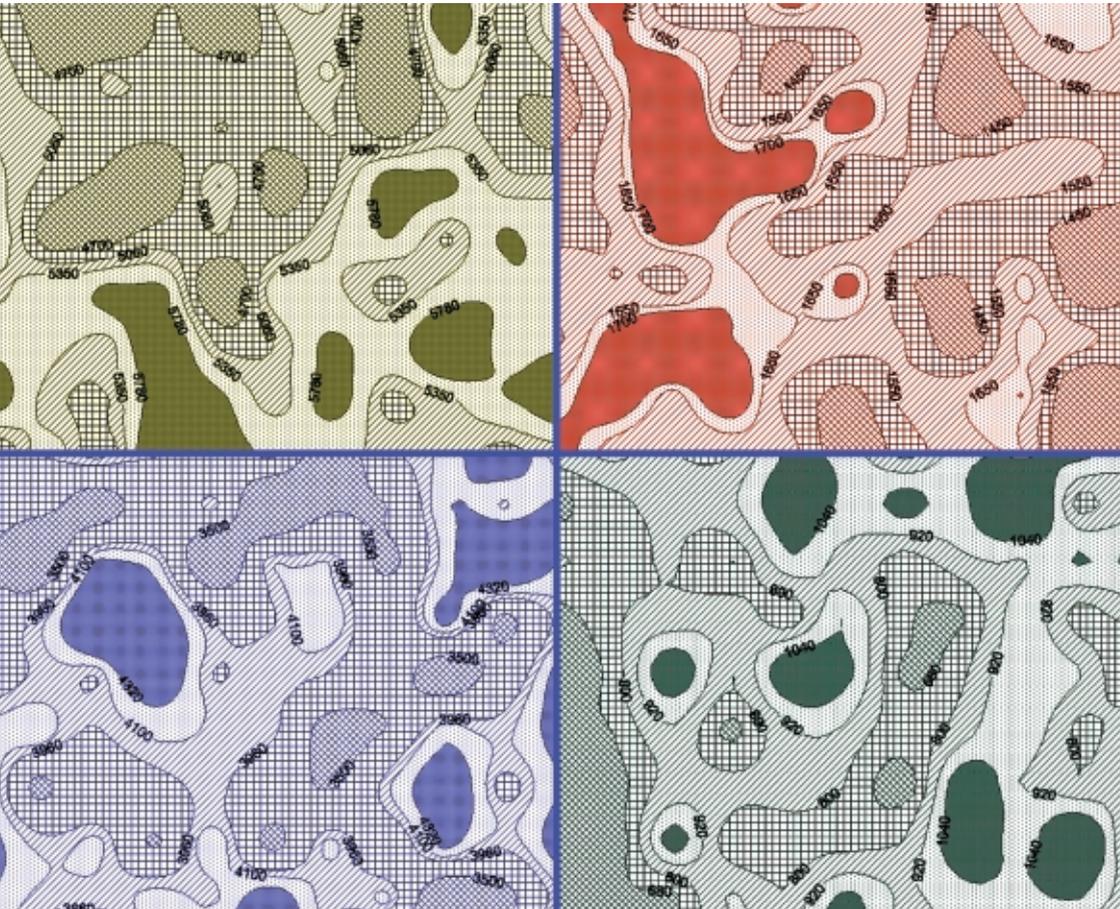


Como Avaliar Similaridades entre Mapas de Produtividade



República Federativa do Brasil

Fernando Henrique Cardoso
Presidente

Ministério da Agricultura, Pecuária e Abastecimento

Marcus Vinicius Pratini de Moraes
Ministro

Empresa Brasileira de Pesquisa Agropecuária - Embrapa

Conselho de Administração

Márcio Fortes de Almeida
Presidente

Alberto Duque Portugal
Vice-Presidente

Dietrich Gerhard Quast
José Honório Accarini
Sérgio Fausto
Urbano Campos Ribeiral
Membros

Diretoria Executiva da Embrapa

Alberto Duque Portugal
Diretor-Presidente

Bonifácio Hideyuki Nakasu
Dante Daniel Giacomelli Scolari
José Roberto Rodrigues Peres
Diretores-Executivos

Embrapa Informática Agropecuária

José Gilberto Jardine
Chefe-Geral

Tércia Zavaglia Torres
Chefe-Adjunto de Administração

Kleber Xavier Sampaio de Souza
Chefe-Adjunto de Pesquisa e Desenvolvimento

Álvaro Seixas Neto
Supervisor da Área de Comunicação e Negócios

Relatório Técnico 10

Como Avaliar Similaridades entre Mapas de Produtividade

José Ruy Porto de Carvalho

Sidney Rosa Vieira

Regina Célia Camargo Pinto Moran

Embrapa Informática Agropecuária
Área de Comunicação e Negócios (ACN)

Av. Dr. André Tosello s/nº
Cidade Universitária "Zeferino Vaz" – Barão Geraldo
Caixa Postal 6041
13083-970 – Campinas, SP
Telefone/Fax: (19) 3789-5743
URL: <http://www.cnptia.embrapa.br>
Email: sac@cnptia.embrapa.br

Comitê de Publicações

Amarindo Fausto Soares
Francisco Xavier Hemerly (Presidente)
Ivanilde Dispato
José Ruy Porto de Carvalho
Marcia Izabel Fugisawa Souza
Suzilei Almeida Carneiro

Suplentes

Fábio Cesar da Silva
João Francisco Gonçalves Antunes
Luciana Alvin Santos Romani
Maria Angélica de Andrade Leite
Moacir Pedrosa Júnior

Supervisor editorial: *Ivanilde Dispato*
Normalização bibliográfica: *Marcia Izabel Fugisawa Souza*
Capa: *Intermídia Publicações Científicas*
Edição eletrônica: *Intermídia Publicações Científicas*

1ª edição

Todos os direitos reservados

Carvalho, José Ruy Porto de.

