Métodos

Descrição dos dados

A área de abrangência desse estudo são os estados de GO, TO, MT e RO. As variáveis climáticas utilizadas foram precipitação, temperatura máxima e mínima do ar, radiação solar global, TAVE (média entre a temperatura máxima e mínima de cada dia) e TAMP (diferença entre a temperatura máxima e a mínima de cada dia), todas provenientes da base de dados climáticos da Embrapa Arroz e Feijão. As estações meteorológicas utilizadas estão descritas na Tabela 1. Nesse estudo utilizou-se a média mensal das séries históricas das variáveis climáticas de 1980 a 2012 descritas acima. Essas médias mensais foram agrupadas em três conjuntos de dados: a) conjunto 1 (composto pela média anual proveniente das médias mensais das variáveis climáticas – dados resumidos), b) conjunto 2 (composto pela média mensal dos meses de safra – novembro a abril) e c) conjunto 3 (composto pelas médias mensais dos 12 meses do ano). Assim, formou-se três diferentes matrizes. A primeira matriz, referente aos dados resumidos, é composta por sete colunas (média anual da precipitação pluvial, temperatura máxima, mínima do ar, radiação solar global, TAVE, TAMP e latitude). A segunda matriz é composta por 25 colunas, sendo seis colunas que representam a precipitação, seis colunas para a temperatura máxima do ar, seis colunas para temperatura mínima do ar, seis colunas para a radiação solar, nos meses da safra, e uma coluna para a latitude. A terceira matriz é composta 49 colunas, sendo 12 colunas que representam a precipitação, 12 colunas para a temperatura máxima do ar, 12 colunas para temperatura mínima do ar, 12 colunas para a radiação solar referentes aos meses do ano e uma coluna para a latitude. Para as três matrizes, as linhas são compostas pelas 54 estações meteorológicas.

Tabela 1. Lista das estações meteorológicas e seus respectivos identificadores (ID) e número de anos.

<i>Cidade</i>	<i>UF</i>	<i>ID</i>	<i>Número de anos</i>	<i>Cidade</i>	<i>UF</i>	<i>ID</i>	<i>Número de anos</i>
Araguaína	TO	1	22	Goiás	GO	28	32
Campos Lindos	TO	2	22	Faina	GO	29	32
Palmas	TO	3	22	Aruanã	GO	30	32
Gurupi	TO	4	22	Aragarças	GO	31	32
Porangatu	GO	5	32	Nova Xavantina	MT	32	32
Monte Alegre de Goiás	GO	6	32	Canarana	MT	33	32
Posse	GO	7	32	Paranatinga	MT	34	32
Formosa	GO	8	32	Campo Verde	MT	35	32
Luziânia	GO	9	32	Rondonópolis	MT	36	32
Cristalina	GO	10	32	Santo A. do Leverger	MT	37	32
Catalão	GO	11	32	Cuiabá	MT	38	32
Ipameri	GO	12	32	São José do Rio Claro	MT	39	32
Caldas Novas	GO	13	32	Diamantino	MT	40	32
Morrinhos	GO	14	32	Matupá	MT	41	32
Itumbiara	GO	15	32	Alta Floresta	MT	42	32
Quirinópolis	GO	16	32	Porto dos Gaúchos	MT	43	32
Bom Jesus de Goiás	GO	17	32	Juína	MT	44	32
Paraúna	GO	18	32	Comodoro	MT	45	32
Rio	GO	19	32	Vila Bela da S. T	MT	46	32
Jataí	GO	20	32	Vilhena	RO	47	32
Mineiros	GO	21	32	Cacoal	RO	48	32
Caiapônia	GO	22	32	Ji-Paraná	RO	49	32
Varjão	GO	23	32	Aripuanã	MT	50	32
Goiânia	GO	24	32	Machadinho D'oeste	RO	51	32
Santo Antônio de Goiás	GO	25	32	Ariquemes	RO	52	32
Anápolis	GO	26	32	Guajará-Mirim	RO	53	32
Pirenópolis	GO	27	32	Porto Velho	RO	54	32

Análise de componentes principais (PCA)

Nesse estudo aplicou-se o método de análise de componentes principais com o objetivo de reduzir a dimensão dos dados originais e facilitar a interpretação das análises realizadas. As componentes principais não correlacionadas apresentam variâncias iguais aos valores próprios da matriz de covariância (Σ) (FERREIRA, 2000). A esperança (equação 1) e variância da componente Y_i (equação 2) são, respectivamente, iguais a:

$$E[Y_i] = e'_i \mu = e'_{i1} \mu_1 + e'_{i2} \mu_2 + \dots + e'_{ip} \mu_p \quad (\text{eq. 1})$$

$$\text{Var}[Y_i] = e'_i \Sigma_{p \times p} e_i = \lambda_i \quad (\text{eq. 2})$$

Sendo $\text{Cov}[Y_i, Y_k] = 0, j \neq k$. Cada autovalor λ_i representa a variância de uma componente principal Y_i . Como os autovalores estão ordenados em ordem decrescente, a primeira componente é a de maior variabilidade e a p -ésima é a de menor. A proporção da variância total de X que é explicada pela i -ésima componente principal está descrita pela equação 3. Já a proporção da variância total que é explicada pelas k primeiras componentes principais está descrita na equação 4 (MINGOTI, 2005).

$$\frac{\text{Var}[Y_i]}{\text{Variância Total de } X} = \frac{\lambda_i}{\text{Traço}(\Sigma_{p \times p})} = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \quad (\text{eq. 3})$$

$$\frac{\sum_{i=1}^k \text{Var}[Y_i]}{\text{Variância Total de } X} = \frac{\sum_{i=1}^k \lambda_i}{\text{Traço}(\Sigma_{p \times p})} = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (\text{eq. 4})$$

As componentes principais obtidas a partir de uma matriz de covariância $\Sigma_{p \times p}$ são, em geral, influenciadas pelas variáveis de maior variância, sendo, portanto, de pouca utilidade nos casos em que existe uma discrepância muito acentuada entre essas variâncias. A discrepância pode ser causada pela diferença das unidades de medidas das variáveis. Este problema é amenizado por meio da transformação dos dados originais, de modo a equilibrar os valores de variância ou colocar os dados na mesma escala de medida. Nesse estudo foi realizada transformação na qual cada variável é padronizada pela matriz de covariâncias das variáveis padronizadas.

