



ÂNGELA CRISTINA DA FONSECA MIRANTE

**GEOESTATÍSTICA E ANÁLISE DE
AGRUPAMENTO NO ESTUDO DA
MICROFAUNA DE FORAMINÍFEROS**

LAVRAS - MG

2013

ÂNGELA CRISTINA DA FONSECA MIRANTE

**GEOESTATÍSTICA E ANÁLISE DE AGRUPAMENTO NO ESTUDO DA
MICROFAUNA DE FORAMINÍFEROS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

Orientador

Dr. João Domingos Scalon

Coorientadora

Dra. Tânia Maria Fonseca Araújo

LAVRAS – MG

2013

**Ficha Catalográfica Elaborada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Mirante, Ângela Cristina da Fonseca.

Geoestatística e análise de agrupamento no estudo da microfauna
de foraminíferos / Ângela Cristina da Fonseca Mirante. – Lavras :
UFLA, 2013.

153 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2013.

Orientador: João Domingos Scalon.

Bibliografia.

1. Semivariância. 2. Análise multivariada. 3. Método de Ward.
4. Método *k*-Médias. 5. Krigagem. I. Universidade Federal de
Lavras. II. Título.

CDD – 519.535

ÂNGELA CRISTINA DA FONSECA MIRANTE

**GEOESTATÍSTICA E ANÁLISE DE AGRUPAMENTO NO ESTUDO DA
MICROFAUNA DE FORAMINÍFEROS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

APROVADA em 13 de junho de 2013.

Dra. Tânia Maria Fonseca Araújo	UFBA
Dr. Marcelo Tavares	UFU
Dra. Helisângela Acris Borges de Araújo	FATEC-BA
Dr. Renato Ribeiro de Lima	UFLA
Dr. Agostinho Roberto de Abreu	UFLA

Dr. João Domingos Scalon
Orientador

LAVRAS – MG

2013

Aos meus pais, José Antônio (*in memoriam*) e Dinalva, esposo, José Albino, filhos Ricardo e Leila e neto Rodrigo, pelo amor e alegrias.

DEDICO

AGRADECIMENTOS

A Deus, força presente em todos os momentos da minha vida;

À Universidade Federal de Lavras (UFLA), ao Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela oportunidade de concretizar esse trabalho;

Ao professor Dr. João Domingos Scalon, pela orientação, competência, apoio, incentivo e amizade;

A professora Dra. Tânia Maria Fonseca Araújo pela coorientação, competência, sugestões, apoio e carinho; e por ter disponibilizado os dados utilizados neste trabalho;

Aos professores do Departamento de Ciências Exatas da Universidade Federal de Lavras (DEX-UFLA), pela competência, ensinamentos e comprometimento;

Ao professor Dr. Enio Seidel Júnior, pelo apoio e sugestões;

Ao professor Dr. Lurimar Smera Batista, do Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA), pelo apoio e comprometimento;

Aos colegas e amigos da Pós-Graduação: Azly, Cleide, Edmary, Jailson, Jaime, Jose Otaviano (*in memoriam*), Luiz Vásquez, Marcio, Nelson, Norma e Regilson, pelo companheirismo e convivência durante todo o curso, principalmente nos momentos de estudo;

Em especial, aos amigos Tânia Jussara, Isabel e Walter, pelo apoio, convívio, solidariedade e incentivo em todos os momentos; e a minha gratidão à Isabel, pelas horas de estudo em que compartilhou seus conhecimentos com competência e dedicação;

Aos meus pais; José Antônio (*in memoriam*) e Dinalva pelo amor, compreensão e educação; as minhas irmãs; Tânia Fonseca Araújo e Nádia Fonseca Vita Souza e meus cunhados; Jair Araújo e Antônio Luiz Vita Souza, pelo apoio, incentivo e carinho;

Aos meus filhos; Ricardo e Leila, minha nora Graciete e meu neto Rodrigo, pelo afeto, confiança, incentivo e momentos descontraídos e alegres;

Ao meu esposo, amigo e companheiro, José Albino, pelo amor, compreensão, apoio, estímulo e confiança em todos esses anos;

Meus sinceros agradecimentos a todos que de alguma forma contribuíram para elaboração desse trabalho.

