

# ANÁLISE DE AGRUPAMENTO EM AQUIFERO FRATURADO

Mário L.D. Wrege \*  
Gilberto Mobus \*

## ABSTRACT

This paper presents the analysis of a fractured aquifer through cluster analysis. The area is in northwestern Rio Grande do Sul, Brazil. The statistical analysis, based on the geological data, aided on the understanding of the fractured medium as a hydrogeological aquifer.

## INTRODUÇÃO

Na geologia é comum ter-se grandes quantidades de dados para fundamentar um determinado trabalho ou estudo. Trabalhar estes dados, de modo a tirar o máximo de informação, nem sempre é uma tarefa fácil. O número elevado de dados, por si só, já se constitui num problema. Também o fato de ter-se, muitas vezes, dados medidos em unidades diferentes (diversos parâmetros observados), e com intervalos de variação distintos, criam uma dificuldade a mais na visualização de possíveis relações existentes entre as amostras a serem trabalhadas.

Quando medidas são feitas, envolvendo diversos parâmetros, um número elevado de amostras, a tabela de dados resultantes é grande e de difícil interpretação. Por vezes é desejável compartimentar estas amostras em grupos, em função da similaridade dos parâmetros observados; seja em termos quantitativos, seja de comportamento. Ainda, é desejável que a base desse agrupamento seja objetiva (impessoal) e que se possa obter uma medida da qualidade desse agrupamento, qualquer que seja o critério empregado.

Análise de agrupamento (*Cluster Analysis*) é uma das técnicas empregadas no processo de classificação de amostras caracterizadas por duas ou mais variáveis. Ela foi desenvolvida com base em conceitos estatísticos.

Dados de caráter hidrogeológico da Formação Serra Geral, da região noroeste do Estado do Rio Grande do Sul, foram submetidos a análise de agrupamento, visando tanto fornecer subsídios ao conhecimento do aquífero fraturado, como testar a eficiência e operacionalidade da técnica empregada.

## FUNDAMENTOS TEÓRICOS

Quando se deseja tratar estatisticamente amostras onde foram feitas duas ou mais medidas, recorre-se à análises estatísticas multivariantes, pois estas permitem considerar as variações em diversas propriedades simultaneamente, ou seja, fornecem informações na razão direta do número de variáveis utilizadas, considerando, ao mesmo tempo, a variabilidade nas diversas propriedades medidas e a relação entre estas variáveis (SILVA, 1984)

A análise de agrupamento é uma das técnicas multivariantes empregadas para analisar a similaridade entre indivíduos (amostras) caracterizados por variáveis. Por este método são definidos grupos de indivíduos com determinado grau de similaridade, sendo considerado simultaneamente todas as variáveis de cada indivíduo (DAVIS, 1973). Por agrupamento de grupos de itens, procura-se um número conveniente de grupos, relacionados através de coeficientes de similaridade (NETO & CAMPOS, 1982).

Para análise de agrupamento, parte-se de uma matriz inicialmente simétrica, constituída por coeficientes de associação (similaridade) entre itens, gerados a partir das vari-

---

\* Setor de Águas Subterrâneas do Instituto de Pesquisas Hidráulicas Universidade Federal do Rio Grande do Sul

áveis normalizadas. Em seguida, grupos iniciais de semelhança (com maiores coeficientes de similaridade) são encontrados, os quais admitirão, posteriormente, novos membros, em ordem sucessiva no sentido de decréscimo de similaridade (agrupamento hierárquico), até que todos os indivíduos estejam corretamente correlacionados, ou seja, corretamente agrupados.

Embora várias medidas de similaridade tenham sido propostas nas literaturas especializadas (SOKAL & MICHENER, 1958; IMBRIE & PURDY, 1962; SOKAL, 1961; DAVIS, 1974; etc.), somente duas geralmente são usadas. São elas: o coeficiente de correlação linear do produto momento (Pearson), que procura detectar a similaridade existente entre as variáveis (correlação entre elas - modo R), e o coeficiente de distância, usado para verificar a similaridade entre indivíduos (amostras semelhantes - modo Q).