Como avaliar similaridade entre mapas de produtividade / José Ruy Porto de Carvalho, Sidney Rosa Vieira, Regina Célia Camargo Pinto Moran. — Campinas : Embrapa Informática Agropecuária, 2001.

24 p. : il. — (Relatório técnico / Embrapa Informática Agropecuária ; 10)

ISSN 1517-0330

1. Análise multivariada. 2. Análise de correspondência. 3. Variação temporal. 4. Variação espacial. 5. Geoestatística. I. Vieira, Sidney Rosa. II. Moran, Regina Célia Camargo Pinto. III. Título. IV. Série.

CDD – 519.535 (21.ed.)

Sumário

Resumo	5
Abstract	7
Introdução	9
Material e Métodos	10
Resultados e Discussão	14
Conclusões	22
Referências Bibliográficas	23

Como Avaliar Similaridades entre Mapas de Produtividade

José Ruy Porto de Carvalho¹

Sidney Rosa Vieira²

Regina Célia Camargo Pinto Moran³

Resumo

Um dos problemas emergentes na Agricultura de Precisão é a interpretação de mapas de produtividade. Como avaliar suas similaridades? Seria somente sua comparação visual ou através de técnicas analíticas? O objetivo do presente relatório foi introduzir a Análise Multivariada de Correspondência como um método estatístico que pode auxiliar na interpretação analítica de mapas de produtividade. Com isto pretende-se levantar a hipótese de que a utilização destes mapas para delimitar regiões homogêneas no campo depende da estabilidade espacial e temporal desta variável, tornando-se preocupante seu uso indiscriminado como ferramenta de informação preditiva. Quatro mapas de produtividade foram

¹ P.hD. em Estatística, Pesquisador da Embrapa Informática Agropecuária, Caixa Postal 6041, Barão Geraldo – 13083-970 – Campinas, SP. (email: jrui@cnptia.embrapa.br)

² P.hD. em Solos, Pesquisador do Instituto Agronômico de Campinas, Centro de Solos e Recursos Agroambientais, Caixa Postal 28 – 13001-970 – Campinas, SP. (email: sidney@cec.iac.br)

³ P.hD. em Estatística, Professora do Instituto de Matemática, Estatística e Computação Científica da Unicamp, Caixa Postal 6065, Barão Geraldo – 13081-970 – Campinas, SP. (email: regina@ime.unicamp.br)

usados neste trabalho, os quais correspondem a amostras coletadas de componentes de produção para arroz em 1994, milho em 1995, algodão em 1996 e milho em 1998, em uma malha com 110 pontos, espaçados de 10 x 10 m, da estação experimental do Instituto Agronômico de Campinas em Votuporanga, SP. Os dados originais de produtividade foram categorizados e analisados usando a Análise de Correspondência Simples e Múltipla. Conclui-se que existe uma grande dissimilaridade entre as culturas ou as classes de culturas, medidas através da distância euclidiana no plano.

Termos para indexação: Interpretação de mapas de produtividade; Variabilidades espacial e temporal; Análise de correspondência simples e múltipla.

How to Evaluate Similarities Among Yield Mapping

Abstract

One of the emerging problems in the Precision Agriculture is the interpretation of yield maps. How to evaluate their similarities? It would be only its visual comparison or through analytical techniques? The objective of the present report was to introduce the Multivariate analysis of Correspondence as a statistical method that can aid in the analytical interpretation of yield maps. The hypothesis that the use of yield mapping is strictly related to the variability of the components was studied. Four yield maps were used corresponding to collected yield samples for rice in 1994, corn in 1995, cotton in 1996 and corn in 1998, in a grid of 110 points, spaced of 10 x 10 m, at the experimental station of the Instituto Agronômico, located at Votuporanga, SP, Brazil. The yield original data were categorised and analysed using Simple and Multiple Correspondence Analysis. A great dissimilarity between the cultures or classes of cultures was found as measured by the Euclidean Distance in the plan.

Index terms: Yield mapping interpretation; Spatial and temporal variability; Simple and multiple correspondence analysis.

Introdução

As tecnologias e práticas emergentes da Agricultura de Precisão desafiam o uso da média (Capelli, 2000) como valor representativo de uma situação. O manejo da variabilidade é o principal foco de sua atenção. Duas formas principais de variabilidade são imediatamente detectadas: espacial e temporal. A variação espacial é a variação vista no campo, afetada pelos diferentes fatores que o compõem. A variação temporal é a variação que ocorre quando se comparam mapas de produtividade de um ano para outro.

O uso de Global Positioning Systems - GPS combinado com dados de produção agrícola permite elaborar mapas de produtividade. Estes mapas constituem uma ferramenta de diagnóstico muito importante para a agricultura de precisão, pois promovem o uso mais eficiente das informações, possibilitando desta maneira a administração da variabilidade no campo (Emmott et al., 1997).

A comparação de mapas de colheita para diferentes anos é uma etapa muito importante no processo de identificação de zonas de manejo. Porém, o que determina similaridade ou dissimilaridade entre esses mapas permanece subjetivo, enquanto não se tiver um critério para tal. Seriam algumas poucas áreas que aparecem moldadas semelhantemente no mapa? Seriam fatores como seleção de cor, número de classes e pontos de ruptura temáticos? Seriam a presença, ausência ou controle da variabilidade fatores de influência em similaridade?