Em muitas situações em que se aplicam os componentes principais, se uma porcentagem de 70% ou mais for atribuída aos primeiros r componentes principais, então, esses podem substituir as p variáveis originais sem perda de informações. A determinação dessa porcentagem da variação explicada pelos primeiros r componentes é subjetiva. Nesse estudo utilizou-se como critério adotar o número de componentes principais que expliquem 90% da variabilidade dos conjuntos de dados.

Agrupamentos

A análise de *cluster* é uma técnica exploratória de dados que pesquisa a existência de grupos naturais de objetos ou de variáveis. A aplicação desta técnica não pressupõe qualquer característica da estrutura do agrupamento; apenas se baseia em medidas de semelhança ou distância entre objetos e na escolha de critérios de agregação (FERREIRA, 2000).

As componentes principais que expliquem mais que 90% da variância serão selecionadas na análise de componentes principais e utilizadas como variáveis na análise de *cluster*. Para aplicar a análise de *cluster*, neste estudo, optou-se pelo método de agrupamento hierárquico aglomerativo. Primeiro, os objetos mais similares são agrupados formando um único grupo. Eventualmente, o processo é repetido, e com o decréscimo da similaridade, todos os subgrupos são agrupados, formando um único grupo com todos os objetos. Para medir a similaridade dos objetos escolheu-se a distância euclidiana e o critério de aglomeração foi realizado pelo método de "Ward". A variabilidade dentro dos grupos para cada etapa do método de "Ward" é utilizada para definir o melhor número de grupos para o conjunto de dados. A partição sugerida é aquela com a maior perda relativa de variabilidade. Este método é definido por Husson et al. (2011) como "*inertiagain*".

Análise de concordância

A semelhança entre os agrupamentos foi quantificada utilizando-se o coeficiente concordância de Kappa. Esse método compara duas classificações de uma mesma variável e quantifica o grau de concordância considerando as classificações idênticas obtidas pelo acaso (NUNES, 2006).

De acordo com os critérios definidos por Landis e Koch (1977), o grau de concordância entre duas observações, dado pelo valor do coeficiente k , pode ser interpretado de acordo com a Tabela 2.

Tabela 2. Coeficiente de concordância de Kappa de Cohen.

<i>Estatística de Kappa</i>	<i>Grau de concordância</i>
< 0,00	Desprezável
0,01 – 0,20	Ruim
0,21 – 0,40	Leve
0,41 – 0,60	Moderado
0,61 – 0,80	Substancial
0,81 – 1,00	Quase perfeito

Espacialização dos agrupamentos

Após a determinação dos agrupamentos, os mesmos foram espacializados na região produtora de arroz de terras altas utilizando-se o método polígono de Thiessen (HEGGEN, 1996), que considera somente a distância física entre as estações meteorológicas, com o objetivo de obter uma transição discreta entre os grupos e facilitar a localização geográfica dos diferentes agrupamentos.

Análise estatística

As análises estatísticas descritivas (gráficos, médias e desvios-padrão) foram feitas utilizando-se o software livre R, versão 2.15.0 (R DEVELOPMENT CORE TEAM, 2012), e os pacotes FactoMineRv (HUSSON et al., 2013) para as análises do componente principal, psych (REVELLE, 2013) para calcular o coeficiente de Kappa e MASS para análise de *cluster*.

Resultados

Resultado da análise de PCA

A análise de componente principal foi realizada para cada conjunto de dados, com o intuito de reduzir a quantidade de variáveis a serem utilizadas na análise de *cluster*. A Figura 1 ilustra para cada conjunto de dados (1, 2 e 3) o número de componentes principais que explicam 90% da variância.