RESUMO GERAL

Objetivou-se com este trabalho aplicar a geoestatística e a análise de agrupamento nos dados da microfauna de foraminíferos presentes nos sedimentos da plataforma e talude continentais do norte do Estado da Bahia. A análise geoestatística foi aplicada para avaliar a presença de variabilidade espacial no número de indivíduos das espécies de foraminíferos e realizar predições por meio de krigagem. Para obter os semivariogramas experimentais das variáveis em estudo foram utilizados os estimadores de Matheron e de Cressie & Hawkins. Os resultados mostraram que as espécies *Amphistegina lessonii*, *Peneroplis carinatus*, *Quinqueloculina bicarinata* e *Reusella spinulosa* apresentaram estruturas de dependência espacial descritas pelos modelos esférico e exponencial. Na etapa seguinte deste trabalho foi realizada a análise multivariada dos dados para caracterizar a plataforma e o talude continentais da Bahia. A metodologia consistiu na aplicação dos métodos de agrupamento de Ward e *k*-Médias. Os resultados revelaram que as plataformas interna e média foram caracterizadas pela presença de areia siliciclástica com maior predominância das espécie *Peneroplis carinatus*; a plataforma externa foi caracterizada por areia carbonática cascalhosa tendo a espécie *Amphistegina lessonii* como a mais predominante; o talude foi caracterizado por sedimento formado por areia carbonática lamosa e pela maior predominância da espécie *Amphistegina lessonii*.

Palavras-chave: Geoestatística. Semivariância. Krigagem. Análise multivariada. Método de Ward. Método *k*-Médias.

GENERAL ABSTRACT

This work aimed at applying geostatistics and cluster analysis on data from foraminifera microfauna present in the sediments of the continental shelf and slope in the northern state of Bahia, Brazil. The geostatistical analysis was applied to evaluate the presence of spatial variability in the number of individuals of the foraminifera species and to make predictions using kriging. In order to obtain the experimental semivariograms of the studied variables, we used the Matheron and Cressie & Hawkins estimators. The results showed that the species *Amphistegina lessonii*, *Peneroplis carinatus*, *Quinqueloculina bicarinata* and *Reusella spinulos* presented spatial dependence structures described by the spherical and exponential models. In the following stage of this work, we performed the multivariate analysis in order to characterize the continental shelf and slope of Bahia. The methodology consisted in the application of Ward and *k*-Means' clustering methods. The results revealed that the inner and middle shelves were characterized by the presence of siliciclastic sand with a larger predominance of the specie *Peneroplis carinatus*; the outer shelf was characterized by gravelly carbonate sand, in which the specie *Amphistegina lessonii* was the most predominant; the slope was characterized by sediment formed by mudflat carbonate sand with a predominance of the specie *Amphistegina lessonii*.

Keywords: Geostatistics. Semivariance. Kriging. Multivariate analysis. Ward's method. *k*-means' method.

LISTA DE FIGURAS

PRIMEIRA PARTE

Figura 1 Esquema do semivariograma experimental 28

ARTIGO 1

Figura 1 Localização da área de estudo e das estações de coleta das amostras 65

Figura 2 Gráficos boxplot para as espécies *Reusella spinulosa* (A); *Quinqueloculina bicarinata* (B) e *Quinqueloculina disparilis curta* (C) 75

Figura 3 Gráficos boxplot para as espécies *Amphistegina lessonii* (D); *Peneroplis carinatus* (E); *Peneroplis bradyi* (F) e *Cibicides pseudoungerianus* (G) 76