O objetivo deste relatório é apresentar a análise de correspondência como um método estatístico multivariado que pode ajudar na interpretação de mapas de produtividade. Com isto pretende-se estudar a hipótese de que a utilização de mapas de produtividade para identificação de zonas de manejo é dependente da variabilidade de rendimento no ano e entre anos, tornando-se, desta maneira, preocupante seu uso indiscriminado sem se considerar as variações espaciais e temporais inerentes.

Material e Métodos

Na área experimental do Instituto Agronômico de Campinas - IAC, localizada no Núcleo de Agronomia da Noroeste em Votuporanga – São Paulo, em latossolo vermelho amarelo arenoso, foram coletadas amostras de componentes de produtividade para arroz em 1994 – A4, milho em 1995 – M5, algodão em 1996 – A6 e milho em 1998 – M8, em uma malha com 110 pontos, espaçados de 10 x 10 m, conforme pode ser visto na Fig. 1. Em cada ponto, foram coletados as partes aéreas de 5 m², onde também foi medida a produtividade.

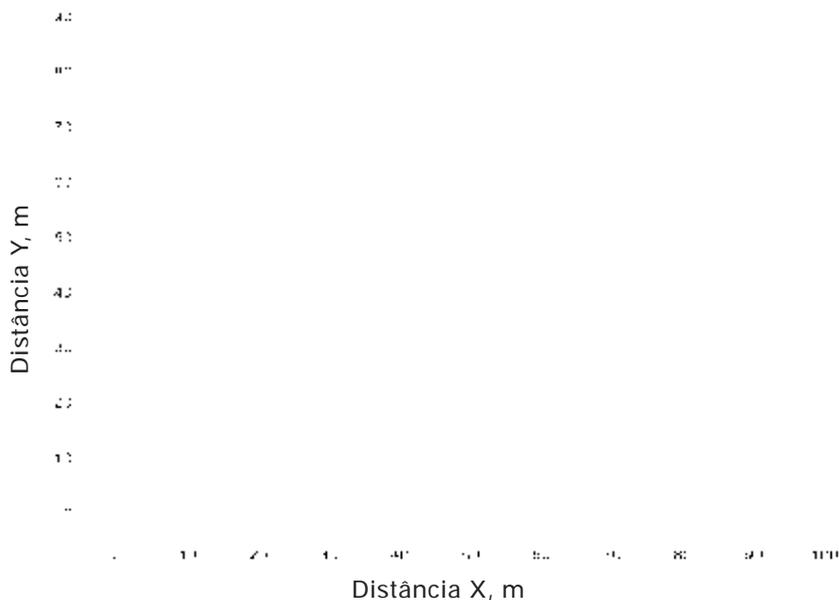


Fig. 1. Esquema de amostragem.

Os mapas de produtividade ou mapa de isolinhas para as culturas nos quatro anos, foram obtidos pela interpolação dos valores em qualquer posição no campo de estudo, sem tendência e com variância mínima, através do método de interpolação chamado krigagem. Segundo Vieira et al. (1983), a precisão da localização das isolinhas entre dois pontos é extremamente dependente da densidade de pontos por área e, conseqüentemente, da distância entre os pontos.

O método estatístico usualmente utilizado para avaliar diferenças de médias de duas amostras de dados é o teste t de Student. Entretanto, as condições para que este teste seja apropriadamente aplicado são de que as amostras devem ser independentes e normalmente distribuídas. O problema é que estas condições raramente se aplicam em mapas de produtividade devido às dependências espacial e temporal das observações.

Uma outra forma de efetuar a comparação de mapas é através do índice KAPPA (K) de concordância. Esta medida estatística foi introduzida por psicólogos e adaptada como medida de concordância para mapas por Congalton & Mead (1983). Este índice testa a associação entre mapas. Ele nos ajuda a entender se os mapas diferem devido a alguma variação causal ou se há uma real concordância. Ele também permite a comparação de dados com estruturas definidas (Rosenfield & Fitzpatrick-Lins, 1986). O índice KAPPA varia de 0 a 1 onde 0 indica que os resultados acontecem totalmente ao acaso e 1 indica concordância perfeita e é definido por:

$$K = \left\{ n \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i}) \right\} / \left\{ n^2 - \sum_{i=1}^r (x_{i+} * x_{+i}) \right\}$$

onde r é o número de linhas em uma tabela de classificação cruzada; x_{ii} é o número de combinações na diagonal; x_{i+} é o total de observações na linha i; x_{+i} o total de observações na coluna i e n o número total de observações.

Em resumo, existem diversas maneiras de se tratar o problema em questão. Há também os métodos de classificação cruzada, coeficiente de similaridade de Qui-Quadrado e Cramer (Lorup, 2000), entre outros. Entretanto, a maioria deles depende de condições pré-estabelecidas que normalmente não se verificam para dados temporais e espacialmente dependentes.

A Análise de Correspondência é uma técnica multivariada para análise exploratória de dados categorizados. A preocupação com a análise destes tipos de dados já vem desde o começo do século. Mas somente na década de 60, por intermédio de Jean-Paul Benzécri que definiu um método mostrando suas propriedades algébricas e geométricas denominado "Analyse Factorielle des Correspondences", esta técnica se difundiu (Benzécri, 1992). Ela converte uma matriz de dados não negativos em um

particular tipo de gráfico que exhibe as linhas e colunas da matriz como pontos de um espaço vetorial de dimensão menor que a original, de maneira que as relações entre as linhas, entre as colunas e entre linhas e colunas possam ser interpretadas. É exatamente por estar interessada em estudar a correspondência entre variáveis, que esta técnica recebeu o nome de Análise de Correspondência. Sua geometria e álgebra fazem com que pertença a uma família de técnicas de disposição gráfica que são baseadas em aproximação de uma matriz por outra de posto menor, por meio da decomposição em valores singulares. Ou seja, o objetivo desta análise é achar um subespaço que melhor se ajuste ao conjunto (nuvem) de pontos no espaço euclidiano. Este ajuste é feito pelo método de quadrado mínimo ponderado onde a distância euclidiana generalizada (ponderada) é utilizada em um sistema de massas pontuais (Greenacre & Hastie, 1987). Um dos pontos relevantes desta técnica é que não há exigência de normalidade para a resposta estudada. Em consequência, os testes estatísticos inferenciais não são aqui utilizados, estando a solução sugerida pela distribuição gráfica de seus resultados, o que se constitui na base da escola francesa para as análises multivariadas (Benzécri, 1992).