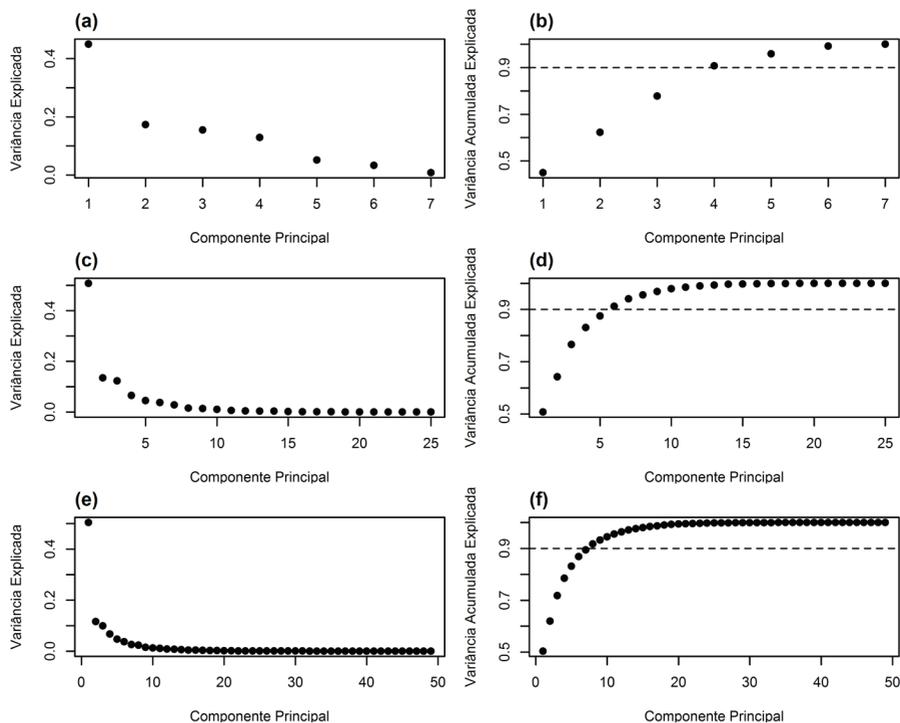


Figura 1. Variância explicada pelas componentes principais para os conjuntos de dados 1 (dados resumidos - a e b), 2 (meses da safra - c e d) e 3 (todos os meses do ano -e e f).

Para o conjunto de dados 1 (dados resumidos), Figura 1 (b), os quatro primeiros componentes principais explicam 90% da variância. Já para o conjunto de dados 2 (meses da safra), Figura 2 (d), os seis primeiros componentes principais explicam 90% da variância e para o conjunto de dados 3 (todos os meses do ano) os oito primeiros componentes principais explicam 90% da variância. Esses componentes principais, que explicam no mínimo 90% da variância, foram utilizados para a análise de agrupamento (*cluster*) para os conjuntos de dados 1, 2 e 3.

Agrupamentos

Para a realização dos agrupamentos para cada conjunto de dados (1, 2 e 3) utilizou-se a análise de *cluster*. A Figura 2 ilustra o número de grupos

definidos para os conjuntos de dados 1, 2 e 3, pelo método "inertiagain". Para o conjunto de dados 1 (dados resumidos), 2 (meses da safra) e 3 (todos os meses) foram obtidos 4 (Figura 2a), 6 (Figura 2b) e 4 (Figura 2c) grupos, respectivamente.

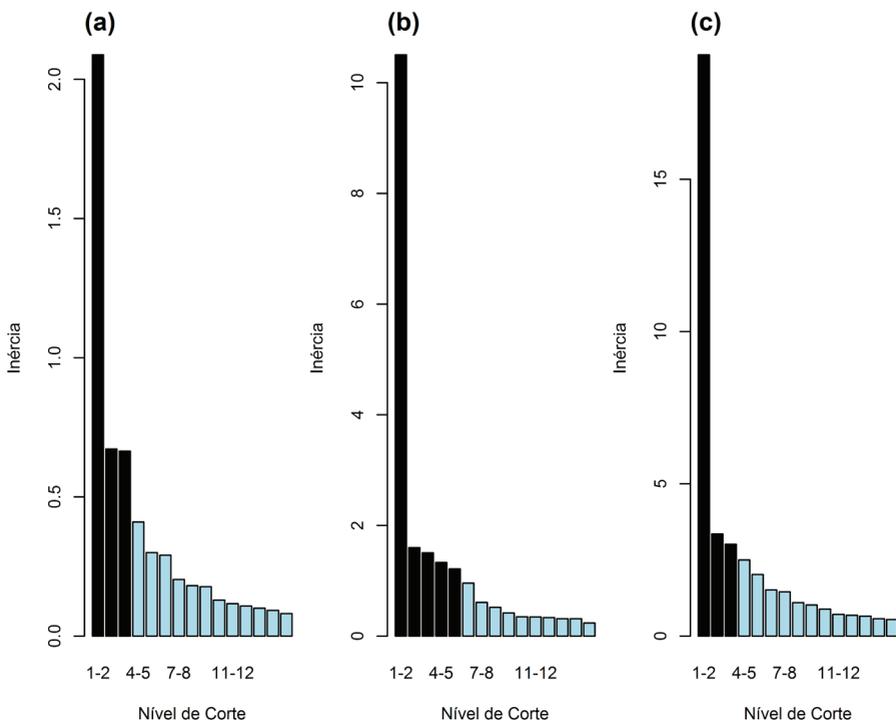


Figura 2. Número de grupos definidos pelo método "inertiagain" para os conjuntos de dados 1 (a), 2 (b) e 3 (c).

A quantificação da semelhança entre os agrupamentos, calculado pelo coeficiente de Kappa é descrita na Tabela 3.

Tabela 3. Coeficiente de Kappa para a quantificação das semelhanças entre os conjuntos de dados.

Conjunto de dados	Valor para o coeficiente de Kappa
Conjuntos 1 e 3	0,59
Conjuntos 1 e 2	0,11
Conjuntos 2 e 3	0,23

Para os conjuntos de dados 1 e 3, o valor do coeficiente de Kappa foi de 0,59 (Tabela 3), havendo evidências de semelhança moderada entre esses conjuntos. Já, na comparação entre os conjuntos 2 e 3, o valor do coeficiente de Kappa é baixo, de apenas 0,23. Isso indica que o agrupamento realizado pelo conjunto 2 (meses safra) é diferente do agrupamento do conjunto 3 (todos os meses). O valor do coeficiente de Kappa entre os conjuntos 1 e 2, também foi baixo, 0,11, indicando diferença entre esses conjuntos.

Nos conjuntos de dados 1 e 3 (Figuras 3 e 4), os grupos 1 e 2 são formados, predominantemente, por estações meteorológicas localizadas no Estado de Goiás. Para o conjunto de dados 3 (todos os meses) (Figura 7), o grupo 3 é formado apenas por estações meteorológicas do Estado de Mato Grosso. Para o conjunto de dados 1 e 3, o grupo 4 é mais heterogêneo.

