

# Considerações sobre a identificação dos fatores de formação do preço da soja no estado de Mato Grosso

*Bryan Mariano Martinez<sup>1</sup>*

*Mariano Martinez Espinoza<sup>2</sup>*

*Mamadu Lamarana Bari<sup>3</sup>*

*Arturo Alejandro Zavala Zavala<sup>4</sup>*

*João de Sousa<sup>5</sup>*

**Resumo:** O objetivo do presente trabalho foi identificar fatores que mais influenciaram na formação do preço da soja no estado de Mato Grosso em 2011. Considerando que a soja é uma das culturas de grande relevância para a economia do estado e, de maneira geral, para o Brasil, devido ao país ocupar lugar de destaque no cenário mundial e do complexo de soja representar aproximadamente um terço do agronegócio brasileiro, é importante analisar quais são os principais fatores que mais influenciam na formação do preço da soja. Para este estudo, foram consideradas as técnicas de análise multivariada, especificamente Análise Fatorial e Análise de Agrupamento. Os resultados do estudo mostram que as variáveis que mais contribuem com a formação do custo da soja, das variáveis consideradas,

---

1 Mestre em Agronegócios e Desenvolvimento Regional da Faculdade de Economia da Universidade Federal de Mato Grosso (FE/UFMT).

2 Mestre em Agronegócios e Desenvolvimento Regional da Faculdade de Economia da Universidade Federal de Mato Grosso (FE/UFMT).

3 Professor Adjunto da Faculdade de Economia da Faculdade de Economia da Universidade Federal de Mato Grosso (FE/UFMT).

4 Professor Adjunto da Faculdade de Economia da Faculdade de Economia da Universidade Federal de Mato Grosso (FE/UFMT).

5 Professor Adjunto do Instituto Exatas da Terra da Universidade Federal de Mato Grosso (UFMT).

foram a soja Chicago Board of Trade (CBOT), operações agrícolas, o preço do dólar e o preço do diesel.

**Palavras-Chave:** Preço da soja. Análise fatorial. Análise de agrupamentos.

**Abstract:** The aim of this study was to identify factors that more influenced in the formation of the price of soybeans, in Mato Grosso State, Brazil, in 2011. Considering that the soybeans is one of the cultures of great relevance for the state economy in general and to Brazil, due to Brazil occupy a prominent place on the world stage, and soybean complex to represent approximately one third of the Brazilian agribusiness, therefore it is important to analyze what are the main factors that influence the formation of the price of soybeans. This study, we considered the multivariate analysis, specifically factor analysis and cluster analysis. The study results show that the variables that contribute most to the formation of the cost of soybeans, of the considered variables was the CBOT soybeans, agricultural operations, the price of the dollar and the price of diesel.

**Keywords:** Soybean prices. Factor analysis. Cluster analysis.

## I Introdução

No Brasil, os produtos do agronegócio que mais têm se destacado pelo volume exportado e pela geração de divisas são os do complexo soja. Esses produtos são constituídos pela soja em grão, pelo farelo e óleo de soja. (IMEA, 2011)

O complexo constitui uma fonte geradora de divisas no mercado internacional e é responsável por resultados positivos na balança comercial brasileira. Ele tem passado por um processo de transformação tecnológica, estrutural e competitiva que tem revolucionado as atividades no campo no Brasil de 1970 a 2009, e especificamente em Mato Grosso, de 1980 a 2009 (CONAB).