A Análise de Correspondência Simples parte de uma matriz de dados representado por uma tabela de contingência. O desenvolvimento do algoritmo e sua geometria fornecem as regras básicas para a interpretação (Greenacre, 1984; Barioni Junior, 1995).

Na interpretação de mapas de produtividade, além do interesse na similaridade dos mapas, é interessante o estudo das similaridades de classes ou áreas dentro dos mapas. Essas diferenças entre classes ou áreas são observáveis através da investigação das associações existentes entre as categorias dentro de cada classe e/ou entre as categorias de diferentes classes. Ela segue o mesmo objetivo da Análise Discriminante Canônica de Fisher, ou seja, o de maximizar a variabilidade entre classes, representado na Análise de Correspondência pela inércia entre classes.

O tratamento matemático para a execução da Análise de Correspondência em classes segue o mesmo princípio para a Análise de Correspondência Múltipla. Enquanto na Análise de Correspondência Simples a tabela de contingência é usada, na Análise de Correspondência Múltipla usa-se a tabela de incidência ou tabela Burt. Na Análise de Correspondência Múltipla, todas as associações entre pares de variáveis são analisadas, bem como a associação entre uma variável e ela mesma. A inclu-

são das associações de uma variável com ela mesma, o que não acontece na Análise de Correspondência Simples, é que permite a Análise de Correspondência Múltipla ser eficiente em sua parte computacional.

Em muitas técnicas multivariadas, há sempre a escolha de basear as análises na matriz de dados originais ou em alguma forma transformada destes dados. Na maioria das vezes é mais apropriado escolher a forma transformada, para que valores de cada variável não dependam de possíveis unidades arbitrárias de medida.

Para tornar as variáveis de produtividade comparáveis, as 110 amostras de cada uma das quatro variáveis - A4, M5, A6 e M8 foram classificadas em cinco grupos de aproximadamente igual tamanho, de conformidade com seus valores de produtividade e foram atribuídas a elas números correspondentes às suas posições na classificação conjunta do grupo. Os grupos foram definidos através do módulo Proc Rank do software SAS Institute (1998). Assim, para a variável A4 por exemplo, a menor produtividade pertencerá ao grupo zero e a maior ao grupo quatro, respeitando suas posições originais.

Para estudar se existe similaridade entre classes de cada variável, a matriz de dados antes composta com 04 variáveis (A4, M5, A6, M8), foi transformada em 20 variáveis, onde para cada variável produtividade, 05 variáveis foram criadas, correspondentes aos grupos 0,1,2,3 e 4. Assim sendo, as novas variáveis foram definidas como se segue:

A40, A41, A42, A43, A44 \Rightarrow Produtividade de arroz de 1994 nas classes 0,1,2,3,4;

M50, M51, M52, M53, M54 \Rightarrow Produtividade de milho de 1995 nas classes 0,1,2,3,4;

A60, A61, A62, A63, A64 \Rightarrow Produtividade de algodão de 1996 nas classes 0,1,2,3,4;

M80, M81, M82, M83, M84 \Rightarrow Produtividade de milho de 1998 nas classes 0,1,2,3,4.

A matriz de dados para aplicação da Análise de Correspondência tem a dimensão 20 colunas por 110 linhas e é composta de 0 e 1's, denominada matriz de incidência, indicando presença ou ausência de cada nova variável criada para cada amostra coletada.

Resultados e Discussão

As Fig. 2, 3, 4 e 5 mostram os mapas de produtividades para as culturas de arroz, milho e algodão para os anos de 1994, 1995, 1996 e 1998, respectivamente. Os mesmos grupos anteriormente obtidos foram usados para subdividir os dados de produtividade. Verifica-se que, apesar de serem medidas na mesma área experimental e de se tratarem de culturas de anos diferentes, existem grandes diferenças numéricas entre os diversos grupos para os quatro mapas, sugerindo que as variabilidades espacial e temporal apresentam influência nos mapas de produtividade.



Fig. 2. Mapa de produtividade de arroz (A4) - Votuporanga 1994.

