

APLICAÇÃO DA ANÁLISE FATORIAL AO ESTUDO DE ELEMENTOS PRINCIPAIS NAS ROCHAS BASÁLTICAS DA BACIA DO PARANÁ

ARLEI B. MACEDO*, N. R. RÜEGG*

ABSTRACT

The technique of factor analysis is applied to the study of factors which describes the variability in 197 chemical analyses of basaltic rocks found throughout the magmatic province of the Paraná Basin. The mathematical techniques are delineated and the available computer programs compared. The analysis are then recalculated in terms of percentual factor coefficients, and the results compared with equivalent data available in the literature for basaltic rocks of world distribution.

RESUMO

A técnica da análise fatorial é aplicada na determinação de fatores que descrevem a variabilidade de 197 análises químicas de rochas basálticas distribuídas por toda a província magnética da Bacia do Paraná. São delineadas as técnicas matemáticas e comparadas as características dos programas de computador disponíveis. As análises são recalculadas em termos de coeficientes fatoriais percentuais e os resultados comparados a dados equivalentes disponíveis na literatura para rochas basálticas de distribuição mundial.

INTRODUÇÃO

A análise fatorial constitui, modernamente, útil instrumento no estudo de grandes séries de dados constituídos de observações multivariáveis, como é, por exemplo, o caso de análises químicas de rochas. A técnica estatística para a sua utilização foi desenvolvida no início do século por psicólogos visando estabelecer fatores que controlassem a inteligência. Estendeu-se, porém, mais tarde, a outros ramos da Ciência, inclusive a Geologia em cuja área a técnica foi incorporada na última década. Entre os trabalhos pioneiros nesta área distingüem-se os de Mertie, 1961 e Imbrie, 1962.

Utilizando a técnica, mais recentemente, Manson, 1967 e Prinz, 1967, estudaram as características geoquímicas de rochas basálticas de distribuição mundial; o primeiro no que se refere à distribuição dos elementos principais e o segundo no que se refere aos elementos traços. Hoje considerados importantes trabalhos de referência, neles, por meio da análise fatorial, foram representados em termos de variáveis derivadas (fatores) em número menor que as originais (teores dos elementos expressos na forma de óxidos), as características do total das observações reunidas (análises químicas). Mantiveram-se, porém, por meio dos recursos propiciados pela análise fatorial, as características totais de variabilidade química eventualmente existente no conjunto reunido de observações. Criaram desta forma arcabouços de referência muito úteis quando se trata de, por comparação, traçar em bloco as características químicas de determinada província basáltica. Comparação desse gênero em relação ao modelo fornecido por Manson, *op. cit.*, será ora apresentada com dados referentes à província basáltica do Brasil meridional, com base em amplo documentário químico de elementos principais dosados em rochas que ocorrem na Bacia do Paraná. Resultados de estudo semelhante, utilizando-se porém, grande número de análises de elementos traços dosados nas mesmas rochas e os esquemas fornecidos por Prinz, *op. cit.*, estão sendo divulgados em trabalho paralelo.

* IG/USP

GENERALIDADES SOBRE ANÁLISE FATORIAL

A técnica estatística da análise factorial pode ser empregada para explorar matrizes de dados, que exprimam observações multivariáveis, com o objetivo de determinar possíveis relações existentes entre as variáveis e/ou entre as observações e o número mínimo de fatores capazes de explicar tais relações. Emprega-se a técnica também para testar hipóteses que explicam a variabilidade de conjuntos de dados por meio de fatores empíricos idealizados pelos pesquisadores. Serve, ainda, para comparar dados multivariáveis, expressos por meio de coeficientes derivados dos fatores.

Para serem alcançados os objetivos acima expostos, vários artifícios de cálculo são necessários, artifícios estes de matemática e interpretação complexa. A interpretação, entretanto, é facilitada por meio de exemplos de aplicação da análise factorial na determinação da composição química dos minerais e de que forma essa técnica estatística permite expressar por meio de um número menor de fatores a variabilidade existente na composição química desses minerais. Exemplos simples desta natureza podem ser encontrados em Manson, *op. cit.*, porém, o caso dos plagioclásios, descrito por McIntire, 1969, é mais ilustrativo.