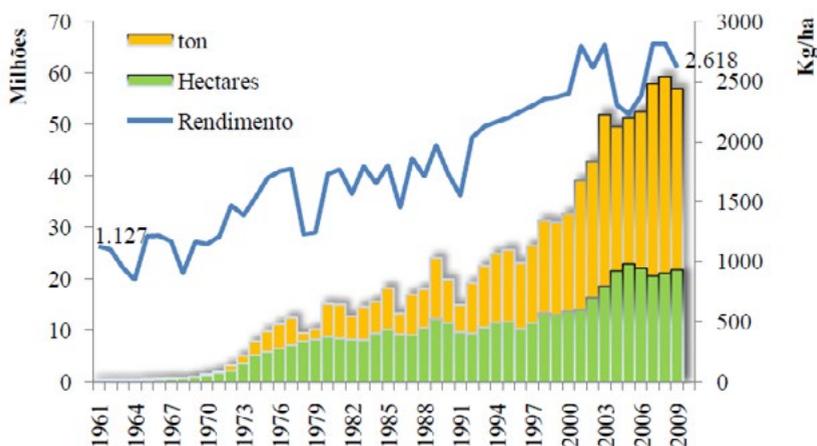
Considerando a importância dessa cultura para o estado de Mato Grosso e que o mercado do complexo soja é principalmente internacional, na formação do preço da soja interferem diversos fatores, tais como, o frete da soja, a soja Chicago Board of Trade (CBOT), o preço do dólar, entre outros. (APROSOJA, 2011)

Portanto, o objetivo do presente estudo foi identificar os fatores que mais influenciaram, no estado de Mato Grosso em 2011, na formação do preço da soja e como eles se relacionam.

## 2 Caracterização

Até o ano de 1950, a pequena produção da oleaginosa era consumida como forragem para bovinos ou para a engorda de suínos nas pequenas unidades produtoras do interior gaúcho. Sua trajetória de crescimento no país começou na década de 1960 e, em menos de 20 anos depois, converteu-se na principal cultura do agronegócio brasileiro. (IBGE, 2007) A Figura 1 ilustra a evolução da soja no Brasil:

Figura 1 – Evolução da soja no Brasil



Fonte: elaborado pelo autor, baseado em Faostat (2010).

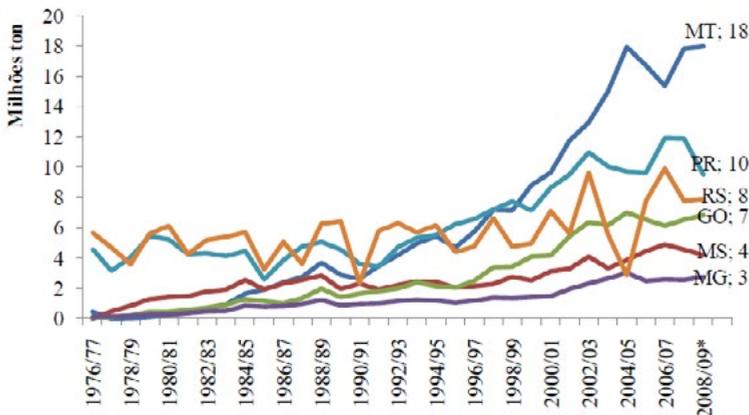
No contexto mundial, o Brasil figura como segundo produtor, com claras indicações de que será o primeiro ainda no correr desta década, dadas as limitações de área para expansão nos demais países produtores e pelo domínio tecnológico que o país possui para produzir em regiões tropicais com baixas latitudes. (AGNOL et al., 2007)

A primeira referência de produção comercial de soja no Brasil data de 1941 (área cultivada de 640 ha, produção de 450 toneladas e rendimento de 700 kg/ha), e o primeiro registro internacional do Brasil como exportador de soja data de 1949, com uma produção de 25 mil toneladas; alcançou as 100 mil em meados dos anos 50, e na década de 1960, a soja se estabeleceu definitivamente como cultura econômica importante para o Brasil, passando de 206 mil toneladas (1960) para 1.056 milhão (1969). Cerca de 98% desse volume era produzido nos três estados da região Sul, em áreas onde prevalecia a combinação: trigo no inverno e soja no verão. (IBGE, 2007)

Apesar do significativo crescimento da produção ao longo dos anos 1960, foi na década seguinte que a produção da soja mais evoluiu e se consolidou como a principal cultura do agronegócio nacional, passando de 1,5 milhão de toneladas em 1970 para mais de 15 milhões de toneladas em 1979. Esse crescimento se deve não apenas ao aumento da área cultivada (1,3 milhão ha para 8,8 milhões ha), mas também ao expressivo incremento da produtividade. (1.140 kg/ha, para 1.730 kg/ha) (CONAB).