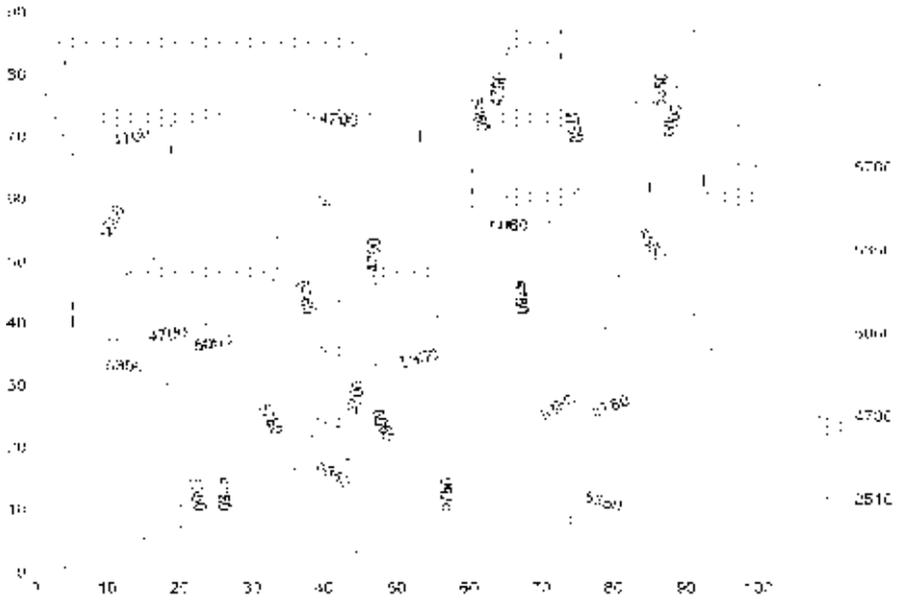


Fig. 3. Mapa de produtividade de milho (M5) – Votuporanga 1995.

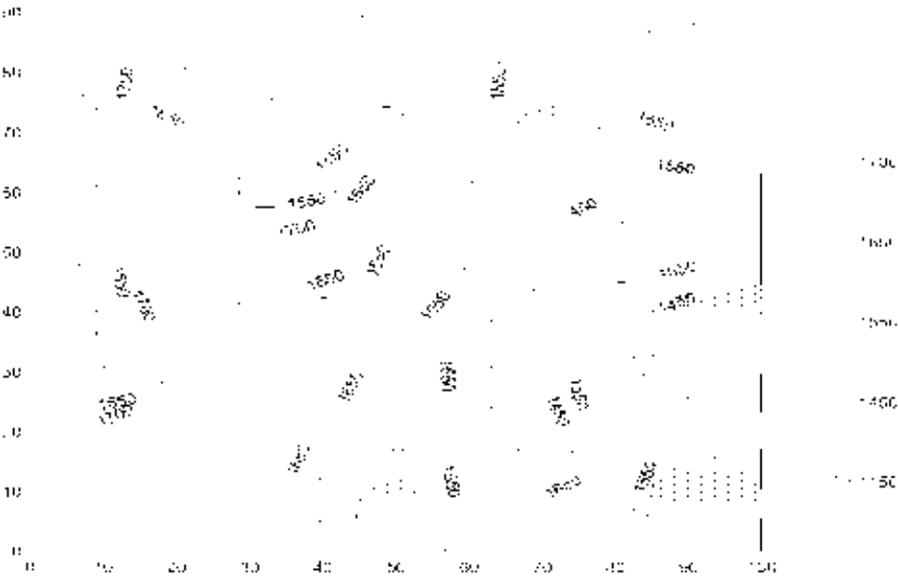


Fig. 4. Mapa de produtividade de algodão (A6) – Votuporanga 1996.

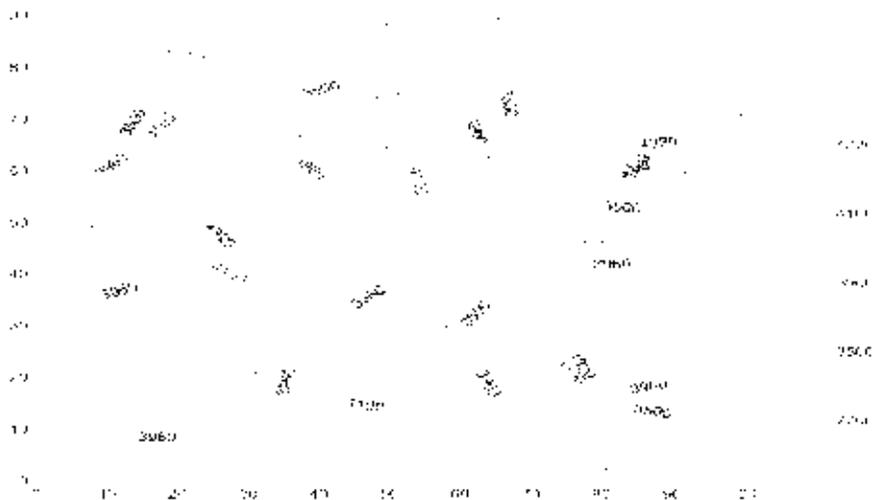


Fig. 5. Mapa de produtividade de milho (M8) – Votuporanga 1998.

Para verificar se há similaridade entre os mapas de produtividade, utilizou-se a Análise de Correspondência Simples através do módulo Proc Corresp do software SAS Institute (1998) cujos resultados estão resumidos nas Tabelas 1 e 2.

Ao analisar um conjunto de dados utilizando a Análise de Correspondência como instrumento, faz-se uso de gráficos para visualizar as relações entre as variáveis que se deseja analisar. Ao representar um conjunto de pontos, não interessa ver as freqüências absolutas de cada linha, mas sim o perfil de uma linha ao longo das categorias das colunas. Ao introduzir os perfis, padronizam-se os pontos, reduzindo assim a informação à proporção de indivíduos em cada linha. A massa é atribuída a cada ponto sendo a freqüência relativa de cada linha na matriz de dados.

Para obtenção dos eixos a serem analisados, observam-se os percentuais acumulados das inércias principais e definem-se, baseados na porcentagem explicada pela inércia, quantos eixos forem necessários para a representação dos pontos.

Tabela 1. Decomposição da Inércia e do Qui-Quadrado. Análise de Correspondência Simples para as variáveis de produtividade.