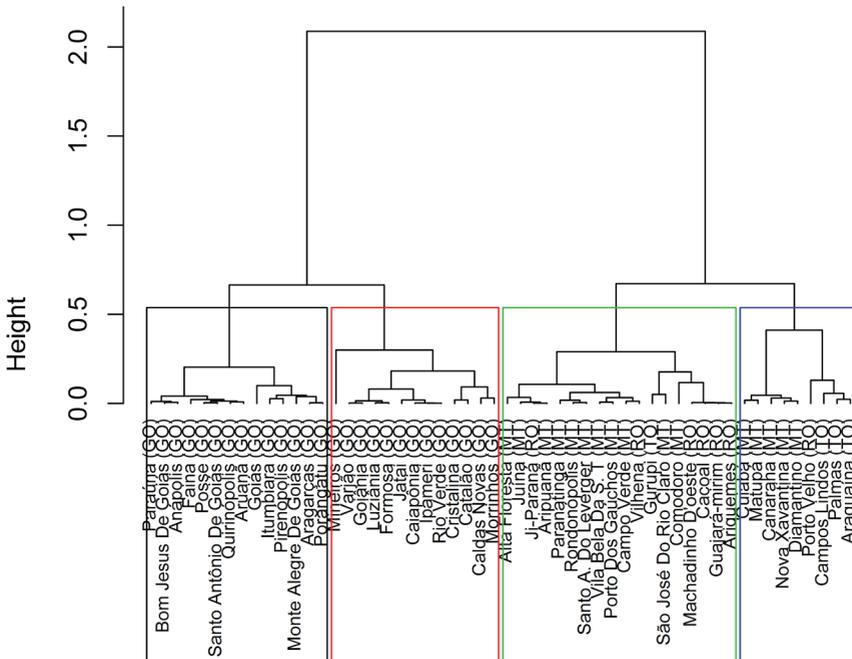


Figura 3. Dendrograma com agrupamento pelo método Ward para o conjunto de dados 1 (dados resumidos). Grupo 1 - Vermelho, Grupo 2 - Preto, Grupo 3 - Azul, Grupo 4 - Verde.

Os grupos do conjunto de dados 2 (Figura 5), são os que mais diferenciaram em relação aos outros conjuntos de dados, tanto em quantidade quanto em composição. Nesse conjunto as estações meteorológicas de Morrinhos e Caldas Novas, ambas de Goiás, formaram dois grupos isolados.

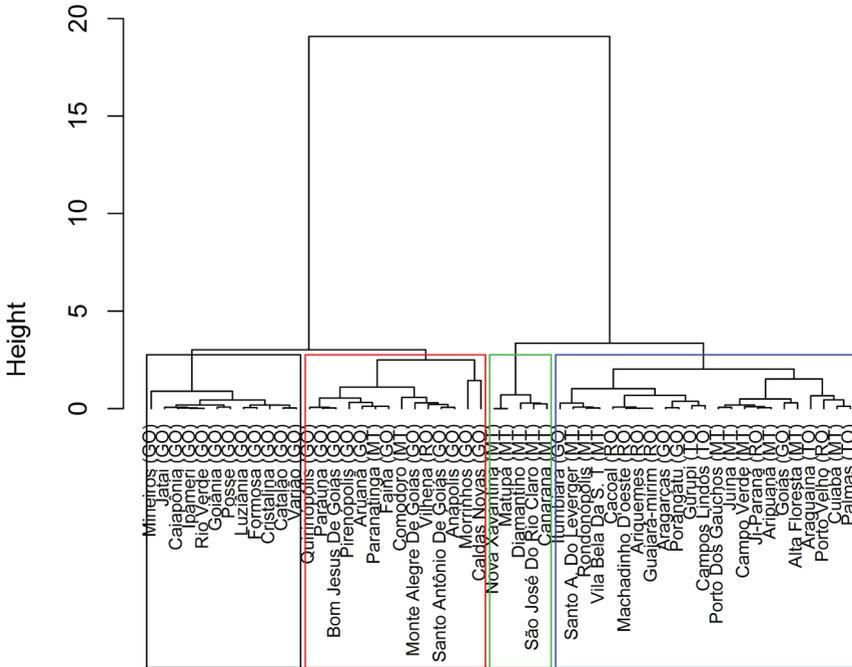


Figura 4. Dendrograma com agrupamento pelo método de Ward para o conjunto de dados 3 (dados com todos os meses). Grupo 1 - Preto, Grupo 2 - Vermelho, Grupo 3 - Azul, Grupo 4 - Verde.

Para uma análise baseada somente nos valores do coeficiente de Kappa (Tabela 3), a utilização do conjunto de dados 1 (dados resumidos) tem o mesmo efeito no agrupamento de se utilizar o conjunto de dados 3 (dados de todos os meses). Entretanto, o mesmo não é verdadeiro para o conjunto de dados 2 (meses da safra).