Figura 4 Modelo efeito pepita puro (linha contínua) ajustado ao semivariograma experimental (pontos) para as espécies *Cibicides pseudoungerianus* (A); *Peneroplis bradyi* (B) e *Quinqueloculina disparilis curta* (C) 77

Figura 5 Modelo exponencial (linha contínua) ajustado ao semivariograma experimental (pontos) para a espécie *Peneroplis carinatus* 78

Figura 6 Modelo esférico (linha contínua) ajustado ao semivariograma experimental (pontos) para as espécies *Amphistegina lessonii* (A); *Quinqueloculina bicarinata* (B) e *Reusella spinulosa* (C) 79

Figura 7 Mapa de krigagem do número de indivíduos da espécie *Amphistegina lessonii* 81

Figura 8 Mapa de krigagem do número de indivíduos da espécie *Quinqueloculina bicarinata* 82

Figura 9	Mapa de krigagem do número de indivíduos da espécie <i>Reusella spinulosa</i>	83
Figura 10	Mapa de krigagem do número de indivíduos da espécie <i>Peneroplis carinatus</i>	84

ARTIGO 2

Figura 1	Localização da área de estudo e das estações de coleta das amostras	104
Figura 2	Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies da plataforma interna, a partir das amostras (localização).....	117
Figura 3	Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies da plataforma média, a partir das amostras (localização).....	121
Figura 4	Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies da plataforma externa, a partir das amostras (localização).....	125
Figura 5	Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies do talude, a partir das amostras (localização).....	129
Figura 6	Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as amostras considerando os teores sedimentológicos (Carbonato, Siliciclástico e Granulometria).....	133

ANEXOS

- Figura 1 *Peneroplis carinatus*. 2. *P. bradyi*. 3. *P. proteus*. 4. *Archaias angulatus*. 5. *Massilina pernambucensis*. 6. *Quinqueloculina bicarinata*. 7. *Q. bicostata*. 8. *Q. microcostata*. 9. *Q. moynensis*. 10. *Sigmavirgulina tortuosa*. 11. *Cassidulina subglobosa*. 12. *C. curvata*..... 151
- Figura 2 *Elphidium discoideale*. 2. *Siphonina reticulata*. 3. *Dicorbis mira*. 4. *Amphistegina gibbosa*. 5. *A. lessonii*. 6. *Bigenerina nodosaria*. 7. *Textularia agglutinans*. 8. *T. candeiana*. 9. *T. gramen*. 10. *Cibicides pseudoungerianus*, (lado dorsal). 11. *C. pseudoungerianus*, (lado ventral). 12. *Globigerinoides ruber*. 152

LISTA DE TABELAS

ARTIGO 1

Tabela 1	Profundidade, localização e posição geográfica das 38 amostras na região de estudo	66
Tabela 2	Medidas descritivas da contagem do número de indivíduos por espécie	74
Tabela 3	Parâmetros estimados no semivariograma.....	80

ARTIGO 2

Tabela 1	Profundidade e coordenadas das amostras de fundo estudadas	106
Tabela 2	Número de indivíduos das espécies da plataforma interna.....	119
Tabela 3	Medidas descritivas dos agrupamentos de espécies encontrados na plataforma interna pelo método de Ward.....	120
Tabela 4	Número de indivíduos das espécies da plataforma média	123
Tabela 5	Medidas descritivas dos agrupamentos de espécies encontrados na plataforma média pelo método de Ward	124
Tabela 6	Número de indivíduos das espécies da plataforma externa	127
Tabela 7	Medidas descritivas dos agrupamentos de espécies encontrados na plataforma externa pelo método de Ward	128
Tabela 8	Número de indivíduos das espécies do talude.....	131
Tabela 9	Medidas descritivas dos agrupamentos de espécies encontrados no talude pelo método de Ward	132
Tabela 10	Valores percentuais de carbonato, siliciclasto, cascalho, areia e lama dos sedimentos em cada amostra. Carb = carbonato; Silic = Siliciclasto	135
Tabela 11	Medidas descritivas dos agrupamentos de amostras formados pelo método de Ward.....	137