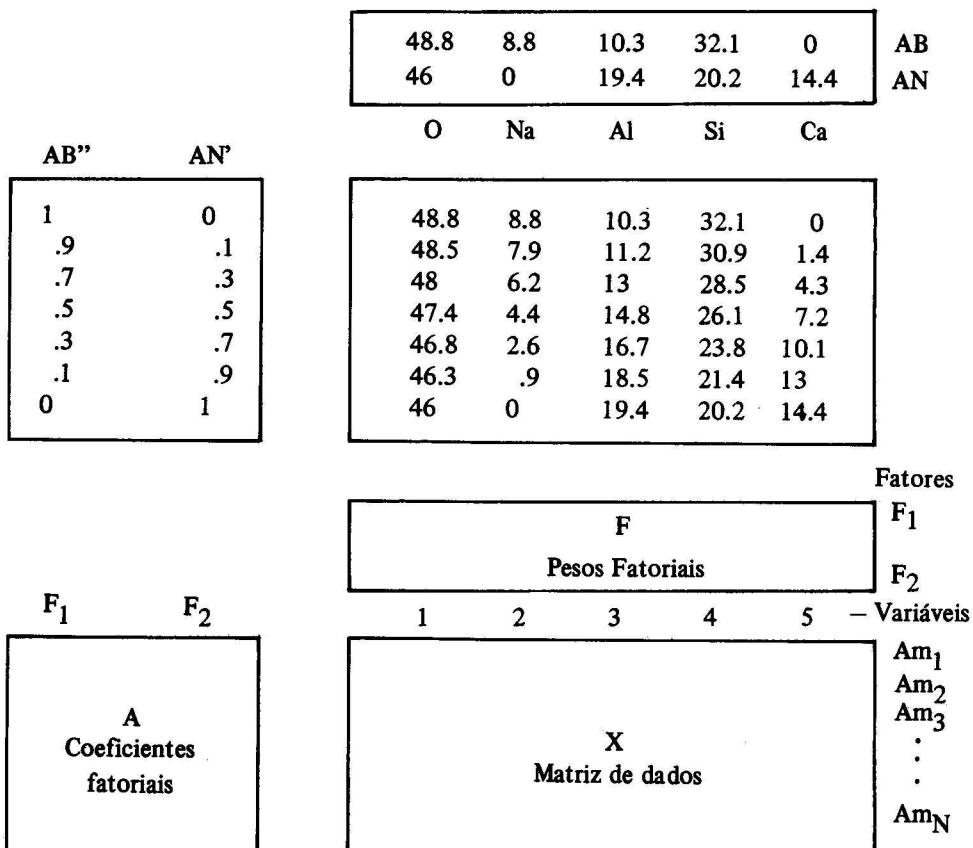


Fig. 1: Acima: matrizes que representam a composição química de plagioclásios expressa segundo teores de elementos químicos e suas respectivas composições expressas segundo proporções de Ab e An, incluindo a composição de Ab e An. Embaixo: Estrutura das referidas matrizes expressa segundo a nomenclatura da análise factorial, conforme McIntire, 1969.

Conforme demonstrou McIntire, *op. cit.*, a composição dos plagioclásios pode ser representada pelas matrizes dispostas na Figura 1 que, operadas adequadamente, resultam em composições expressas em termos de 2 fatores F₁ e F₂. Estes correspondem, respectivamente, à forma mais usual para descrever a composição de plagioclásios Ab e An. No momento que quisermos saber a composição de um determinado plagioclássio em termos de seus principais elementos químicos, bastará multiplicar o vetor que exprime esta composição, por meio de fatores,

pela matriz dos pesos fatoriais. Reconstituiremos dessa forma, a linha da matriz X que corresponde à composição desejada. É evidente que a matriz A de coeficientes fatoriais é muito mais fácil de interpretar que a matriz X de coeficientes não reduzidos. A partir dela é possível agora lançar num simples gráfico linear, pontos que correspondam à composição de plagioclásios expressos como coeficientes fatoriais e compará-los entre si, operação muito mais difícil de realizar com os 5 elementos químicos. Isto se tornaria ainda mais complicado, se as composições envolvessem maior número de elementos que é o caso de análises químicas de rochas.