No final dos anos 70, mais de 80% da produção brasileira de soja ainda se concentrava nos três estados da região Sul, embora os cerrados da região Centro-Oeste, Norte e Nordeste do país sinalizassem que participariam como importante ato no processo produtivo da oleaginosa, o que efetivamente ocorreu a partir da década de 1980. Em 1970, menos de 2% da produção nacional foi colhida nessa região e estava concentrada no estado de Mato Grosso do Sul. Em 1980, essa porcentagem passou para 20%, em 1990 já era superior a 40%, e em 2007 superou os 58%, com tendências a ocupar maior espaço a cada nova safra. Considerando-se a produção média dos anos 1970, com a produção de 2007, observa-se que a produção da região Sul cresceu apenas 3,36 vezes (7,3 milhões de toneladas para 24,5 milhões de toneladas), enquanto que na região central do Brasil cresceu 69 vezes (500 toneladas para 34,5 milhões de toneladas) Essa transformação promoveu e consolidou o estado de Mato Grosso (MT) como líder nacional da produção e da produtividade da soja, conforme exposto na Figura 2 (CONAB).

Figura 2 – Evolução da Produção por Estado



Fonte: elaborado pelo autor, baseado em CONAB (2010).

### 3 Metodologia

Neste estudo foram utilizados os dados dos boletins semanais do Instituto Mato-Grossense de Economia Agropecuária (IMEA) e dos preços do diesel do Sistema de Levantamento de Preços da Agência Nacional de Petróleo (SLP-ANP) do ano de 2011 para MT. (IMEA, 2011; SLP-ANP) Para as análises, os dados dos boletins semanais do IMEA foram organizados para ficarem em formato mensal e se constituem de 12 meses que correspondem ao ano de 2011.

Para análise dos dados, foram utilizadas as técnicas multivariadas de Análise Fatorial e Análise de Agrupamento.

Segundo Corrar e colaboradores (2007), a Análise Fatorial relaciona uma quantidade volumosa de variáveis de forma que se tornem variáveis latentes com quantidade reduzida.

A Análise Fatorial (AF) é uma técnica estatística que busca, através da avaliação de um conjunto de variáveis, a identificação de dimensões de variabilidade comuns existentes em um conjunto de fenômenos; o intuito é desvendar estruturas não observáveis diretamente. Cada uma destas dimensões de variabilidade recebe o nome de FATOR. (CORRAR; PAULO; DIAS FILHO, 2007, p. 74)

Em outras palavras, a AF pressupõe que as variáveis apresentam correlações elevadas de tal forma que se torna possível analisá-las em agrupamentos. Nesse caso, ao relacionar, por meio da AF, os 12 meses, pode-se constatar a existência ou não de correlações entre elas suficientes para a formação de fatores (grupos) que facilitem a análise, ou seja, é possível identificar quais variáveis estão relacionadas entre si e, principalmente, como estão relacionadas. (CORRAR; PAULO; DIAS FILHO, 2007)

Segundo Rezende e colaboradores (2007), cada variável pode ser definida por meio de uma linearidade combinada entre os fatores comuns, que explicam não só parte da variância de cada variável, que recebe o nome de Comunalidades, como um desvio que tenta minimizar os impactos da variância total não explicada pelos fatores, as especificidades. Para Corrar e colaboradores (2007), as Comunalidades assumem valores de 0 a 1, sendo que, quanto mais próximo de um, maior o nível de explicação das variâncias por parte dos fatores.