Tabela 12 Medidas descritivas dos agrupamentos de amostras formados pelo método k-Médias	137
---	-----

SUMÁRIO

	PRIMEIRA PARTE	
1	INTRODUÇÃO	15
2	REFERENCIAL TEÓRICO	17
2.1	Foraminíferos	17
2.2	Geoestatística	19
2.2.1	Teoria das variáveis regionalizadas	20
2.2.2	Hipóteses restritivas	21
2.2.3	Semivariograma	25
2.2.4	Modelos teóricos de semivariogramas	30
2.2.5	Métodos de estimação de parâmetros	33
2.2.6	Krigagem	34
2.3	Análise multivariada de agrupamento	34
2.3.1	Medidas de similaridade e dissimilaridade	36
2.3.2	Técnicas de agrupamento	41
2.3.3	Métodos de agrupamento hierárquicos	42
2.3.3.1	Método de ligação simples	42
2.3.3.2	Método de ligação completa	43
2.3.3.3	Método de ligação média	44
2.3.3.4	Método do centroide	44
2.3.3.5	Método de Ward	45
2.3.4	Método de agrupamento não hierárquico	48
2.3.4.1	Método das <i>k</i>-Médias	48
2.3.5	Dendrograma	49
2.3.6	Determinação do número de grupos	50
2.3.7	Validação do agrupamento	50
2.4	Métodos estatísticos na análise de foraminíferos	51
	REFERÊNCIAS	54
	SEGUNDA PARTE - ARTIGOS	60
	ARTIGO 1 Variabilidade espacial de foraminíferos nos sedimentos da plataforma e talude continentais do estado da Bahia	60
	ARTIGO 2 Caracterização da plataforma e do talude continentais do Estado da Bahia, a partir da análise de agrupamento dos dados sedimentológicos e da microfauna de foraminíferos	96
	CONSIDERAÇÕES FINAIS	153

PRIMEIRA PARTE

1 INTRODUÇÃO

O estudo da evolução dos ambientes marinhos constitui uma das principais linhas de pesquisa na área de geociências. Por meio da reconstrução paleoambiental pode-se determinar significativas mudanças no clima e na dinâmica dos oceanos, o que permite projetar alterações futuras para diferentes ecossistemas.

Alguns organismos marinhos são potenciais indicadores dessas mudanças, por possuírem esqueletos cuja análise revela algumas propriedades físico-químicas dos oceanos em que viveram. Dentre os diferentes grupos de microfósseis explorados com esta finalidade estão os foraminíferos.

Estudos revelaram que existe uma grande fauna de foraminíferos presentes nos sedimentos de superfície e subsuperfícies da plataforma e no talude continental que podem ser utilizadas como bioindicadoras ambientais, paleoclimáticas e paleobatimétricas.

Nas últimas décadas, devido ao avanço tecnológico, intensificou-se a obtenção de dados geológicos e biológicos quantitativos de organismos marinhos. Entretanto, a maioria das informações coletadas não é analisada estatisticamente. Em estudos para descrição biológica das associações de foraminíferos, geralmente, são utilizadas análises baseadas em estatísticas descritivas e índices de diversidade.

Em estudos oceanográficos, as assembleias de foraminíferos obtidas de diferentes localidades requerem análise e interpretação de dados provenientes de um grande número de variáveis. Nesse caso, a análise estatística multivariada se faz necessária, pois os procedimentos desenvolvidos por essas técnicas analisam os dados por meio da combinação das inúmeras variáveis, facilitando a

compreensão da variabilidade total dos dados. Assim, as técnicas estatísticas multivariadas começam a ser utilizadas. Entretanto, pode-se observar que a maioria das análises de dados em foraminíferos não leva em conta a componente espacial. Assim, antes de qualquer análise estatística deveria ser verificada a presença de dependência espacial nas variáveis analisadas.