Para efetuar a redução de grande número de variáveis, que é o caso de análises químicas de rochas, a coeficientes fatoriais, podemos representar uma matriz de dados na forma de um produto de duas matrizes mais simples, segundo o procedimento proposto por Manson, *op. cit.*, conforme a equação abaixo:

$$X_{(N,n)} = A_{(N,m)} \cdot F_{(m,n)} + E_{(N,n)}$$

onde os símbolos representam:

N = número de observações (análises químicas)

n = número de variáveis (teores dos vários elementos dosados)

$X_{(N,n)}$ = matriz inicial de dados

m = número de fatores capaz de explicar a variabilidade dos dados. Este número é determinado por 2 critérios: deve ser capaz de explicar a maior porcentagem possível da variabilidade e para facilidade de interpretação e representação gráfica, o menor possível.

$A_{(N,m)}$ = matriz de coeficientes fatoriais, descrevendo as N análises como proporções de m membros finais (fatores)

$F_{(m,n)}$ = matriz de pesos fatoriais que descreve a composição de cada um dos fatores (membros finais) como combinação das n variáveis.

$E_{(N,n)}$ = matriz da variabilidade não representada por A e F. Uma boa solução matemática deve fazer com que esta matriz tenha o valor mais próximo possível de zero. Geralmente esta condição não é atingida, porque quando reduzimos o número de fatores em relação às variáveis, perdemos em precisão uma parte do que ganhamos em simplicidade. Além disso os erros de medidas também contribuirão para formar a matriz E.

Para extrair os fatores da matriz inicial são empregadas técnicas de cálculo matricial cujo procedimento pormenorizado pode ser encontrado, por exemplo, em Harman, 1960, ou resumidamente, porém com maior enfoque geológico, em Lafitte, 1972, ou, ainda, em Nie *et alii* 1970, que usa linguagem matemática mais simples.

Esquematicamente, porém, os passos necessários para se obter os fatores são os seguintes:

1. extrair da matriz inicial de dados uma matriz de coeficientes de similaridade. Estes podem ser coeficientes de correlação, de distância, tabelas de variância-covariância, co-seno de Θ , etc,
2. calcular os valores característicos da matriz de coeficientes de similaridade e, por meio deles, decidir quantos fatores são necessários para explicar a similaridade dos dados;
3. calcular a matriz fatorial inicial (de pesos fatoriais);
4. fazer girar os fatores de modo a tornar os pesos fatoriais mais facilmente interpretáveis;
5. calcular a matriz de estimativa de fatores, que será utilizada para obter os coeficientes fatoriais a partir da matriz inicial.

Devido ao enorme volume de cálculo necessário para a execução de processamento estatístico, ele é realizado em computador eletrônico. Os programas de computador para efetuar a análise fatorial encontram-se incorporados a bem conhecidos conjuntos padrões (*packages*) de programas estatísticos de aplicação geral. Vários desses conjuntos são disponíveis na Universidade de São Paulo. São os chamados SSP (*Scientific Subroutine Package*) fornecido pela IBM e que

inclui o programa FACTOR destinado ao processamento da análise fatorial; o *BMD* (*Biomedical Computer Programs*) que inclui para esse tipo de análise os programas BMDO3M e BMDX72; e, ainda, o *SPSS* (*Statistical Package for the Social Sciences*) que inclui o programa FACTOR com o mesmo objetivo. A utilidade dos programas mencionados é variável, por isso, foi elaborada a Tabela 1, onde para facilidade dos interessados, são comparadas as características de cada um.