O coeficiente de distância expressa o grau de similaridade como distância, baseando-se na forma generalizada de Pitágoras. Quanto maior a distância, menor a similaridade, e vice-versa. O coeficiente de distância é calculado por:

$$D_{ij} = \sqrt{\frac{\sum_{k=1}^m (X_{ik} - X_{jk})^2}{m}}$$

onde:  $D_{ij}$  = coeficiente de distância entre as amostras;  
 $X_{ik}$  = k-ésima variável na amostra i;  
 $X_{jk}$  = k-ésima variável medida na amostra j;  
 $m$  = número de variáveis

A forma usual para a representação gráfica mostrando o relacionamento das amostras é o dendrograma (ou diagrama ramificado), o qual projeta, num espaço bidimensional, o agrupamento hierárquico.

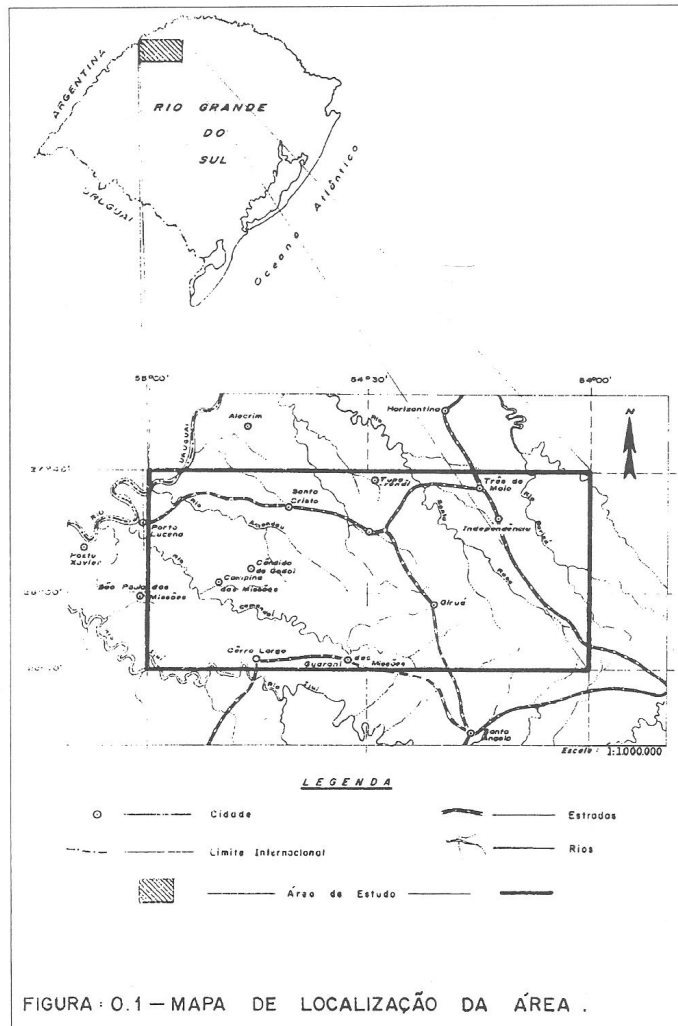


FIGURA : 0.1 - MAPA DE LOCALIZAÇÃO DA ÁREA .

## ANÁLISE DO AQUIFERO

Na Análise Estrutural e Hidrogeológica do Aquífero Fraturado da Formação Serra Geral - Região Noroeste do Rio Grande do Sul (MOBUS, 1987) foi empregado a técnica de análise de agrupamento aos dados (físicos, químicos e hidrodinâmicos) de poços tubulares, como ferramenta auxiliar na caracterização hidrogeológica da região em questão.