Tendo em vista o que foi apresentado anteriormente, este trabalho tem como objetivo aplicar métodos de geoestatística e de análise de agrupamentos para caracterizar a microfauna de assembleias de foraminíferos da plataforma e talude continentais do norte do Estado da Bahia.

Na primeira etapa do trabalho será realizada uma ampla revisão bibliográfica sobre os foraminíferos presentes nos sedimentos marítimos que podem ser utilizadas como bioindicadoras ambientais e como os métodos estatísticos têm sido utilizados na análise de dados dos sedimentos e da microfauna de foraminíferos.

Na segunda etapa do trabalho será realizada uma análise espacial dos dados com o objetivo de avaliar a variabilidade espacial das variáveis em estudo. A análise será baseada em métodos da geoestatística tais como ajuste de modelos de semivariogramas e construção de mapas de krigagem.

Na terceira etapa do trabalho será realizada uma análise estatística multivariada dos dados com o objetivo de caracterizar a plataforma e talude continentais do norte da Bahia. A metodologia consistirá na aplicação da técnica de agrupamentos. Essa análise permitirá avaliar as condições sedimentológica e hidrodinâmica atuais, além de servirem de subsídio para futuros estudos de biomonitoramento da margem continental em estudo.

2 REFERENCIAL TEÓRICO

2.1 Foraminíferos

Os foraminíferos são organismos unicelulares, predominantemente marinhos, cujas características gerais e potencialidades para a pesquisa paleoambiental se resumem ao fato do grupo apresentar número populacional elevado, potencial de preservação, grande variabilidade morfológica, restrita movimentação e distribuição controlada por diferentes fatores abióticos. Tais fatores, em conjunto, podem agir controlando a reprodução e a mortalidade e, conseqüentemente, influenciam na distribuição das espécies, além de delimitarem e controlarem a formação das associações (ERNST et al., 2006).

Estes microorganismos são classificados sistematicamente como pertencentes ao Reino Protista, filo Sarcomastigophora (BRUSCA; BRUSCA 1990), podem ser bentônicos (vivem sobre ou dentro do substrato) ou planctônicos (vivem livres flutuando na superfície da água do oceano) e segregam uma testa, estrutura que protege o corpo celular, que é composta por carbonato de cálcio e por outros compostos inorgânicos, como sílica e sulfato de magnésio. Para fora da testa estendem-se os pseudópodos utilizados na movimentação e captura de alimentos.

Os foraminíferos são geologicamente importantes porque suas testas são adicionadas ao sedimento quando o organismo morre ou se reproduz. O estudo dessas estruturas é utilizado na identificação de espécies e as testas dos foraminíferos bentônicos constituem parte integrante do sedimento, permitindo, inclusive, o registro de evidências de estresse ambiental (YANKO; KRONFELD; FLEXER, 1994). Essas testas se comportam como grãos sedimentares, podendo ser transportadas e erodidas, e constituem porção significativa de carbonato de cálcio nos sedimentos da plataforma continental.

O modelo de distribuição de foraminíferos está condicionado a diversos fatores físicos, químicos e biológicos, incluindo correntes, turbulência, luz, salinidade, pH, temperatura, disponibilidade de nutrientes, concentração de oxigênio, competição, bioturbação e as propriedades do sedimento, tais como, conteúdo de carbono orgânico, tamanho do grão, e concentração de oxigênio na água intersticial (JORISSEN, 1999).

A maioria das espécies de foraminíferos varia de tamanho entre 0,1 e 1,0 mm de diâmetro. Porém, o crescimento máximo individual varia em tamanho de aproximadamente 100 µm (micrômetros) até quase 12 cm de comprimento (ARAÚJO, 2004). No Brasil, na Formação Pirábas, Mioceno do Estado do Pará e no Terciário do Golfo do México e do Caribe, são encontrados foraminíferos de até 20 mm de tamanho (VICALVI; FERREIRA, 1992).