TABELA 1

Comparação entre os diversos programas de computador para análise fatorial disponíveis na Universidade de São Paulo.

	FACTOR	BMDO3M	BMDX72	FACTOR
Nº máximo de variáveis	35	80	198	99
Nº máximo de amostras	99 999	9.999		
Entrada por matrizes	não	matrizes de coeficientes de similaridade e fatoriais		
Matrizes de similaridade calculada pelo programa	correlação	correlação	correlação covariância	correlação covariância
Processos de extração de fatores	componentes principais	componentes principais	componentes principais	componentes principais Fat. canônica de Rao Fat. Imagem Fat. Alpha
Métodos de rotação	ortogonal —Varimax	ortogonal —Varimax	ortogonal —Varimax —quartimax oblíqua —quartimin —covarimin	ortogonal —Varimax —quartimax —equimax oblíqua (com controle de ângulo)
Cálculo de coeficientes fatoriais	não	sim	sim	sim
Plotagem de pesos fatoriais	não	não	não	sim

ANÁLISE FATORIAL APLICADA AO ESTUDO GEOQUÍMICO DE ROCHAS BASÁLTICAS

Manson, *op. cit.*, por meio de levantamento sistemático na literatura, reuniu grande número de dados analíticos referentes a elementos principais dosados em rochas basálticas de distribuição mundial - com exceção da América do Sul. Selecionou, por meio de vários critérios, dentre os dados reunidos, 1996 análises consideradas como representativas de todos os tipos de rochas basálticas, excluindo-se tipos diferenciados. Constituiu, dessa forma, elenco analítico de referência que corresponde, essencialmente, a caracteres geoquímicos de origem, supostamente não afetada

por rochas basálticas que representam estágios mais avançados de evolução magmática. Utilizando a técnica da análise factorial, reduziu o conjunto mencionado de observações multivariáveis, a 4 fatores, de acordo com a matriz factorial reunida na Tabela 2. Esses fatores são suficientes para explicar 99% da variabilidade existente na amostra estatística e, como se observa na matriz factorial, cada um deles é respectivamente dominado pelas seguintes variáveis: Al_2O_3 (SiO_2), MgO (CaO), Fe_2O_3 (SiO_2) e SiO_2 (Fe_2O_3 , CaO).

TABELA 2

	FATOR			
	1	2	3	4
SiO_2	.471	.379	.459	-.627
TiO_2	-.086	.059	.185	.115
Al_2O_3	.771	-.035	-.077	.373
Fe_2O_3	-.235	.275	.793	.550
MnO	-.001	.002	.008	.004
MgO	-.246	.735	-.373	.142
CaO	.118	.518	-.016	.521
Na_2O	.076	-.018	.035	.014
K_2O	.017	-.007	.003	-.031

Matriz de estimativa de fatores para rochas basálticas de distribuição mundial, segundo Manson, 1967. Os valores correspondentes aos vários óxidos representam o significado proporcional de cada um deles na determinação da composição dos fatores apropriados.

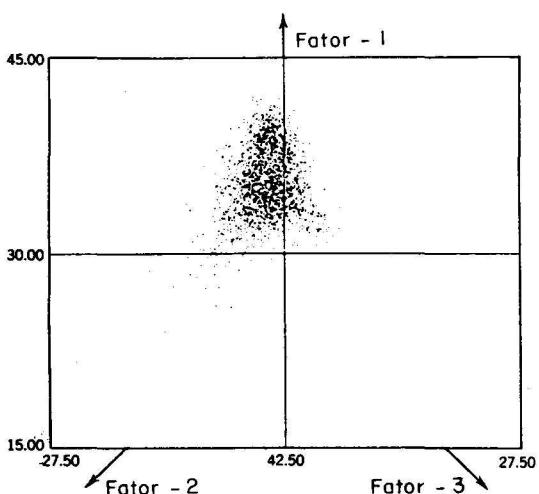


Fig. 2: Projeção em diagrama ternário das principais variações químicas de rochas basálticas. A área focalizada corresponde a detalhe de triângulo maior ao qual pode ser referido pelos parâmetros apostos na figura. Os pontos projetados correspondem à composição de 1996 análises de rochas basálticas expressas em termos das proporções dos 3 primeiros fatores, dos 4 a que foram reduzidas as variáveis analíticas conforme a Tabela 2. Diagrama segundo Manson, 1967.