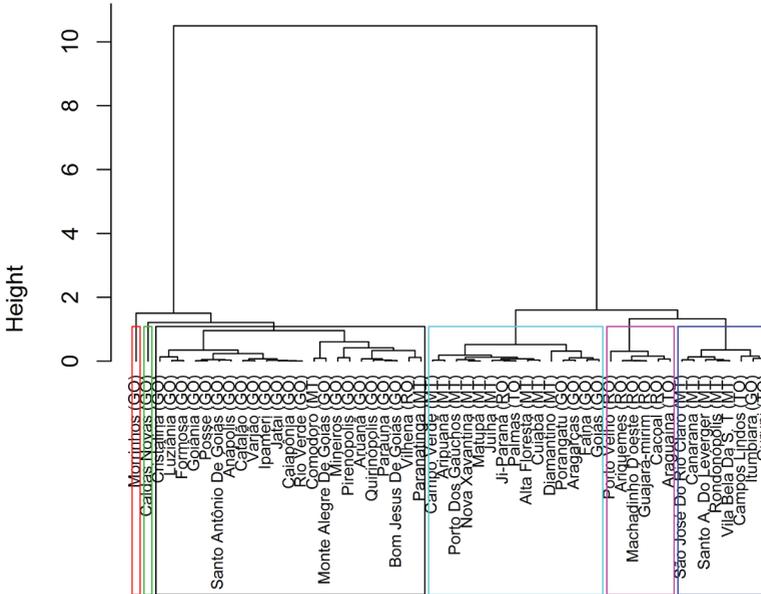


Figura 5. Dendrograma com agrupamento pelo método de Ward para o conjunto de dados 2 (dados dos meses da safra). Grupo 1 - Vermelho, Grupo 2 - Preto, Grupo 3 - Verde, Grupo 4 - Azul Claro, Grupo 5 - Roxo, Grupo 6 - Azul.

Caracterização dos conjuntos de dados

Dados resumidos (conjunto de dados 1)

A Figura 6 ilustra a distribuição espacial dos diferentes grupos pertencentes ao conjunto de dados 1. Entretanto, a Figura 7 ilustra a variabilidade das variáveis climáticas: precipitação pluvial, temperatura máxima e mínima do ar e radiação solar global para os quatro agrupamentos. Nesse conjunto de dados, o grupo 3 é representado pelas estações do Estado de Mato Grosso localizadas no norte do estado, que apresentou a maior mediana para a precipitação pluvial (Figura 7a). Nesse grupo também está incluído Campos Lindos, no Estado do Tocantins, que é responsável pelo “outlier” inferior. Esse grupo apresenta o maior valor para o desvio padrão da precipitação pluvial (302 mm, Tabela 4). Os grupos 1 e 2 possuem uma menor variância para a precipitação pluvial e, em média, as estações meteorológicas desses grupos contêm os menores índices de precipitação pluvial, como pode ser observado na Tabela 4. Já,

o grupo 4 destaca-se por sua alta variabilidade. Comodoro (MT) é a estação com menor quantidade de precipitação pluvial nesse grupo e Alta Floresta (MT) com a maior quantidade entre todas as estações meteorológicas utilizadas nesse estudo.

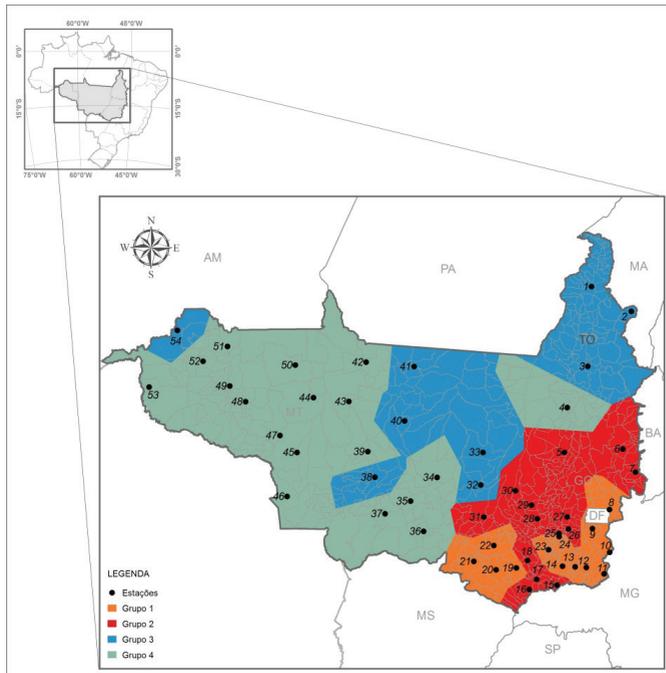


Figura 6. Distribuição espacial dos agrupamentos para o conjunto de dados 1 (dados resumidos). Os números representam os identificadores (IDs) das estações meteorológicas descritos na Tabela 1.

Tabela 4. Média e desvio padrão das variáveis do conjunto de dados 1 (médias anuais).

	<i>Precipitação anual (mm)</i>		<i>Temp. máxima anual (°C)</i>		<i>Temp. mínima anual (°C)</i>		<i>Radição solar global acumulada (MJ/m².ano)</i>	
	<i>Média</i>	<i>Desvio padrão</i>	<i>Média</i>	<i>Desvio padrão</i>	<i>Média</i>	<i>Desvio padrão</i>	<i>Média</i>	<i>Desvio padrão</i>
Grupo 1	1455,25	120,31	27,57	0,71	15,82	0,80	6765,06	294,85
Grupo 2	1503,95	203,40	29,71	1,10	17,94	0,72	6909,09	152,96
Grupo 3	1713,39	302,38	34,78	1,32	17,14	1,04	7155,49	251,47
Grupo 4	1542,82	284,08	31,94	0,79	15,01	0,67	6907,48	94,80

Os grupos 3 e 4 apresentaram as maiores medianas para a temperatura máxima, sendo acima de 30 °C (Figura 7b). O grupo 4 apresentou menor variância entre as estações meteorológicas. O grupo 2, para

a temperatura mínima do ar (Figura 7c), apresentou a maior mediana seguido pelo grupo 3 que possui a maior variabilidade e desvio padrão (Tabela 4). Porto Velho (RO) e Diamantino (MT), pertencentes a esse grupo, apresentaram como média a temperatura mínima de 18,6 °C e 15,7 °C, respectivamente. As estações goianas pertencentes ao grupo 1 apresentaram as menores temperaturas mínimas. Já o grupo 4 apresentou a menor mediana para a temperatura mínima do ar.