A partir da matriz de correlação entre as variáveis aplicadas a cada um dos meses, torna-se possível calcular os fatores por meio do método de Componentes Principais e, posteriormente, a quantidade dos fatores a serem considerados. (HAIR et al., apud REZENDE; FERNANDES; SILVA, 2007) Segundo Corrar e colaboradores (2007), o principal método utilizado para definir a quantidade de fatores necessários para explicar a variância das variáveis é o autovalor (*eigenvalue*). Os autovalores são utilizados para definição (aceitação) do número de fatores por meio do teste de Bartlett, em que se espera que a hipótese de não explicação das variâncias dos fatores seja rejeitada.

Para incrementar ainda mais a interpretação dos fatores, é utilizada uma rotação dos fatores de forma a permitir a análise de casos em que mais de um fator explica uma única variável. O método de rotação de fatores utilizado foi o varimax, que se define como uma rotação ortogonal e permite a minimização do fato descrito anteriormente, em outras palavras, permite a ocorrência de cargas fatoriais de cada variável para apenas um fator.

São as cargas fatoriais que definirão a qual fator determinada variável pertence. Assim, um fator será constituído por um grupo de variáveis e sua porcentagem de explicação perante o total de variáveis será dado de acordo com o tamanho desse grupo. A maior carga fatorial de uma variável frente aos fatores identificará a qual fator essa pertence, pois significa que é esse fator que melhor explica sua variância. Tendo definido quais variáveis pertencem a cada fator, cabe ao pesquisador definir o significado de cada um dos fatores e interpretá-los.

Já a Análise de Agrupamento comporta uma variedade de algoritmos de classificação diferentes, todos voltados para uma questão importante, ou seja, dirigidos pelos dados observados de forma a agrupar-se segundo características comuns que neles ocorram.

Esse processo deve levar em conta a possibilidade de se realizar uma organização hierárquica de grupos, em que, a cada nível de abstração maior, são também maiores as diferenças entre elementos contidos em cada grupo.

O método de Agrupamento pode ser realizado por dois procedimentos, sendo eles o do Agrupamento em Árvore e Agrupamento por k-Médias. Para o estudo, será apenas utilizado o Agrupamento em Árvore.

A Análise de Agrupamentos tem por objetivo agrupar as variáveis conforme a sua proximidade ou características comuns, buscando mostrar a homogeneidade dentro do grupo e a heterogeneidade entre os grupos. Essa análise torna-se interessante principalmente sob o aspecto descritivo, pois seu resultado final, nos métodos, é um gráfico de esquemas hierárquico denominado dendograma, também chamado de gráfico em árvore, onde o mesmo é uma representação gráfica que sintetiza o comportamento das variáveis.

Dada uma amostra de  $n$  objetos (ou indivíduos), cada um medido segundo  $p$  variáveis, procura-se um esquema de classificação que reúna os objetos em  $g$  grupos, exigindo conceitos científicos mais sofisticados de semelhança. Devem ser determinados também o número e as características desses grupos. (BUSSAB; MIAZAKI; ANDRADE, 1990)

A Análise Fatorial foi realizada com o programa Statistical Package for the Social Sciences (SPSS versão 20) e, para Análise de Agrupamento, foi utilizado o Minitab versão 16.

## 4 Análise dos resultados

A pesquisa foi feita através de dados secundários do IMEA e do SLP-ANP. Os dados envolvem nove variáveis.

No presente estudo, inicialmente, utilizou-se a Análise Fatorial de componentes principais a fim de identificar os fatores da formação do preço da soja. Para isso, aplicou-se nessa análise uma rotação nos dados para transformar os coeficientes das componentes principais retidas numa estrutura mais simplificada. O objetivo é dividir o conjunto inicial de variáveis em subconjuntos com maior grau de independência possível. Optou-se pela rotação varimax. Esse processo pretende que, para cada componente principal, existam apenas alguns pesos significativos e todos os outros sejam próximos de zero, através da maximização da variância entre os fatores para a rotação das matrizes fatoriais.