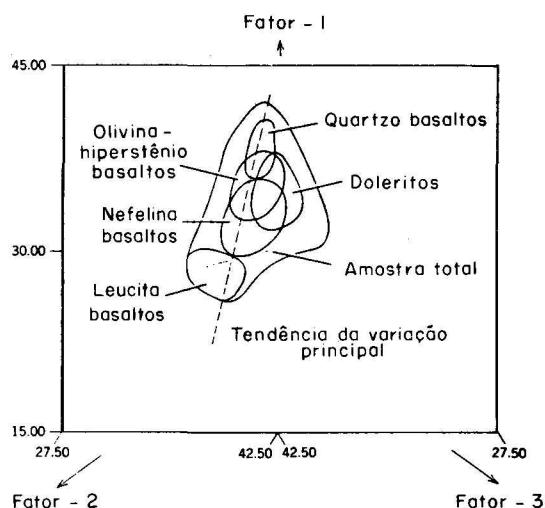


Fig. 3: Projeção em diagrama ternário das principais variações químicas de rochas basálticas. Corresponde à mesma área detalhada na Figura 2, porém indicadas as áreas de distribuição máxima para os diversos tipos de rochas basálticas conforme indicado. A linha tracejada representa a tendência principal de variação química entre rochas basálticas. Diagrama segundo Manson, 1967.

A representação gráfica dos resultados obtidos por Manson, *op. cit.*, para as rochas basálticas de distribuição mundial, está reproduzida nas Figuras 2 e 3, por meio de diagramas ternários onde foram projetados 3 dos 4 fatores determinados. As Figuras 2 e 3 são auto-explicativas e não necessitam maiores comentários.

A ANÁLISE FATORIAL APLICADA AO ESTUDO GEOQUÍMICO DAS ROCHAS BASÁLTICAS DA BACIA DO PARANÁ

A técnica da análise factorial foi aplicada na forma proposta por Manson, *op. cit.*, ao estudo de amostra estatística de análises químicas de rochas basálticas da Bacia do Paraná. A referida amostra é constituída por 197 análises químicas de elementos principais. Cerca de 90 dessas análises foram reunidas por Rüegg, 1969, da literatura e de dados então inéditos. As demais constituem parte de conjunto ainda maior que visa retratar aspectos geoquímicos das rochas da província, e cujo estudo pormenorizado será apresentado em outro contexto (Rüegg, em preparação).

As análises químicas ora utilizadas correspondem indistintamente aos mais variados tipos de rochas basálticas que ocorrem associadas na Bacia do Paraná. As amostras analisadas são de ampla distribuição geográfica e foram coletadas ao longo de toda a área de ocorrência. Os tipos petrográficos são variados e as amostras foram coletadas em derrames e corpos intrusivos rasos, *sills*, *sheets* e diques. Petrologicamente constituem termos de série toleítica, diagnóstico este realizado quer na mineralogia modal, quer na mineralogia normativa ou ainda utilizando-se parâmetros geoquímicos apropriados. Constituem, em geral, termos supersaturados dessa série onde basaltos, senso estrito, são raros, predominando termos petrográficos em vários estágios de evolução como demonstraram Leterrier, De La Roche e Rüegg, 1972. Parte das novas análises foram realizadas em amostras geocronologicamente controladas e, em conjunto, não mostram qualquer peculiaridade adicional, em termos de idade, às já descritas nos trabalhos de Amaral *et. alii.*, 1969, Melfi, 1967 e Miniolli *et. alii.*, 1960, nos quais foram originalmente descritas.