Os foraminíferos são largamente usados nas pesquisas de: paleoceanografia, delimitando correntes oceânicas; paleoclimatologia, determinando mudanças climáticas do passado; paleogeografia, delimitando antigos mares e continentes; paleoecologia, interpretando ambientes antigos de sedimentação; estratigrafia, datando e correlacionando camadas geológicas; bioestratigrafia, identificando zonas que podem ser correlacionadas; geologia econômica, nas investigações para exploração de petróleo. Além disso, os foraminíferos são utilizados nos estudos sobre variação do nível relativo do mar, tectônica de placas e impacto ambiental (CARVALHO, 2000). Nas bacias costeiras de Sergipe e Alagoas foram feitas reconstituições paleogeográficas por meio de estudos com os foraminíferos (KOUTSOUKOS et al., 1993). Na Formação Itaituba, Carbonífero da Bacia Amazonas, ocorrem os fósseis-guia *Fusulinella* e *Millerella*, que dataram estas camadas (CARVALHO, 2000). As biológicas: marinha, molecular, celular e evolução aplicam estudos com base em foraminíferos em suas pesquisas (HALLOCK et al., 1995).

Os foraminíferos estão amplamente distribuídos na plataforma continental, como estuários salobres ou mangues intertidais salinos, lagoas costeiras e recifes de coral até as zonas batial e abissal, e dos polos até os trópicos (ARAÚJO; ARAÚJO, 2010; WETMORE, 1998).

2.2 Geoestatística

Geralmente, os métodos convencionais de análise estatística são baseados em amostras independentes e identicamente distribuídas, o que não acontece em dados que apresentam estrutura de dependência espacial, uma vez que as amostras são correlacionadas.

A ideia de dependência espacial surge com a primeira lei da geografia quando estabelece que todas as coisas possam ser parecidas, entretanto coisas mais próximas se parecem mais que coisas mais distantes (TOBLER, 1979). Segundo Cressie (1993), a dependência (espacial) está presente em todas as direções e torna-se mais fraca à medida que aumenta a dispersão na localização dos dados. Entende-se por dependência espacial o fato de que a maior parte das ocorrências naturais ou sociais apresenta entre si uma relação que depende da distância (CAMARGO; FUCKS; CÂMARA, 2004).

A técnica da estatística espacial nasce da necessidade de quantificação da dependência espacial presente num conjunto de dados, distinguindo-se das demais técnicas empregadas em análises estatísticas por considerar explicitamente as coordenadas dos dados no processo de coleta, descrição ou análise (CÂMARA et al., 2002).

A taxonomia mais utilizada para descrever o problema da análise espacial considera três tipos de dados: eventos ou padrões pontuais, áreas com contagens ou taxas agregadas e superfícies contínuas (geoestatística)

O conceito de geoestatística surgiu em 1951, por meio dos trabalhos desenvolvidos com dados de concentração de ouro, pelo engenheiro sul africano Daniel Gerhardus Krige que concluiu que as variâncias obtidas pela estatística clássica não faziam sentido se não considerasse a distância entre os locais onde foram observadas as amostras.

Os dados de superfícies contínuas são originários a partir de um conjunto de amostras de campo, que podem estar regularmente ou irregularmente distribuídas. Usualmente, este tipo de dados é resultante de levantamento de recursos naturais, e que incluem mapas geológicos, topográficos, ecológicos, fitogeográficos e pedológicos. Assim, a geoestatística é aplicada em vários ramos da ciência como metodologia para análise de dados georeferenciados, com o objetivo de determinar o grau de dependência espacial entre observações de uma variável, com base na direção e na distância entre elas e predizer valores em locais não amostrados (krigagem).