Para efetuar o procedimento referido, utilizou-se dos testes de Kaiser-Meyer-Olkin (KMO) e de Esfericidade de Bartlett. Os resultados para esses testes são apresentados na Tabela 1. O primeiro, teste de KMO, teve o resultado de 0,560. Como esse valor de explicação foi maior do que 0,50, significa que os fatores encontrados na Análise Fatorial conseguem descrever, satisfatoriamente, as variações dos dados originais. O segundo, teste de esfericidade de Bartlett, teve resultado menor que o nível de significância de  $\alpha = 5\%$ , concluindo-se que vários fatores poderão explicar grande proporção da variabilidade dos dados.

Tabela 1 – Teste de KMO e Bartlett

Medida Kaiser-Meyer-Olkin de adequação de amostragem.		0,560
Teste de Esfericidade de Bartlett	Chi-quadrado aprox.	87,730
	df	36
	Sig.	0,000

Fonte: Análise própria no software SPSS 20.0

Também contemplando a Comunalidade das variáveis analisadas, pode ser observado na Tabela 2 que praticamente todas as variáveis possuem um poder de explicação alto, considerando todos os fatores obtidos, com exceção da variável prêmios, que apresentou uma explicação abaixo de 0,70.

Tabela 2 – Comunalidades

Variáveis	Inicial	Extração
Frete soja	1,000	0,775
Preço diesel	1,000	0,899
Dólar	1,000	0,941
Prêmios	1,000	0,552
Soja CBOT	1,000	0,961
Soja BM&AMP;F	1,000	0,871
Insumos	1,000	0,868
Operações agrícolas	1,000	0,948
Outros custos	1,000	0,827

Fonte: análise própria no software SPSS 20.0.

Além da tabela de Comunalidades, a Tabela 3, de variação total explicada, também demonstra de forma bastante razoável como se seguir com o modelo de análise. O percentual de variação explicada foi de aproximadamente 68,11% com uso de dois fatores, e de 84,91% com o uso de três fatores.

Tabela 3 – Variação total explicada

	Valores próprios iniciais			Somadas de extração de carregamentos ao quadrado			Somadas rotativas de carregamentos ao quadrado		
	Total	Varia. %	Cumula. %	Total	Varia. %	Cumula. %	Total	Varia. %	Cumula. %
1	5,130	56,997	56,997	5,130	56,997	56,997	3,500	38,894	38,894
2	1,451	16,120	73,116	1,451	16,120	73,116	2,629	29,212	68,106
3	1,061	11,793	84,909	1,061	11,793	84,909	1,512	16,803	84,909
4	0,628	6,978	91,887						
5	0,386	4,285	96,172						
6	0,244	2,716	98,888						
7	0,075	0,830	99,718						
8	0,018	0,195	99,912						
9	0,008	0,088	100,000						

Fonte: análise própria no software SPSS 20.0.

Como se observa na Tabela 3, a coluna “total” dos “valores próprios iniciais” está ordenada de acordo com seu tamanho. Na situação inicial, a soma dos “valores próprios iniciais” iguala-se ao número de variáveis em análise, que, nesse caso, são de 18 variáveis. Desses valores, apenas sete são maiores do que um, e, por isso, o número de fatores retidos é de somente sete. Desse modo, em vez de se trabalhar com 18 indicadores, pode-se utilizar apenas sete fatores, uma vez que esses fatores são responsáveis por explicar 84,91% da variação total entre os indicadores.

Dessa forma, acredita-se ter chegado a um grau de relacionamento e explicação das variáveis capaz de ser útil na avaliação. Cabe, agora, identificar quais indicadores fazem parte de cada um dos fatores. A tabela matriz de componentes rotacionados (Tabela 4) permite verificar qual dos fatores melhor explica cada um dos indicadores considerados. Essa matriz indica quais variáveis pertencem a cada fator, identificando os principais indicadores da pesquisa.