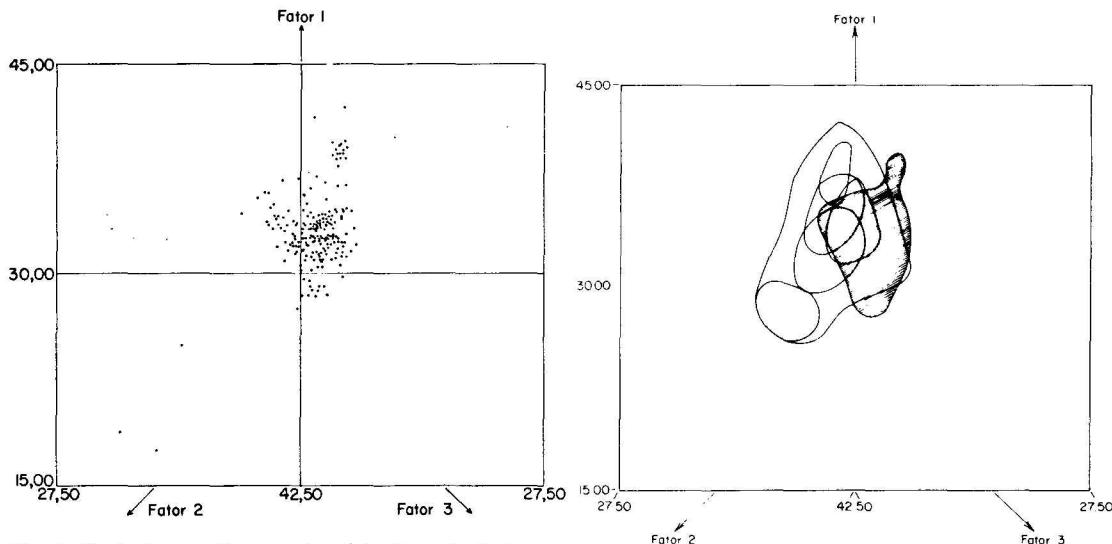


Fig. 4: Projeção em diagrama ternário das principais variações químicas determinadas para as rochas basálticas da bacia do Paraná, utilizando-se os mesmos referenciais da Figura 2. Os pontos projetados correspondem à composição de 197 análises químicas expressas na figura em termos das proporções dos 3 primeiros fatores dos 4 a que foram reduzidas as variáveis analíticas de acordo com os valores da matriz de estimação de fatores apresentada por Manson, 1967 e reproduzida na Tabela 2.

Fig. 5: Área ocupada pelas rochas da Bacia do Paraná em relação aos principais tipos de rochas basálticas conforme focalizado na Figura 3.

A análise fatorial foi realizada com os dados químicos acima referidos, aplicando-se a matriz de estimação de fatores determinada por Manson, *op. cit.* (Tab. 2). Os cálculos necessários foram conduzidos por meio de pequeno programa (COEFAT) para computador IBM 360/44. Os resultados obtidos encontram-se ilustrados na Figura 4, que consiste em diagrama ternário dos 3 primeiros coeficientes fatoriais. Dessa forma são diretamente comparáveis aos dados de referência.

Dessa comparação, salientada na Figura 5, verifica-se que as análises correspondentes às rochas da Bacia do Paraná, ocupam, generalizadamente, posição distinta em relação ao arcabouço de referência, elaborado para os principais tipos de rochas basálticas. A posição ocupada por elas, em relação aos tipos de referência, caracteriza-se por deslocamento para posições de maior importância para o fator 3, dominado como se verifica na Tabela 2, pelo ferro e, subsidiariamente, pela sílica. Confirma-se, dessa forma, mais uma vez, a já conhecida feição das rochas basálticas da Bacia do Paraná que consiste na sua extrema riqueza em ferro. Aos altos teores em ferro, correspondem baixos teores de magnésio e cálcio, óxidos que dominam o fator 2. Esta relação evidencia que as rochas basálticas da província se apresentam de forma geral marcadas por um ou mais estágios de diferenciação em profundidade, antes de sua emissão na superfície. Estas circunstâncias justificam sua posição destacada na Figura 5, em relação aos tipos de referência, pois estes não incluem rochas diferenciadas.