Valores Singulares	Inércia Principal	Qui-Quadrado	%	7	14	21	28	35
				-----+	-----+	-----+	-----+	-----+
0,378	0,143	127,294	36,92	*****				
0,371	0,138	122,725	35,59	*****				
0,326	0,106	94,780	27,49	*****				
Total	0,388	344,799		Graus de Liberdade = 327				

Como pode ser visto na Tabela 1, o total da estatística de Qui-Quadrado é 344,79, a qual é uma medida de associação entre as linhas e colunas nas três dimensões da tabela de dados transformada obtida pela decomposição dos valores singulares. Cerca de 37% da estatística de Qui-Quadrado e Inércia são explicadas pelo primeiro eixo vetorial, enquanto que cerca de 36% são explicadas pelo segundo eixo vetorial. No trabalho em estudo verifica-se que os pontos serão representados em dois eixos, indicando que a associação entre linhas e colunas é bidimensional, com 72% da inércia explicada. Como a menor dimensão da Tabela1 é a terceira, não há perda de informação quando somente duas dimensões são consideradas.

Para a interpretação dos eixos, verifica-se como a inércia dos pontos se decompõe em cada eixo. São os componentes da inércia que formam os coeficientes que indicam o quanto cada ponto contribuiu para a determinação da direção dos eixos e o quanto cada ponto está representado em cada eixo indicados pelos coeficientes C1 e C2, como pode ser visto na Tabela 2.

Tabela 2. Resumo de estatísticas para as variáveis de produtividade. Análise de Correspondência Simples.

Variáveis	Eixo1	Eixo2	C1	C2	Qr1	Qr2
A4	1,565	0,656	0,602	0,105	0,815	0,138
M5	-0,999	1,281	0,246	0,405	0,346	0,548
A6	0,201	-1,216	0,010	0,375	0,016	0,565
M8	-0,745	-0,666	0,140	0,112	0,228	0,176

Eixo1 e Eixo2 referem-se respectivamente à decomposição da inércia para o primeiro e segundo eixos.
 C1 e C2 referem-se respectivamente aos pontos que melhor explicam a inércia para o primeiro e segundo eixos.
 Qr1 e Qr2 referem-se respectivamente à qualidade de representação dos pontos para o primeiro e segundo eixos.

Na mesma Tabela 2, a qualidade de sua representação, indicada pelas coeficientes Qr1 e Qr2, também pode ser vista. A qualidade da representação é o coeficiente que nos dá o quanto cada ponto está ou não bem representado no subespaço considerado. Os resultados das Tabelas 1 e 2 possibilitam a geração e a interpretação do gráfico bidimensional – Produtividade para as culturas, conforme pode ser visto na Fig. 6, definido pelas coordenadas dos dois primeiros eixos vetoriais.

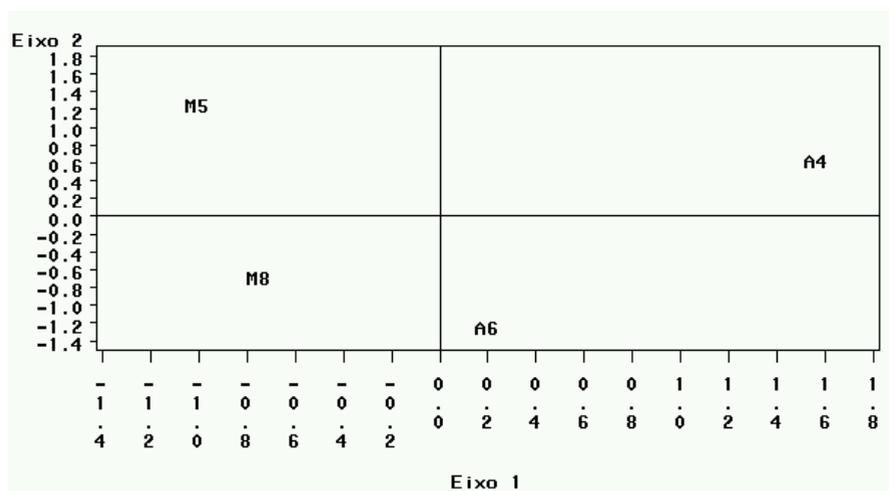


Fig. 6. Produtividade para as culturas - Produtividade de arroz 94 (A4)/Produtividade de milho 95 (M5)/Produtividade de algodão 96 (A6)/Produtividade de milho 98 (M8).

Cada ponto representa o perfil de produtividade de cada uma das culturas ao longo das 110 amostras. Assim sendo, M5 representa o perfil da produtividade do milho em 95 ao longo das 110 amostras, A4 o perfil de arroz para 94, M8 o perfil de milho para 98 e A6 o perfil de algodão para 1996, respectivamente. Como a Análise de Correspondência aloca os pontos em um espaço euclidiano, as projeções dos mesmos sobre os eixos permitem tirar conclusões sobre suas similaridades ou dissimilaridades.

Pela Fig.6 e Tabela 2, verifica-se que a variável A4 foi a que mais contribuiu para o primeiro eixo vetorial com 60% da inércia explicada. A variável M8 contribuiu com 14% para este mesmo eixo. A distância de suas projeções sobre os eixos mostra uma grande dissimilaridade entre elas. O primeiro eixo vetorial mostra um claro contraste entre as culturas de arroz e de algodão com a de milho. No segundo eixo vetorial, a variável M5 contribuiu com 41% e a A6 com 38% para a inércia explicada. Suas distâncias apresentam-se bastante dissimilares. Neste eixo, estão relacionados os anos de 1994 e 1995 com 1996 e 1998, estando evidenciado um contraste entre os primeiros anos e os últimos. Verifica-se que os mapas de produtividade para arroz em 94 e milho em 98 são completamente distintos. O mesmo acontece para milho 95 e algodão 96, ou seja, as variabilidades espacial e temporal influenciaram a variável produtividade dentro e entre anos. Assim conclui-se que não são somente as exigências nutricionais das diferentes culturas que podem alterar os mapas de produtividade, as variações espaciais e temporais são também causas de sua alteração, tornando-se um motivo maior de preocupação o seu uso indiscriminado.