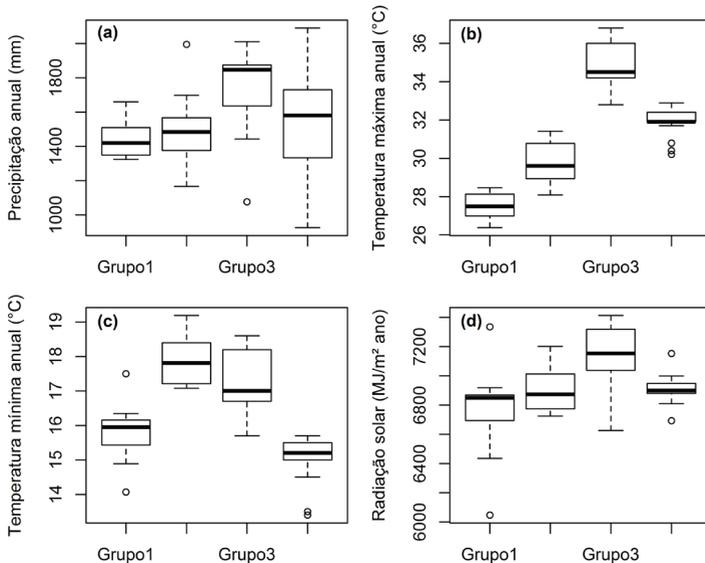


Figura 7. Caracterização climática dos agrupamentos formados pelo conjunto de dados 1 (dados resumidos): (a) precipitação pluvial anual (mm), (b) temperatura máxima anual do ar (°C), (c) temperatura mínima anual do ar (°C) e (d) radiação solar global acumulada por ano (MJ/m².ano).

Para a radiação solar global acumulada, Figura 7(d), o grupo 1 contém duas estações meteorológicas com valores extremos em relação à média do grupo, Caldas Novas e Mineiros, com 6046,59 e 7334,59 MJ/m².ano, respectivamente. O grupo 3 se destaca por apresentar os maiores valores de precipitação pluvial anual, temperatura máxima anual do ar e radiação solar global anual acumulada e o grupo 1, por possuir os menores valores de precipitação pluvial, temperatura máxima do ar e radiação solar. Define-se, então, basicamente duas regiões distintas, uma formada por estações mato-grossenses (grupo 3) e outra por estações goianas (grupo 1).

Para esse conjunto de dados, o grupo 3 apresentou as melhores condições climáticas para o crescimento e desenvolvimento do arroz de terras altas. Esse grupo possui alta precipitação, maior amplitude térmica (diferença entre temperatura máxima e mínima) e maior radiação solar acumulada.

Dados dos meses da safra (conjunto de dados 2)

A Figura 8 ilustra a distribuição espacial dos diferentes grupos pertencentes ao conjunto de dados 2. A Figura 9 apresenta a variabilidade das variáveis climáticas para os seis grupos pertencentes a esse conjunto de dados que apresentou dois agrupamentos (1 e 2) que são compostos por apenas uma estação, Morrinhos e Caldas Novas, ambas localizadas no Estado de Goiás. Devido ao fato de essas duas localidades não serem representativas na produção de arroz de terras altas, esses dois grupos de apenas uma estação não serão comentados neste estudo.

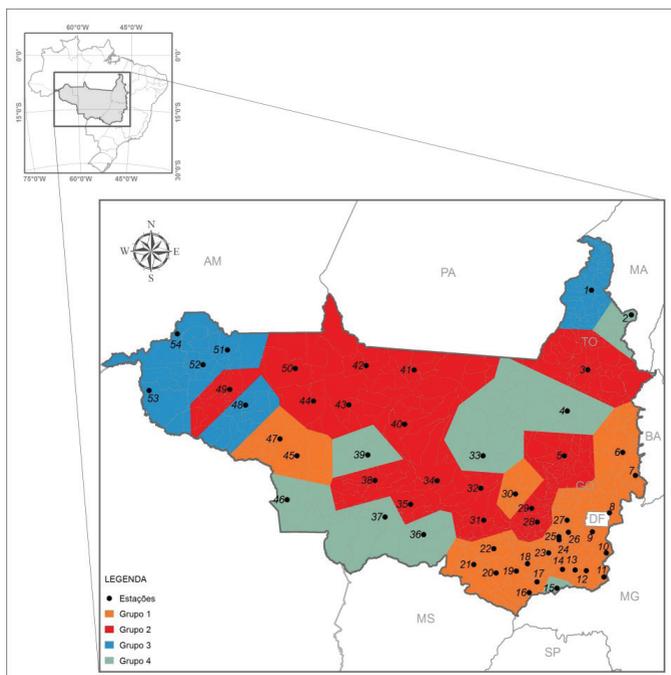


Figura 8. Distribuição espacial dos agrupamentos para o conjunto de dados 2 (meses safra). Os números representam os identificadores (IDs) das estações descritas na Tabela 1.

Para a precipitação pluvial (média dos seis meses da safra) (Figura 9a), o grupo 4 se destacou, apresentando um maior valor para a mediana e média (1456,30 mm) (Tabela 5). Para os outros grupos (1, 2, 3, 5 e 6) existe um equilíbrio para os valores da mediana para a precipitação pluvial (média dos seis meses da safra).

Tabela 5. Médias e variâncias das variáveis do Conjunto de dados 2 (média dos seis meses da safra).