Embora, segundo Rüegg, *op. cit.*, as rochas basálticas da província constituam termos toleíticos supersaturados, ou os chamados quartzo-basaltos e/ou diabásios, na acepção conferida por esses termos por Yoder e Tilley, 1962, elas não coincidem na Figura 5 com a área de referência ocupada por essas rochas. Aliás, dela se distanciam quase em igual medida das rochas basálticas altamente insaturadas, os leucita-basaltos. Em contrapartida considerável superposição se observa com as áreas correspondentes aos doleritos, aos hiperstênio basaltos (olivina tholeítos, *apud* Yoder e Tilley, *op. cit.*) e ainda as correspondentes aos nefelina basaltos (basaltos alcalinos). Se a superposição com os doleritos é normal, uma vez que a amostra estatística analisada inclui muitos exemplares do gênero, o mesmo não ocorre com os demais tipos. A mineralogia, quer modal, quer normativa, não mantém qualquer afinidade com esses tipos de rochas, a não ser pontos de contato numericamente inexpressivos com os olivina toleítos. A superposição observada entre as rochas supersaturadas da Bacia do Paraná e rochas insaturadas de referência provavelmente se deve à

TABELA 3

Matriz obtida pela análise fatorial de modo R de dados analíticos referentes aos elementos principais dosados em rochas basálticas da Bacia do Paraná

	FATOR			
	1	2	3	4
SiO ₂	—. 743	—. 234	—. 793	. 062
TiO ₂	—. 009	—. 029	. 177	—. 021
Al ₂ O ₃	. 011	. 357	—. 053	—. 004
Fe ₂ O ₃	. 236	—. 070	. 463	. 209
MnO	. 002	—. 001	. 005	. 001
MgO	. 302	. 002	. 069	. 042
CaO	. 404	. 007	. 152	—. 027
Na ₂ O	—. 074	. 001	. 005	—. 031
K ₂ O	—. 153	—. 006	—. 062	—. 032

intersecção de caracteres de origem das últimas com caracteres de diferenciação das primeiras, conforme mecanismo descrito por Leterrier, *et alii*, 1972.

ANÁLISE FATORIAL INDEPENDENTE DAS ROCHAS DA BACIA DO PARANÁ

Além da comparação dos dados relacionados às rochas da Bacia do Paraná com os esquemas de referência de rochas equivalentes de distribuição mundial, fornecida por Manson, *op. cit.*, foi efetuada com os mesmos dados, análise fatorial independente.

Para isto as análises químicas não reduzidas de rochas da Bacia do Paraná foram submetidas a diversos processos de análise fatorial, empregando-se os programas de computador BMDX72 e FACTOR, já citados. Os resultados obtidos são de modo geral, concordantes com os apresentados por Manson, *op. cit.* Deste trabalho, são apresentadas apenas duas matrizes fatoriais, consideradas as mais representativas dos resultados. Os leitores interessados em mais pormenores poderão obtê-los diretamente dos autores.

A Tabela 3 reúne os resultados da análise fatorial de modo R (procurando inicialmente relações entre variáveis). Verifica-se na composição dos fatores a participação preponderante dos mesmos elementos que controlam a matriz fatorial proposta por Manson, *op. cit.*, e reproduzida na Tabela 2. Neste caso, os 4 fatores são capazes de descrever 93% da variabilidade apresentada pelas variáveis.

A Tabela 4 reúne os resultados obtidos pela análise fatorial de modo Q (procurando inicialmente relações entre as amostras). Novamente aparecem nas suas composições os mesmos óxidos predominantes. Esta matriz descreve 99% da variabilidade das amostras. Em ambos os casos é relevante a importância do ferro, confirmado, por caminho diferente, as considerações finais realizadas no capítulo precedente.