Na matriz componente rotativo pertencerão a um fator as cargas com maior valor absoluto em módulo. No estudo foram consideradas relevantes apenas as cargas fatoriais (correlações) maiores que 0,50, destacadas em negrito na Tabela 4.

Tabela 4 – Matriz de componente rotativa

Variável	Fator		
	1	2	3
Frete soja	0,745	-0,241	-0,401
Preço diesel	-0,911	-0,257	0,060
Dólar	-0,742	0,508	0,364
Prêmios	-0,591	0,017	-0,449
Soja CBOT	0,451	0,867	0,068
Soja BM&F	0,259	-0,528	0,725
Insumos	-0,926	0,063	0,076
Operações agrícolas	0,973	-0,024	-0,039
Outros custos	0,877	0,181	0,158

Método de extração: Análise do Componente principal. Método de rotação: Varimax com normalização de Kaiser. Rotação convergida em 11 iterações.

Fonte: análise própria no software SPSS 20.0.

Assim, os indicadores que sugerem cada um dos fatores são compostos a seguir:

- o Fator 1 é composto por frete soja, preço diesel, dólar, prêmios, insumos, operações agrícolas e outros custos;
- o Fator 2 é composto por dólar, soja cbot, soja BM&F;
- o Fator 3 é composto por soja BM&F;

No modelo foi possível interpretar o primeiro fator como sendo o “custo de produção soja”, o segundo fator pode ser interpretado “cotações” e o último fator sendo “BM&F da soja”.

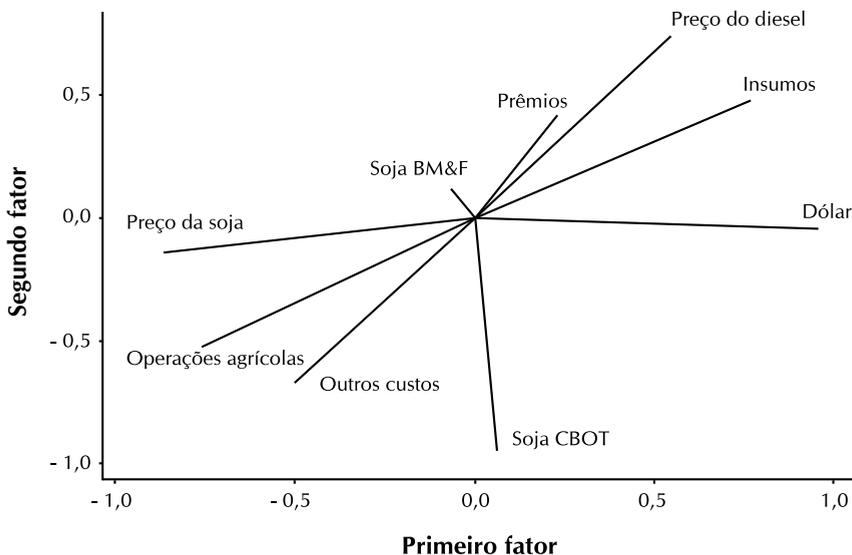
O primeiro fator (“custo de produção soja”) é responsável por 38,89% da variância explicada e o segundo fator (“cotações”) por 29,12%. A “BM&F da soja” é constituída por 16,80% da variância.

A Análise Fatorial finaliza apresentando o gráfico das cargas considerando os dois primeiros fatores (Figura 3).

Na Figura 3 observa-se que as variáveis que mais contribuem com o custo da soja entre as aqui consideradas são a soja CBOT, operações agrícolas, o dólar e o preço do diesel.

