

# O MÉTODO DE ANÁLISE FATORIAL APLICADO AO ESTUDO DE RESULTADOS DE ANÁLISE DE SOLOS

**Jorge Roberto Grobe (1) & Jair M. Marques (2)**

(1) Professor licenciado em Ciências habilitação Matemática, M.Sc. em Métodos Numéricos em Engenharia pela UFPR, curso de Sistemas de Informação, UTFPR Pato Branco. (2) Professor Jair Mendes Marques, D.Sc, Programa de Pós – Graduação em Métodos Numéricos em Engenharia-UFPR.

[grobe@ufpr.edu.br](mailto:grobe@ufpr.edu.br); [jair.marques@utp.br](mailto:jair.marques@utp.br)

**Resumo** – Este trabalho tem o objetivo de explicar o uso de técnicas de estatística multivariada em resultado de análises de solos. Os resultados de análises de solos foram obtidos através de um experimento agrônômico, realizado na Fazenda Guamirim, município de Água Doce, SC, região denominada “Campos de Palmas”. A população estudada é composta de 11 variáveis e 18 amostras. A ferramenta utilizada para interpretação dos dados foi análise fatorial via componentes principal, usada para fornecer as ponderações, escolher o conjunto de variáveis adequadas, e também, os pesos atribuídos a cada variável. Nesta técnica foram obtidos dois fatores de variáveis relacionados pelas suas correlações. Os dois fatores juntos explicam 81,40% dos dados coletados. Para determinação dos resultados, foi utilizado o software STATISTIC 6.0.

**Palavras – chave** – estatística multivariada, análise fatorial, componentes principais.

# **O MÉTODO DE ANÁLISE FATORIAL APLICADO AO ESTUDO DE RESULTADOS DE ANÁLISES DE SOLOS**

## **1. INTRODUÇÃO**

Para FERREIRA (1996), os métodos estatísticos multivariados são estudados a partir de um conjunto de muitas variáveis, onde os dados são coletados ou mensurados. Para analisar os dados coletados existem alguns objetivos de técnicas multivariadas:

- Redução de dados ou simplificação estrutural;
- Ordenação e agrupamento;
- Investigação da dependência entre variáveis;
- Predição de uma ou mais variáveis com base na observação de outras variáveis.

Segundo SENA (1999), o solo é uma estrutura complexa formada pela influência de fatores geológicos, topográficos, climáticos, temporais, sendo a parte mais importante da geosfera, e de onde provêm a maior parte dos nossos alimentos. O interesse na análise de solos é a avaliação de seus parâmetros químicos (pH, Ca, H+Al, NO<sub>3</sub>), etc. Com a técnica de análise fatorial é possível explicar o máximo de intercorrelação entre as variáveis e examinar as relações entre os objetos através de gráficos dos “pesos” projetados pelos fatores.

Em BARROSO (2003), esta técnica possibilita saber o quanto cada fator está associado a cada variável e o conjunto de fatores explica a variabilidade total dos dados originais. A análise fatorial possibilita trabalhar com um número reduzido de variáveis sem muita perda significativa de informações. Um dos métodos mais utilizados baseia-se na análise de componentes principais, sendo que neste método não há pressuposição da normalidade dos dados. Na técnica multivariada, a análise fatorial via método das componentes principais tem o intuito de resumir o padrão de correlação entre as variáveis e muitas vezes é possível chegar a

um conjunto de variáveis que não sejam correlacionadas umas com as outras. Algebricamente, as componentes principais são combinações lineares das variáveis originais.

## **2. MATERIAL E MÉTODOS**

GROBE (2005) em seu trabalho para fins de estudos utilizou uma matriz de dados dos resultados de análises de solos conforme tabela 1, proveniente de um experimento agrônomico realizado na fazenda Guamirim , “Campos de Palmas”, localizada no município de Água Doce (SC). Os resultados das análises da rotina de solos foram coletados em três profundidades: 0-5 cm, 5-10 cm e 10 -20 cm e as variáveis analisadas pelo laboratório de solos da UTFPR - Pato Branco foram no total de 11 : potencial de hidrogênio, alumínio, hidrogênio mais alumínio (acidez total), cálcio, fósforo, potássio e saturação por bases (V), nitrato e amônia.

Segundo SENA (2000), este tipo de análise de dados deve ser sempre autoescalados (média zero e variância um) para assegurar que todas as variáveis contribuam igualmente para o modelo, independente da escala.

O objetivo geral deste trabalho é explicar as correlações e a variabilidade entre um conjunto grande de variáveis e um conjunto de poucas variáveis aleatórias não observáveis chamadas fatores.

Na tabela 2 tem-se a matriz das correlações da matriz de dados.