	<i>Precipitação acumulada nos seis meses de safra (mm)</i>		<i>Temp. máxima nos seis meses de safra (°C)</i>		<i>Temp. mínima nos seis meses de safra (°C)</i>		<i>Radiação solar global acumulada (MJ/m².seis meses)</i>	
	<i>Média</i>	<i>Desvio padrão</i>	<i>Média</i>	<i>Desvio padrão</i>	<i>Média</i>	<i>Desvio padrão</i>	<i>Média</i>	<i>Desvio padrão</i>
Grupo 1	1315,30	*	28,57	*	17,98	*	3420,75	*
Grupo 2	1281,00	*	28,85	*	19,33	*	3213,22	*
Grupo 3	1189,14	137,28	28,10	0,82	18,58	0,90	3612,90	74,57
Grupo 4	1456,30	108,15	31,72	0,47	21,20	0,44	3602,24	66,23
Grupo 5	1214,30	126,25	31,62	0,26	21,66	0,49	3452,40	68,79
Grupo 6	1143,40	185,02	31,85	0,40	21,01	0,42	3673,81	50,07

*Os grupos 1 e 2 possuem apenas uma estação meteorológica e por isso não apresentam desvio-padrão.

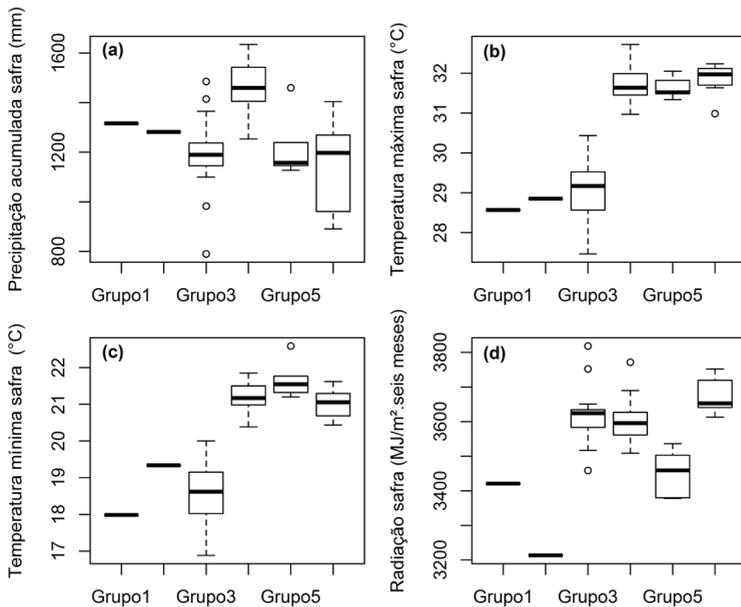


Figura 9. Caracterização climática dos agrupamentos formados pelo conjunto de dados 2 (dados dos meses da safra): (a) precipitação pluvial acumulada por seis meses de safra (mm), (b) temperatura máxima do ar nos seis meses de safra (°C), (c) temperatura mínima do ar nos seis meses de safra (°C), e (d) radiação solar global acumulada por 6 meses (MJ/m².seis meses de safra).

Para a temperatura máxima (Figura 9b), os grupos 4, 5 e 6 se destacaram e apresentaram os maiores valores para as medianas da temperatura máxima em relação aos grupos 1, 2 e 3. Dentre os grupos 3, 4, 5 e 6, o 6 apresentou o menor desvio padrão (0,26 °C, Tabela 5) e o grupo 3 o maior desvio padrão (0,82 °C, Tabela 5). Já, para a temperatura mínima (Figura 9c), os grupos 3 e 5 apresentaram os menores e maiores valores de mediana e média (18,55 °C e 21,66 °C, respectivamente (Tabela 5). Para o grupo 5 a estação localizada em Porto Velho (RO) obteve o maior valor para temperatura mínima (22,58 °C), sendo considerada um “outlier” nesse grupo.

Para a radiação solar acumulada por seis meses (Figura 9d), os grupos 5 e 6 apresentaram as menores e as maiores quantidades da soma da incidência de radiação nos meses da safra, respectivamente. Contudo, ao observar a Tabela 5, verifica-se que a diferença da média da soma da radiação solar entre as estações pertencentes aos grupos 3, 4 e 6 é pequena, de apenas 72MJ/m² seis meses. Para esse conjunto de dados, o grupo 4 apresenta as melhores condições climáticas para o crescimento e desenvolvimento do arroz de terras altas. Esse grupo possui alta precipitação pluvial, amplitude térmica e radiação solar acumulada para seis meses.

Dados dos meses completos (conjunto de dados 3)

As características dos grupos formados pelo conjunto de dados 3 assemelham-se às características dos grupos do conjunto de dados 1 (dados resumidos) devido à pequena diferença entre os agrupamentos. A variável temperatura mínima do ar é a que apresenta a maior diferença entre as variáveis climáticas para os agrupamentos dos dois conjuntos de dados.

A Figura 10 ilustra a variação espacial dos quatro agrupamentos para esse conjunto de dados. A Figura 11 ilustra a variabilidade das variáveis climáticas para os quatro agrupamentos. Para esse conjunto de dados, o grupo 3, composto em sua totalidade pelas estações mato-grossenses localizadas no norte do estado, apresentou o maior valor para a precipitação pluvial mediana e média (Figura 11a e Tabela 6). A

Para a temperatura máxima (Figura 11b), mínima (Figura 11c) e radiação solar acumulada anualmente (Figura 11d), o grupo 3 apresentou os maiores valores para a média (Tabela 6) e mediana. As estações localizadas nos estados de MT e GO foram responsáveis pela diferença entre os agrupamentos 1 e 2, e 3 e 4.