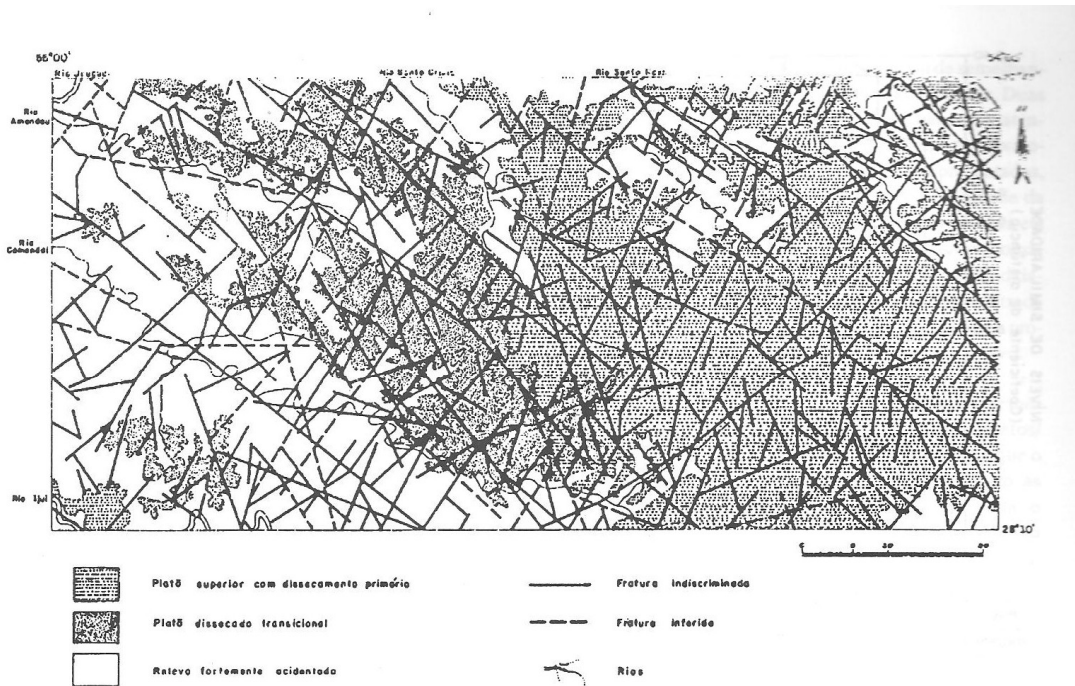


FIGURA: 0.2 — MAPA MORFO-ESTRUTURAL DA ÁREA .

A área estudada situa-se a noroeste do estado do Rio Grande do Sul, entre os paralelos 27°45' e 28°10' de latitude sul e os meridianos 55°00' e 54°00' de longitude oeste (Figura 01).

Litologicamente, a área apresenta-se totalmente coberta pelas rochas efusivas da Formação Serra Geral. Os aluviões quase inexistem e o elúvio, em geral, não são muito espessos. Morfológicamente, constitui-se em uma unidade tipo planalto (Figura 02).

Foram analisados sessenta e oito poços, sendo cada poço caracterizado por sete variáveis. As variáveis selecionadas foram:

- \* número de entradas d'água ..... NEA
- \* profundidade da última entrada d'água (m) .. PUEA
- \* descarga específica (m<sup>3</sup>/dia/m) ..... q
- \* concentração de íons hidrogênio na água ..... pH
- \* indicador geoquímico (relação Cl/CO<sub>3</sub>H) ..... IG
- \* sólidos totais (mg/l) ..... ST
- \* espessura do elúvio (m) ..... E

Os dados de caráter hidrodinâmico (q) e os representativos dos fatores hidrofisicoquímicos (pH, IG e ST) foram inicialmente submetidos a uma análise de consistência, eliminando-se valores anômalos, ora devidos a erro de análise, ora por problema de coletas.

Os valores das variáveis foram padronizados de maneira a ficarem compreendidos no intervalo de 1 a 2, isto visou o equilíbrio de peso que cada variável teria dentro do procedimento da análise.

Como objetivo, dentro do escopo do trabalho, era de verificar semelhança entre as amostras (poços - modo Q), empregou-se o coeficiente de distância. Para o processamento dos dados pela análise geomatemática multivariante foi empregado o programa CLUSTER, desenvolvido por DAVIS (1973). Este programa está em linguagem FORTRAN IV, adaptável para microcomputadores da linha PC.

A Figura 03 mostra o dendrograma resultante da análise de agrupamento. Pode-se notar que, ao nível de similaridade de 0.3, os poços agregam-se em cinco grupos básicos. A Tabela 01 apresenta os valores característicos das variáveis em cada grupo.