TABELA 1 –MATRIZ DE DADOS

Bloco	Prof.	pH	MO	Al	H+Al	Ca	Mg	K	P	V	NH <sub>4</sub>	NO <sub>3</sub>
	cm	CaCl <sub>2</sub>	g dm <sup>-3</sup>	cmol <sub>c</sub> <sup>+</sup> dm <sup>-3</sup>						mg dm <sup>-3</sup>		mg/kg
1	0-5	4,20	53,61	2,20	14,44	2,70	2,20	0,23	16,94	0,26	4,20	57,87
1	5-10	4,00	53,61	3,80	15,77	0,63	1,14	0,10	0,93	0,11	4,20	36,17
1	10-20	4,00	0,00	4,05	13,06	0,40	0,87	0,08	0,93	0,09	4,20	20,67
1	0-5	4,2	53,6	3,1	15,8	1,5	1,4	0,1	10,6	0,2	3,6	20,7
1	5-10	4,0	53,6	4,1	17,2	0,6	1,3	0,1	0,9	0,1	1,8	13,4
1	10-20	4,0	46,9	4,7	15,2	0,5	0,7	0,1	0,9	0,1	2,4	20,7
2	0-5	4,00	56,61	4,24	15,77	0,66	0,94	0,10	1,25	0,10	1,80	33,07
2	5-10	4,00	44,23	4,09	13,06	0,35	0,55	0,05	0,61	0,07	2,10	32,03
2	10-20	4,00	40,21	3,75	13,06	0,33	0,67	0,05	0,61	0,07	2,70	35,13
2	0-5	4,00	53,61	3,11	14,44	1,18	1,72	0,10	8,13	0,17	3,00	24,80
2	5-10	3,90	53,61	4,13	15,77	0,53	1,17	0,10	1,25	0,10	2,70	20,67
2	10-20	3,90	53,61	5,10	18,82	0,40	0,50	0,08	0,61	0,05	1,80	21,70
3	0-5	4,00	60,31	4,58	18,82	0,90	1,17	0,15	1,57	0,11	1,80	21,70
3	5-10	3,90	53,61	4,53	15,77	0,32	0,58	0,08	0,93	0,06	2,40	10,33
3	10-20	3,90	46,91	4,25	14,08	0,20	0,40	0,05	0,00	0,04	2,10	16,53
3	0-5	4,20	60,31	2,25	14,44	2,30	2,00	0,18	15,36	0,24	2,10	26,87
3	5-10	4,00	60,31	5,07	17,23	0,75	0,85	0,13	1,57	0,09	2,10	31,00
3	10-20	3,90	60,31	5,15	18,82	0,40	0,50	0,10	0,61	0,05	1,80	25,83

Fonte: Laboratório de Solos da UTFPR- Pato Branco

TABELA 2 – MATRIZ DAS CORRELAÇÕES

	pH	MO	Al	H+Al	Ca	Mg	K	P	V	NH <sub>4</sub>	NO <sub>3</sub>
pH	1	0,0716	-0,8134	-0,2951	0,8759	0,7843	0,7187	0,8806	0,8663	0,4289	0,4922
MO	0,0716	1	0,0102	0,5717	0,2833	0,2191	0,3788	0,2164	0,1781	-0,4743	0,0957
Al	-0,8134	0,0102	1	0,5558	-0,8239	-0,8479	-0,5493	-0,8629	-0,8937	-0,5382	-0,4677
H+Al	-0,2951	0,5717	0,5558	1	-0,1189	-0,1456	0,2183	-0,2165	-0,2422	-0,4713	-0,2487
Ca	0,8759	0,2833	-0,8239	-0,1189	1	0,9132	0,8915	0,9771	0,9701	0,3525	0,5465
Mg	0,7843	0,2191	-0,8479	-0,1456	0,9132	1	0,8359	0,8623	0,9726	0,4248	0,4398
K	0,7187	0,3788	-0,5493	0,2183	0,8915	0,8359	1	0,7997	0,8372	0,2207	0,4545
P	0,8806	0,2164	-0,8629	-0,2165	0,9771	0,8623	0,7997	1	0,9452	0,3825	0,4855
V	0,8663	0,1781	-0,8937	-0,2422	0,9701	0,9726	0,8372	0,9452	1	0,4472	0,5212
NH <sub>4</sub>	0,4289	-0,4743	-0,5382	-0,4713	0,3525	0,4248	0,2207	0,3825	0,4472	1	0,4219
NO <sub>3</sub>	0,4922	0,0957	-0,4677	-0,2487	0,5465	0,4398	0,4545	0,4855	0,5212	0,4219	1

### 3. RESULTADOS E DISCUSSÃO

Na tabela 3, apresentam-se os fatores Y1 e Y2 da matriz de dados com seus respectivos pesos de cada variável.