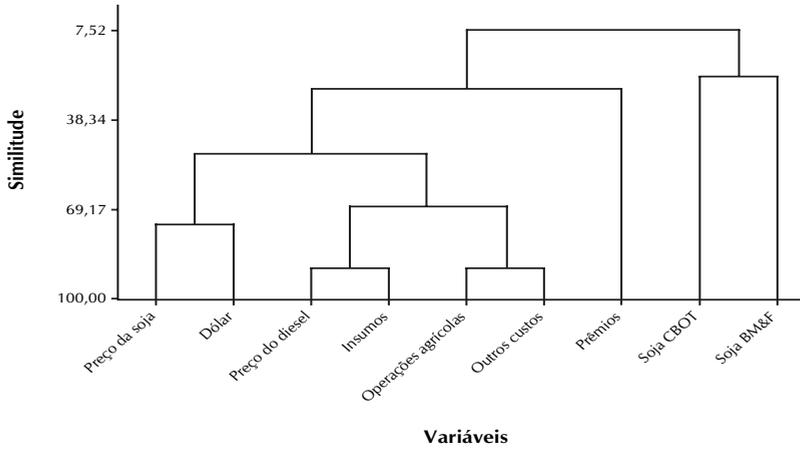
Na Análise de Agrupamento são considerados os gráficos de agrupamentos denominados dendogramas. Na Figura 4 é apresentado um dendograma por variável, considerando o método de ligação completa e a correlação absoluta e utilizando a similitude entre as variáveis.

Figura 3 – Gráfico das cargas



Fonte: análise própria no software Minitab 16.

Figura 4 – Dendograma por variáveis, considerando nove variáveis

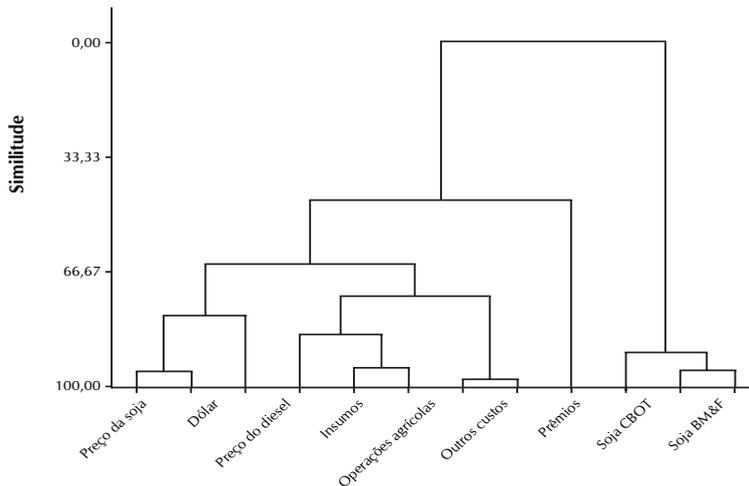


Fonte: análise própria no software Minitab 16.

Na Figura 4 observa-se que as variáveis, segundo os custos da soja, são agrupadas em cinco grupos. As variáveis mais similares são as do segundo grupo, formadas por preço do diesel, insumos, operações agrícolas e outros custos.

Na Figura 5 é apresentado um dendograma por mês, considerando o método de ligação completa, a distância euclidiana e as variáveis padronizadas e utilizando a similitude entre os meses.

Figura 5 – Dendograma por mês, considerando os 12 meses do ano de 2011



Fonte: análise própria no software Minitab 16.

Na Figura 5 observa-se que os meses, segundo os custos da soja, são agrupados em cinco grupos. Os meses mais parecidos são julho e agosto (grupo 3), seguidos de janeiro e fevereiro (grupo 1) e novembro e dezembro (grupo 5). Sendo que o primeiro grupo se refere à colheita da safra da soja, e o último grupo é da época do plantio. (APROSOJA, 2011)

## 5 Conclusão

A partir da Análise Fatorial dos resultados obtidos, foi possível identificar os fatores mais relevantes que influenciaram na formação do preço da soja no estado de Mato Grosso em 2011. Essas variáveis analisadas foram a soja CBOT, operações agrícolas, o dólar e o preço do diesel.

Cabe observar que as nove variáveis podem ser substituídas por três fatores. O primeiro classificado como “custo de produção soja”, o segundo como “cotações” e o terceiro como “BM&F da soja”. Esses três fatores, quando empregados conjuntamente, explicaram aproximadamente 85% do total de variações nos demais índices.