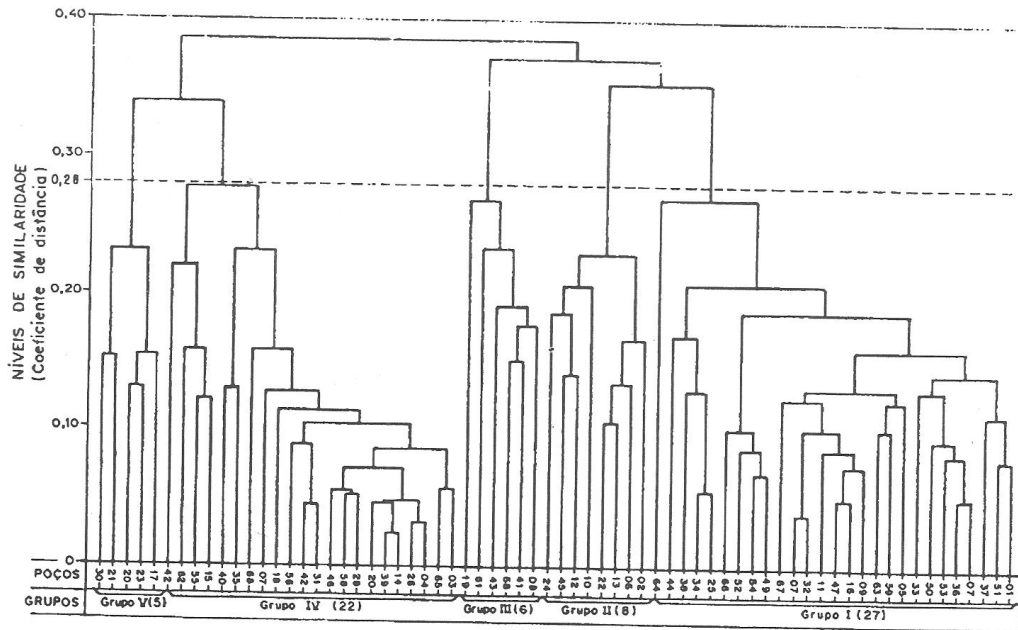


FIGURA 0.3 – DENDROGRAMA DE AMOSTRAS COM DADOS HIDROGEOLÓGICOS

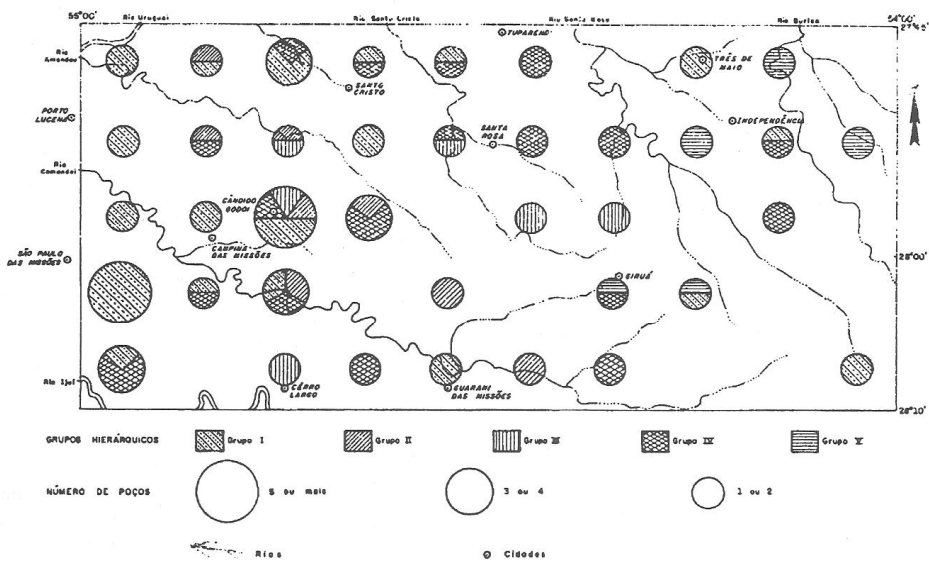


FIGURA 0.4 – MAPA ESQUEMÁTICO DA DISTRIBUIÇÃO ESPACIAL DOS GRUPOS HIERÁRQUICOS.