Para esse conjunto de dados, semelhante ao conjunto de dados 1, o grupo 3 apresentou as melhores condições climáticas para o crescimento e desenvolvimento do arroz de terras altas. Esse grupo possui a maior precipitação pluvial e radiação solar, além de possuir alta amplitude térmica.

Comparando os conjuntos de dados 1 e 3, as variáveis precipitação pluvial e radiação solar acumulada possuem variâncias semelhantes (Tabelas 4 e 6, e Figuras 7 e 11).

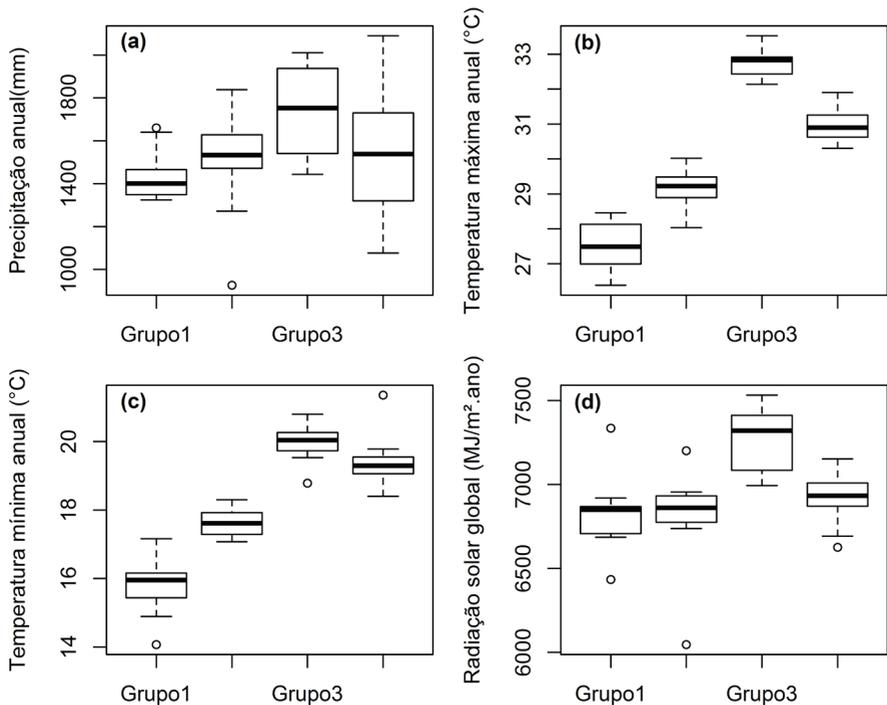


Figura 11. Caracterização climática dos agrupamentos formados pelo conjunto de dados 3 (dados com todos os meses), (a) precipitação pluvial anual (mm), (b) temperatura máxima anual (°C), (c) temperatura mínima anual (°C) e (d) radiação solar global acumulada anualmente (MJ/m².ano).

Conclusões

Apesar dos conjuntos de dados 1 (dados resumidos), 2 (dados dos meses de safra) e 3 (todos os meses do ano) serem oriundos do mesmo banco de dados, os mesmos apresentaram diferenças em função de como as variáveis climáticas foram sumarizadas/observadas.

Os agrupamentos formados pelos conjuntos de dados 1 e 3 apresentaram semelhança moderada, podendo-se assim, substituir o conjunto de dados com todos os meses (3) pelo conjunto de dados resumidos (1).

O conjunto de dados 2 (meses da safra) não apresentou semelhança com os outros conjuntos de dados 1 e 3.

Os agrupamentos formados pelos conjuntos de dados 1, 2 e 3 apresentaram tendência a serem homogêneos em relação a divisão federativa por estados.

Referências

ATLIN, G. N.; BAKER, R. J.; MCRAE, K. B.; LU, X. Selection response in subdivided target regions. **Crop Science**, Madison, v. 40, n. 1, p. 7–13, Jan./Feb. 2000.

FERREIRA, P. L. **Estatística multivariada aplicada**. Coimbra: Universidade de Coimbra, 2000.113 p.

HEGGEN, R. J. (Ed.). **Hydrology handbook**. 2nd ed. New York: American Society of Civil Engineers, 1996. 800 p. (ASCE. Manuals and reports on engineering practice, 28).

HEINEMANN, A. B.; SENTELHAS, P. C. Environmental group identification for upland rice production in central Brazil. **Scientia Agricola**, Piracicaba, v. 68, n. 5, p. 540-547, set./out. 2011.

HUSSON, F.; JOSSE, J.; LE, S.; MAZET, J. **FactoMineR**: multivariate exploratory data analysis and data mining with R: version 1.25.

Disponível em: <<http://CRAN.R-project.org/package=FactoMineR>> .

Acesso em: 5 maio 2013.

HUSSON, F.; LÊ, S.; PAGÈS, J. **Exploratory multivariate analysis by example using R**. Boca Raton: CRC Press, 2011. 228 p.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, Washington, v. 33, n. 1, p. 159-174, Mar. 1977.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte: Ed. UFMG, 2005. 295 p.

NUNES, S. G. **Contribuição da análise de clusters para a identificação de diferentes fenótipos na retinopatia diabética**. 2006. 114 p.

Dissertação (Mestrado em Ciências da Visão) – Faculdade de Medicina, Universidade de Coimbra, Coimbra.

R DEVELOPMENT CORE TEAM. **R**: A language and environment for statistical computing. Vienna, 2012. Disponível em: <<http://www.R-project.org>> . Acesso em: 20 jan. 2013.

REVELLE, W. **Procedures for psychological, psychometric, and personality research**: version 1.2.12. Disponível em: <<http://personality-project.org/r/psych.manual.pdf>> . Acesso em: 20 jan. 2013.