A Análise de Correspondência simples numa tabela de incidência é semelhante à Análise de Correspondência Múltipla. Como era de se esperar, os valores das contribuições e qualidade de representação vistos na Tabela 3, são bem inferiores ao da Análise de Correspondência Simples. Como o plano de amostragem continuou com 110 pontos mas o número de variáveis aumentou de quatro para vinte, a capacidade de justificar a inércia para o mesmo número de eixos é menor.

Tabela 3 apresenta as coordenadas das variáveis para os dois eixos e suas contribuições e qualidade de representação. Verifica-se que as variáveis A40, M50, M51, M53, M54, A63, A64, M81, M83 e M84 apresentam maior contribuição para o primeiro eixo e as variáveis A41, A42, A43, A44, M52, A60, A61, A62, M80 e M82 para o segundo eixo.

Tabela 3. Resumo das estatísticas para as variáveis de produtividade divididas por classes. Análise de Correspondência Múltipla.

Classes	Eixo1	Eixo2	C1	C2	Qr1	Qr2
A40	0,855	-0,413	0,034	0,008	0,067	0,015
A41	-0,581	1,395	0,016	0,092	0,030	0,173
A42	-0,913	-1,140	0,055	0,085	0,115	0,176
A43	0,051	-0,662	0,001	0,016	0,001	0,030
A44	0,902	1,076	0,040	0,057	0,078	0,109
M50	-1,812	0,203	0,164	0,002	0,318	0,003
M51	1,382	-0,591	0,095	0,017	0,185	0,033
M52	0,779	1,128	0,031	0,066	0,062	0,127
M53	0,393	-0,269	0,007	0,003	0,014	0,006
M54	-0,759	-0,535	0,028	0,014	0,055	0,027
A60	-0,980	1,553	0,034	0,087	0,063	0,155
A61	0,504	-1,616	0,015	0,160	0,032	0,321
A62	0,099	0,594	0,001	0,023	0,001	0,047
A63	2,437	0,666	0,162	0,012	0,282	0,020
A64	-1,156	-0,247	0,079	0,003	0,160	0,007
M80	-0,529	2,451	0,012	0,273	0,024	0,505
M81	-1,364	-0,956	0,101	0,049	0,201	0,096
M82	-0,211	-0,487	0,002	0,010	0,003	0,019
M83	1,246	-0,375	0,081	0,007	0,159	0,014
M84	0,821	-0,335	0,035	0,005	0,069	0,011

Eixo1 e Eixo2 referem-se respectivamente à decomposição da inércia para o primeiro e segundo eixos.

C1 e C2 referem-se respectivamente aos pontos que melhor explicam a inércia para o primeiro e segundo eixos.

Qr1 e Qr2 referem-se respectivamente à qualidade de representação dos pontos para o primeiro e segundo eixos.

Estes resultados possibilitam a geração e a interpretação de um gráfico bidimensional – Produtividade por Classes, conforme pode ser visto na Fig. 7, definido pelas coordenadas nos dois eixos ortogonais.

O ponto A60 plotado na Fig. 7, por exemplo, é a melhor representação da distribuição condicionada no plano, a ter a menor produtividade na variável A6. Portanto, refere-se somente às localizações com alguma dimensão na classe zero. O vetor A60, como perfil linha, descreve a distribuição condicional das classes de produtividade ao longo do tempo por cultura. Seguem-se conclusões similares para os demais pontos.

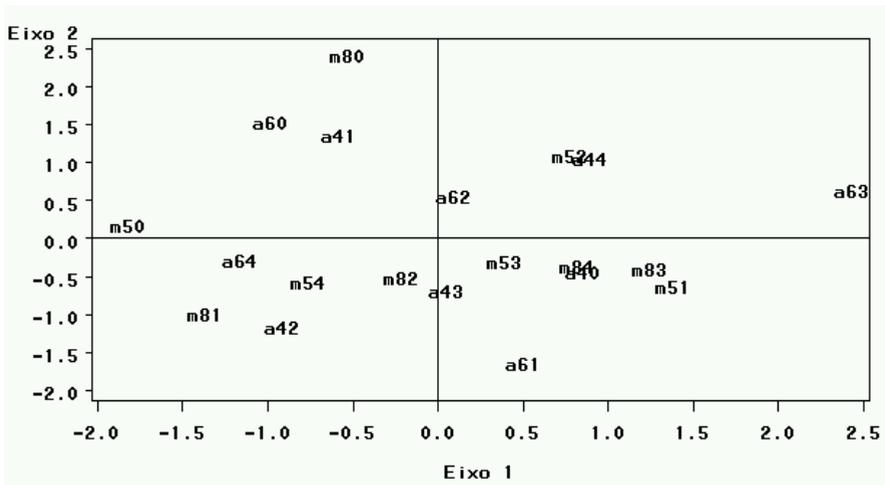


Fig.7. Produtividade por classes - Produtividade de arroz 94 (A40-A44)/ Produtividade de milho 95 (M50-M54)/Produtividade de algodão 96 (A60-A64)/Produtividade de milho 98 (M80-M84).