ESTATÍSTICAS							
VARIÁVEIS		GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5	TOTAL
Nº de amostras		27	8	6	22	5	68
NEA	media	2	2	3	3	3	2
	s	0.7	0.5	1.5	1.3	1.4	1.0
	mín.	1	1	2	2	2	1
	máx.	3	3	5	6	5	6
PUEA	media	51.1	145.7	60.2	58.2	52.0	66.5
	s	24.8	34.4	40.1	44.9	25.8	45.4
	mín.	13.0	103.0	30.0	22.0	29.0	13.0
	máx.	102.0	202.0	132.0	165.0	94.0	202.0
q	media	27.7	7.0	49.5	20.3	20.5	37.2
	s	35.6	4.6	43.5	17.9	10.6	63.4
	mín.	1.8	0.8	9.7	0.6	5.7	0.6
	máx.	120.0	14.0	131.9	79.6	33.6	131.9
pH	media	7.6	9.2	8.0	7.2	7.1	7.6
	s	0.5	0.7	1.0	0.8	0.5	0.9
	mín.	6.4	8.2	7.0	6.2	6.6	6.2
	máx.	8.7	10.0	9.7	9.2	7.9	10.0
IG	media	0.031	0.030	0.085	0.043	0.059	0.04
	s	0.022	0.020	0.043	0.037	0.039	0.04
	mín.	0.010	0.015	0.037	0.008	0.035	0.00
	máx.	0.108	0.075	0.160	0.172	0.126	0.17
ST	media	222.6	382.1	187.7	154.2	99.8	213.8
	s	68.4	98.0	77.5	45.9	24.2	113.0
	mín.	105.5	236.1	102.8	66.8	73.0	66.8
	máx.	378.7	527.5	312.7	242.0	125.0	527.5
E	media	5.3	4.8	6.1	7.6	28.0	7.5
	s	2.2	3.3	6.1	4.4	5.0	6.8
	mín.	0.3	1.4	0.8	3.0	24.0	0.3
	máx.	9.0	12.0	18.0	18.0	36.0	36.0

NEA = Número de Entradas de Água.  
PUEA = Profundidade da Última Entrada de Água.  
q = Descarga Específica (m<sup>3</sup>/d/m)  
pH = Concentração de íons hidrogênio em solução  
IG = Indicador Geoquímico (Cl/CO<sub>3</sub>H)  
ST = Sólidos Totais (mg/l)  
E = Espessura do Manto de Alteração (m)

TABELA 01 - Valores Característicos das Variáveis em cada Grupo

A Figura 04 mostra, de maneira esquemática, a distribuição espacial de cada elemento dos grupos obtidos na análise de agrupamento. Pode-se observar nesta figura que os poços que constituem o grupo I encontram-se quase que totalmente concentrados na porção ocidental da área, próximo ao Rio Uruguai, área esta que apresenta um relevo fortemente dissecado. Este grupo foi o que apresentou o maior número de amostras, o maior coeficiente de similaridade (último nível de agrupamento) foi de 0,27, o que indica

um bom grau de semelhança entre as amostras. Duas variáveis chamam a atenção neste grupo pelos valores médios apresentados, são elas: profundidade da última entrada de água e a relação cloretos/bicarbonatos. A pequena profundidade média da última entrada d'água justifica-se pela própria condição do relevo.

Já o baixo valor médio da relação cloretos/bicarbonatos denota que o aporte de bicarbonato às águas nos poços é alta, o que pode ser um indicativo de que a recarga se de em outras áreas do que as ocupadas pelos poços desse grupo. Também é neste grupo que se encontram os menores valores da variável espessura do elúvio, nunca superior a dez metros, o que reforça o baixo potencial de recarga no local.

O grupo II, constituído por oito amostras, apresentou um nítido padrão de ocorrência, pois os poços encontram-se alinhados segundo uma estreita faixa paralela à diagonal noroeste-sudeste da porção ocidental da área, faixa esta caracterizada por ser uma área de transição entre o platô basáltico à leste e a

porção menos elevada à oeste. O índice de similaridade do último nível de agrupamento é de 0,23 o que confere um alto grau de similaridade a esses poços. Duas variáveis se destacam pelos seus altos valores médios: a profundidade da última entrada de água e o pH. Isto induz, de maneira subjetiva, a crer na possibilidade de haver influência da profundidade das entradas de água sobre o pH da água.

O grupo III é composto por seis amostras, agrupadas

à um nível de similaridade de 0,27, dispostas sempre próximas à porção central da área pesquisada. Este grupo apresenta o maior valor médio da descarga específica, apesar de que, neste grupo, esta variável apresenta um alto índice de variação (C = 88%). Neste grupo também encontra-se o maior valor médio da relação cloretos/bicarbonatos, o que pode ser um indicativo de ser esta a zona preferencial de recarga.