TABELA 3 – CARGAS FATORIAIS PELO MÉTODO DAS COMPONENTES PRINCIPAIS

variáveis	Y1	Y2
pH	<b>-0,9066</b>	-0,0470
MO	-0,1552	<b>0,8694</b>
Al	<b>0,9033</b>	0,2627
H+Al	0,2755	<b>0,8449</b>
Ca	<b>-0,9739</b>	0,1821
Mg	<b>-0,9372</b>	0,1118
K	<b>-0,8290</b>	0,4272
P	<b>-0,9570</b>	0,0856
V	<b>-0,9865</b>	0,0480
NH <sub>4</sub>	-0,5064	<b>-0,6351</b>
NO <sub>3</sub>	<b>-0,6017</b>	-0,1101
Expl.Var	6,7596	2,1940
Prp.Totl	61,45%	19,95%

A análise fatorial obteve dois fatores Y1 e Y2 que podem ser escritos em forma de equações:  $Y1 = -0,9066 \cdot \text{pH} - 0,1522 \cdot \text{MO} + 0,9033 \cdot \text{Al} + 0,2755 \cdot (\text{H} + \text{Al}) - 0,9739 \cdot \text{Ca} - 0,9372 \cdot \text{Mg} - 0,8290 \cdot \text{K} - 0,9570 \cdot \text{P} - 0,9865 \cdot \text{V} - 0,5064 \cdot \text{NH}_4 - 0,6017 \cdot \text{NO}_3$  e  $Y2 = -0,0470 \cdot \text{pH} + 0,8694 \cdot \text{MO} + 0,2627 \cdot \text{Al} + 0,8449 \cdot (\text{H} + \text{Al}) + 0,1821 \cdot \text{Ca} + 0,1118 \cdot \text{Mg} + 0,4272 \cdot \text{K} + 0,0856 \cdot \text{P} + 0,0480 \cdot \text{V} - 0,6351 \cdot \text{NH}_4 - 0,1101 \cdot \text{NO}_3$ . O primeiro fator Y1 explica 61,45% e o segundo fator Y2 explica 19,95% da variabilidade dos dados relativo aos dados das 18 amostras de resultados de análises de solos e das 11 variáveis. Os fatores Y1 e Y2 representam 81,40% da variabilidade dos dados. As variáveis que pertencem ao fator Y1 são: pH, Al, Ca, Mg, K, P, V, NO<sub>3</sub> e ao fator Y2 são: MO, H+Al, NH<sub>4</sub>. Observa-se que as variáveis que obtiveram maior peso ou saturação no modelo são cálcio (Ca) e saturação por base (V) que foram na ordem de 0,9033 e -0,9856 respectivamente, e as menores cargas

fatoriais foram nitrato ( $\text{NO}_3$ ) na ordem de -0,6017 e amônia ( $\text{NH}_4$ ) na ordem de -0,6351 em valores absoluto.

Na figura 2 os fatores com suas cargas fatoriais de suas respectivas variáveis, calculados pelo método das componentes principais:

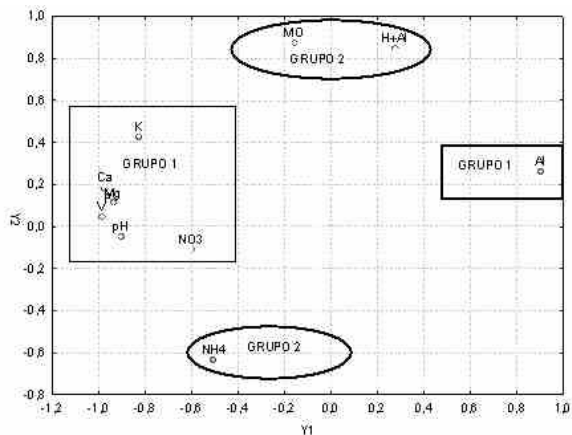


FIGURA – 2 ORDENAÇÃO DAS CARGAS FATORIAIS PELO MÉTODO DAS COMPONENTES PRINCIPAIS

No gráfico da figura 2 estão relacionados as 11 variáveis em dois grupos. As variáveis que pertencem ao grupo 1; são : pH, Al, Ca, Mg, K, P, V e  $\text{NO}_3$  e ao grupo 2, são: MO, H+Al e  $\text{NH}_4$ . As variáveis de cada grupo relacionado possuem uma alta correlação entre si. No grupo 2 as variáveis H+Al,  $\text{NH}_4$  não tiveram uma correlação acima 0,7 devido as suas cargas fatoriais abaixo de 0,7, portanto elas podem ser descartadas do modelo.

#### 4. CONCLUSÕES

Através da técnica de estatística multivariada, análise fatorial pelo método das componentes principais, foram obtidos dois fatores: Y1 e Y2. As cargas fatoriais obtidas foram

importantes para seus agrupamentos quanto à correlação das variáveis.

## 5. REFERÊNCIA

BARROSO, L. P. Análise Multivariada. 48<sup>a</sup> Reunião da RBRAS e 10<sup>o</sup> SEGRO –7 a 11 de julho de 2003 – Lavras MG. Departamento de Ciências e Exatas. Universidade Federal de Lavras

FERREIRA, D. F. Análise Multivariada. Departamento de Exatas. Apostila 400 f. Universidade Federal de Lavras. Lavras. MG. 1996.

GROBE, J. R. Aplicações da Estatística Multivariada na Análise de Resultados em Experimentos com Solos e Animais. Curitiba. 2005. 145 p. Dissertação (Mestrado em Métodos Numéricos em Engenharia) - Setores de Tecnologia e Ciências Exatas, Universidade Federal do Paraná (UFPR).

SENA, M. M.; POPPI, R. J. Avaliação do Uso de Métodos Quimiométricos em Análise de Solos. *Química Nova*, 23(4), p. 547-556, 2000.