TABELA 4

Matriz obtida pela análise fatorial de modo Q de dados analíticos referentes aos elementos principais dos dados em rochas basálticas da Bacia do Paraná

	FATOR			
	1	2	3	4
SiO ₂	— 1.229	2.079	1.048	0.371
TiO ₂	0.769	— 0.296	0.540	— 0.189
Al ₂ O ₃	— 0.697	0.344	— 1.695	— 1.894
Fe ₂ O ₃	— 0.789	— 1.512	1.609	— 0.968
MnO	0.993	0.061	— 0.093	0.226
MgO	— 0.350	— 0.472	— 0.807	1.368
CaO	— 0.980	— 0.842	— 0.693	1.171
Na ₂ O	0.966	0.190	0.026	— 0.065
K ₂ O	1.326	0.448	0.064	— 0.022

BIBLIOGRAFIA

- AMARAL, G.; CORDANI, U. G.; KAWASHITA, K; REYNOLDS, J. H. - 1966 - Potassium-argon dates of basaltic rocks from southern Brazil. *Geochimica et Cosmochimica Acta*, Oxford, 30:169-89.
- IMBRIE, J. & PURDY, E. G. - 1962 - Classification of modern Bahamian carbonate sediments. *Memoir of the American Association of Petroleum Geologists*, Tulsa, Okla, 1:253-72.
- ____ & VAN ANDEL, T. H. - 1964 - Vector analysis of heavy mineral data *Bulletin of the Geological Society of America*, Rochester, N. Y., 75:1131-56.
- HARMAN, H. H. - 1960 - *Modern factor analysis*. Chicago, University of Chicago Press, 417p.
- LAFITTE, P. - 1972 - *Traité d'informatique géologique*. Paris, Masson
- LETERRIER, J.; DE LA ROCHE, H.; RÜEGG, N. R. - 1972 - Composition - chimique et parenté tholeïtique des roches basaltiques du Bassin du Paraná. *Compte Rendu de L'Academie des Sciences*, Paris, 274:1772-5.
- MANSON, V. - 1967 - Geochemistry of basaltic rocks: major elements" In: HESS, H. H. - *Basalts; the Poldervaart treatise on rocks of basaltic composition*. New York, Interscience. V.1, p.215-69.
- MCINTYRE, D. B. - 1969 - "Introduction to the study of data matrices" In: FERNNER, P. - ed. - *Models of geologic processes*. Washington, American Geological Institute.
- MELFI, A. J. - 1967 - Potassium-argon ages for core samples of basaltic rocks from southern Brazil. *Geochimica et Cosmochimica Acta*, Oxford, 31:1079-89.
- MERTIE, J. B. - 1961 - Analytical classification and quadriplanar chartin for analysis with mine or more components. *American Mineralogist*, Lancaster, Pa., 46:613-28.
- MINIOLLI, B.; PONLANO, W. L.; OLIVEIRA, S. M. B. - 1969 - "Extensão geográfica do vulcanismo basáltico do Brasil meridional". In: CONGRESSO BRASILEIRO DE GEOLOGIA, 23º, Salvador: Anais. Salvador, Sociedade Brasileira de Geologia.
- PRINZ, M. - 1967 - "Geochemistry of basaltic rocks: trace elements". In: HESS, H. H. - *Basalts, the Poldervaart treatise on rocks of basaltic composition*, New York, Interscience, v.1, p.271-323.
- RÜEGG, N. R. - 1969 - *Aspectos geoquímicos, mineralógicos e petrográficos da rochas rochas basálticas da Bacia do Parnaíba*, 178P: Tese (Dout.) - Faculdade de Filosofia Ciências e Letras da USP, São Paulo.
- YODER, N. S. & TILLEY, C. E. - 1962 - Origin of basaltic magmas: an experimental study of natural synthetic rock systems. *Journal of Petrology*, Oxford, 3:342-532.