Já o grupo IV é o segundo em tamanho, reunindo 22 amostras (poços), agrupados à um nível de 0,28. Estes poços apresentam-se bem dispersos pela área da pesquisa, não apresentando claramente uma zona de concentração específica. Os termos médios das variáveis desse grupo apresentam valores intermediários em relação aos outros grupos. Destacam-se valores individuais em relação aos outros grupos. Destacam-se valores individuais, pois neste grupo encontra-se o poço com a menor descarga específica e menor pH, assim como o poço com o menor teor de sólidos totais analisado e a maior razão de cloretos por bicarbonatos.

O grupo V é constituído por somente cinco poços agrupados a um nível de similaridade de 0,23. Estes poços estão localizados na porção oriental da área, que apresenta predominantemente um relevo de platô dissecado transicional. Os poços desse grupo mostraram valores característicos das variáveis que se destacam na amostragem total. É o caso do valor médio da espessura do elúvio que, comparado aos outros grupos, é muito superior. As variáveis físico-químicas pH e Sólidos Totais apresentaram, neste grupo, os menores valores médios.

## CONSIDERAÇÕES FINAIS

A análise de agrupamento, aplicável no tratamento de variáveis ou de amostras, permitem uma compartimentação de itens de forma objetiva. Como o agrupamento hierárquico é convenientemente graficado na forma de um diagrama (dendrograma), é possível acompanhar os sucessivos níveis de similaridade do processo, permitindo, assim, a estimativa do grau de similaridade de determinação do número de grupos.

Esta técnica, assim como a maioria das análises estatísticas multivariantes, apresenta limitações na sua aplicabilidade. Dentre estas, pode-se mencionar o número disponível de dados sobre o qual se aplica a análise. Um número de amostras menor que o número de variáveis (parâmetros) compromete a eficácia da análise (Ex: grupos III e V do exemplo exposto). A aplicação dessa técnica neste caso só é permissível quando houver justificativas, mesmo que subjetivas, para a sua utilização.

Um cuidado a ser tomado quando da aplicação da análise de agrupamento sobre amostras, é verificar se as mesmas seguem uma distribuição normal pois, caso contrário, faz-se necessário uma transformação adequada das variáveis.

A operacionalidade dessa técnica geomatemática é grande, uma vez que a mesma acha-se decodificada em programas para computadores de pequeno e médio porte. O tempo de processamento é pequeno e os resultados dispostos, ora na forma de tabelas ou matrizes, ora na forma de gráficos ou mapas, facilitam a interpretação dos resultados.

Em relação ao caso estudado, foi possível obter uma compartimentação objetiva dos poços perfurados no aquífero fraturado da área. A distribuição espacial dos grupos propostos mostrou haver fatores subjetivos (geologia estrutural) que reforçaram tal compartimentação, exceção feita a somente um grupo (grupo IV).

## REFERÊNCIAS BIBLIOGRÁFICAS

1. DAVIS, J.C. - 1973. Statistics and data analysis in geology. New York, John Wiley. 550p.
2. IMBRIE, J. - 1963. Factor and vector analysis programs for analysing geologic data. Tech. Rept. 6, ONR Task nº 389-135.
3. MÓBUS, G. 1987. Análise estrutural e hidrogeológica do aquífero fraturado da Formação Serra Geral - Região noroeste do Rio Grande do Sul. Porto Alegre - Universidade Federal do Rio Grande do Sul Curso de Pós-Graduação em Engenharia Civil. 124f. Dissertação de Mestrado.
4. NETO, A.F.C.; & CAMPOS, H.C.S. - 1980. Análise de Agrupamento - um método auxiliar na caracterização de aquíferos. In: CONGRESSO BRASILEIRO DE GEOLOGIA, Balneário de Camboriú, Santa Catarina, 1980. V.2, p.1041/55.
5. PARKS, J.M. - 1966. Cluster Analysis applied to multivariate geologic problems. Journal of geology, Chicago, 74(5): 703-15, Sep.
6. SILVA, A.B. DA. - 1984. Análise morfoestrutural, hidrogeológica e hidroquímica no estudo do aquífero cárstico do Jaíba, norte de Minas Gerais. São Paulo, Instituto de Geociências da USP. 190p. Tese.
7. SOKAL, R.R.; & MICHENER, C.D. - 1958. A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull., v.38, pt.2, nº 22, p.1409-1438.