Interpretando a Fig. 7 com a ajuda da Tabela 3, verifica-se que a distância das projeções das classes M50, M81, A64 e M54 no primeiro eixo são bastante próximas, indicando que as mesmas podem ser consideradas similares. O mesmo acontece para as classes M53, M84, A40, M83, M51 e com menos rigor para a A63. Entretanto, os dois grupos são bastante dissimilares, pois suas projeções encontram-se em quadrantes opostos em relação ao eixo. Em relação ao segundo eixo, as classes M80, A60, A41, M52, A44 e A62, apresentam pouca distância quando de sua projeção sobre o eixo, podendo ser consideradas similares. O mesmo acontece, em relação ao segundo eixo para as classes A61, A42, A43 e M82. Entretanto, estes dois grupos são dissimilares. Ou seja, para cada eixo, existem grupos compostos por variáveis que são similares dentro dos grupos e dissimilares entre grupos.

Para o primeiro eixo, as classes 0 e 4 para a produtividade de milho 95 (M50 e M54) são similares entre si e similares com a classe 1 da produtividade de milho 98 (M81) e classe 4 de algodão 96 (A64). O segundo grupo neste eixo mostra similaridades entre as classes 3 e 4 para a produtividade de milho 98 (M83 e M84) e entre as classes 1 e 3 para produtivi-

de de milho 95 (M51 e M53) e também com as classes 0 de arroz 94 (A40) e 3 de algodão 96 (A63). Entretanto, ambos os grupos são bastante dissimilares. Para o segundo eixo existe uma clara similaridade para as classes 1 e 4 da produtividade de arroz 94 (A41 e A44) e classes 0 e 2 de algodão 96 (A60 e A62) e elas são similares com a classe 0 da produtividade de milho 98 (M80) e classe 2 de milho 95 (M52). Para o segundo grupo, as classes 2 e 3 da produtividade de arroz 94 (A42 e A43) são similares entre si e com as classes 1 de algodão 96 (A61) e 2 de milho 98 (M82). Comparando os eixos, existe uma clara evidência de que, apesar de serem dissimilares, as produtividades de milho 95 (M5) e 98 (M8) têm a maioria de suas classes contribuindo para o primeiro eixo, enquanto que as classes de produtividade de arroz 94 e de algodão 96 contribuem em maior número para o segundo eixo. Com isto, verifica-se que as variações espacial e temporal apesar de influenciarem a produtividade de milho para 95 e 98, têm maior influência na produtividade de arroz 94 e algodão 96.

Conclusões

Com base nos resultados apresentados, conclui-se:

1. As exigências nutricionais das diferentes culturas alteraram os mapas de produtividade. Entretanto, as variações espaciais e temporais são também causas de sua alteração. Quando as variáveis, produtividade de arroz 94, milho 95, algodão 96 e milho 98 foram comparadas através da Análise de Correspondência Simples, à distância de suas projeções sobre os eixos mostrou-se bastante dissimilar, indicando que os mapas de produtividade sofrem influência de variações espaciais e temporais.
2. Quando as culturas foram categorizadas em grupos de produtividade, encontrou-se classe de culturas similar dentro de grupos. Entretanto existem dissimilaridades entre grupos, mostrando que as variações espaciais e temporais influenciam os mapas de produtividade, tanto no conjunto como quando são subdivididos em classes.
3. Devido à influência das variações espaciais e temporais, torna-se temeroso o uso indiscriminado de mapas de produtividade sem que haja um modelo de produtividade definido. Este modelo seria construído ao se produzir os mapas de produtividade em uma mesma área por um número de anos e as possíveis variações caracterizadas.

Referências Bibliográficas

BARIONI JÚNIOR, W. **Análise de correspondência na identificação dos fatores de risco associados à diarreia e à performance de leitões na fase de lactação**. 1995. 97 f. Tese (Mestrado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba.

BENZÉCRI, J.-P. **Correspondence analysis handbook**. New York: Marcel Decker, 1992. 665 p. (Statistics: Textbooks and Monographs, 125).

CAPELLI, N. L. **Agricultura de precisão – novas tecnologias para o processo produtivo**. Disponível em: <<http://www.cria.org.br/gip/gipap/capelli.htm>>. Acesso em: 28 jul. 2000.

CONGALTON, R. G.; MEAD, R. A. A. Quantitative method to test for consistency and correctness in photointerpretation. **Photogrammetric Engineering and Sensing**, v. 49, p. 69-74, 1983.

EMMOTT, A; HALL, J.; MATTHEWS, R. **Precision farming applied to plantation agriculture**. Disponível na Internet: <<http://www.silsoe.cranfield.ac.uk/cpf>>. Acesso em: 10 fev. 2000.

GREENACRE, M. J. **Theory and applications of correspondence analysis**. London: Academic Press, 1984. 364 p.

GREENACRE, M. J.; HASTIE, T. The geometric interpretation of correspondence analysis. **Journal of the American Statistical Association**, v. 82, n. 398, p. 437-447, 1987.

LORUP, E. J. **IDRISI WWW tutorial**: map pairs analysis tools II. Disponível em: <<http://www.sbg.ac.at/geo/idrisi/wwwtutor/mptools2.htm>>. Acesso em: 29 ago. 2000.

ROSENFELD, G. H.; FITZPATRICK-LINS, K. A coefficient of agreement as a measure of thematic classification accuracy. **Photogrammetric Engineering and Remote Sensing**, v. 52, n. 2, p. 223-227, 1986.

SAS INSTITUTE. **SAS/STAT users guide**: version 6. 4.ed. Cary, 1998. v.1, 943 p.

VIEIRA, S. R.; HATFIELD, J. L.; NIELSEN, D. R.; BIGGAR, J. W. Geostatistical theory and application to variability of some agronomical properties. **Hilgardia**, Berkeley, v.51, n. 3, p. 1-75, 1983.

Embrapa

Informática Agropecuária

**MINISTÉRIO DA AGRICULTURA,
PECUÁRIA E ABASTECIMENTO**

**GOVERNO
FEDERAL**
Trabalhando em todo o Brasil