## Referências

AGNOL, A. D. et al. *O complexo agroindustrial da Soja Brasileira*. Londrina: Embrapa, 2007. 12 p. (Circular Técnica, 43).

APROSOJA – Associação dos Produtores de Soja do Estado de Mato Grosso. Cuiabá, MT, [2005-]. Disponível em: <<http://www.aprosoja.com.br/>>. Acesso em: 21 mar. 2013.

BATALHA, M. O. (Coord.). *Gestão agroindustrial*. São Paulo: Atlas, 2007. 2 v.

BELFIORE, P. P.; FÁVERO, L. P. L.; ANGELO, C. F. Aplicação de técnicas estatísticas multivariadas em empresas de operação logística no Brasil em função de indicadores econômico-financeiros. *REAd: Revista Eletrônica de Administração*, Porto Alegre, v. 12, n. 3, p. 1-22, maio/jun. 2006.

BEZERRA, F. A.; CORRAR, L. J. Utilização da análise fatorial na identificação dos principais indicadores para a avaliação do desempenho financeiro: uma aplicação nas empresas de Seguro. *Revista de Contabilidade e Finanças*, São Paulo, v. 1, n. 42, p. 50-62, set./dez. 2006.

BONJOUR, S. C. M.; FIGUEIREDO, A. M. R.; MARTA, J. M. C. A pecuária de corte no estado de Mato Grosso. In: CONGRESSO BRASILEIRO DE

ECONOMIA E SOCIOLOGIA RURAL DA SOBER, 46., Rio Branco, AC. *Anais...* Viçosa, MG: SOBER, 2008.

BUSSAB, W. O.; MIASAKI, E. S.; ANDRADE, D. F. de. Introdução à Análise de Agrupamentos. In: SIMPÓSIO BRASILEIRO DE PROBABILIDADE ESTATÍSTICA, 9., 1990, São Paulo. *Anais...* São Paulo: ABE: USP, 1990.

CONAB: Companhia Nacional de Abastecimento. Brasília, DF, [1991-]. Disponível em: <<http://www.conab.gov.br/index.php>> . Acesso em: 22 dez. 2011.

CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. *Análise Multivariada: para os cursos de Administração, Ciências Contábeis e Economia*. São Paulo: Atlas, 2007. 541 p.

IBGE – Instituto Brasileiro de Geografia e Estatística. *Estatísticas de Empreendedorismo*. Rio de Janeiro: IBGE, 2011. Disponível em: <<http://www.ibge.gov.br/home/estatistica/economia/empreendedorismo/2011/>> .

IMEA – Instituto Mato-grossense de Economia Agropecuária. *Boletim semanal da soja*. Cuiabá, MT, 1998-. Disponível em <<http://www.imea.com.br/publicacoes.php?categoria=4&subcategoria=2>> . Acesso em: 2 mar. 2013.

MAPA – Ministério da Agricultura, Pecuária e Abastecimento. *Serviço de Inspeção Federal (SIF)*. Brasília, DF, 2001-. Disponível em: <[http://sigsif.agricultura.gov.br/sigsif\\_cons/lap\\_estabelec\\_nacional\\_rep](http://sigsif.agricultura.gov.br/sigsif_cons/lap_estabelec_nacional_rep)> . Acesso em: 21 mar. 2013.

REZENDE, M. L.; FERNANDES, L. P. de S.; SILVA, A. M. R. Utilização da Análise Fatorial para Determinar o Potencial de Crescimento Econômico em uma Região do Sudeste do Brasil. *Revista Economia e Desenvolvimento*, Santa Maria, n. 19, p. 92-109. 2007.

ANP - Agência Nacional de Petróleo: Sistema de Levantamento de Preços da Agência Nacional de Petróleo. [S.l.], Disponível em: <<http://www.anp.gov.br/preco/>> . Acesso em: 2 mar. 2013.