2.2.1 Teoria das variáveis regionalizadas

A geoestatística está fundamentada na teoria das variáveis regionalizadas que foi desenvolvida por Matheron na década de 60, a partir das observações de Krige. A teoria das variáveis regionalizadas leva em consideração que as variáveis são aleatórias, já que os valores numéricos observados podem variar consideravelmente de um ponto a outro do espaço e são estruturadas espacialmente, pois os valores das observações não são completamente independentes de sua localização geográfica.

Na teoria das variáveis regionalizadas, define-se $Z(s)$ como uma variável aleatória que assume diferentes valores em função da posição s em determinada região. O conjunto de variáveis $Z(s)$ medidas em toda a área pode ser considerado uma função aleatória $Z(s)$ uma vez que, segundo Isaaks e Srivastava

(1989) são variáveis aleatórias regionalizadas e assume-se que a dependência entre elas é especificada por algum mecanismo probabilístico.

A função aleatória está definida pela função de distribuição conjunta para quaisquer k componentes de um vetor aleatório.

$$F(z_1, z_2, \dots, z_k; s_1, s_2, \dots, s_k) = P(Z(s_1) \leq z_1, Z(s_2) \leq z_2, \dots, Z(s_k) \leq z_k) \quad (1)$$

A interpretação probabilística de que a variável regionalizada $Z(s)$ é uma particular realização de certa função aleatória $Z(s)$ é consistente quando se pode inferir toda ou, pelo menos, parte da lei de distribuição de probabilidade que define esta função aleatória (JOURNAL; HUIJBREGTS, 1978).

Na prática, em levantamentos amostrais com dados que apresentam dependência espacial não existem habitualmente repetições das observações. Em cada ponto s tem-se apenas uma realização de uma função aleatória ou processo estocástico $Z(s)$, que representa um valor medido $z(s)$ de uma determinada variável aleatória em estudo, tornando-se impossível inferir sobre a distribuição de $Z(s)$.

2.2.2 Hipóteses restritivas

Para inferir estatisticamente na geoestatística é preciso que o processo estocástico $Z(s)$ atenda a algumas hipóteses e a hipótese central da Geoestatística é a da estacionaridade (CÂMARA et al., 2002).

Na análise espacial, um processo estocástico é uma família ou coleção de variáveis aleatórias cujos elementos podem ser identificados ou localizados (indexados) por alguma métrica. Portanto, um processo espacial no plano é uma coleção de variáveis aleatórias indexadas por coordenadas do tipo $s = (x_i, y_i)$.

O processo estocástico $Z(s)$ é dito estacionário, se todos os momentos estatísticos forem constantes para qualquer ponto s , ou seja, não há tendência. Se uma função aleatória é estacionária, as leis da função são invariantes para as translações dos pontos s_1, s_2, \dots, s_k ou seja, esta função não varia no espaço e então, os diferentes momentos que existirem não se alteram para estas translações. Portanto, se $F(z_i)$ for a lei de distribuição das variáveis $Z(s_i)$, $i = 1, \dots, k$ tem-se que:

$$F(z_1, z_2, \dots, z_k; s_1 + h, s_2 + h, \dots, s_k + h) = F(z_1, z_2, \dots, z_k; s_1, s_2, \dots, s_k), \quad (2)$$

o que equivale a:

$$P(Z(s_1 + h) \leq z_1, \dots, Z(s_k + h) \leq z_k) = P(Z(s_1) \leq z_1, \dots, Z(s_k) \leq z_k) \quad (3)$$

Para estudos de Geoestatística exige-se no máximo, a estacionaridade de segunda ordem, ou seja, que o primeiro e o segundo momento em relação à origem sejam constantes (GUIMARÃES, 2004). A hipótese de estacionaridade de primeira ordem é definida por Trangmar, Yost e Uehara (1985) como sendo a hipótese de que o momento de primeira ordem da distribuição da função aleatória $Z(s)$ é constante em toda a área, ou seja,

$$E[Z(s)] = E[Z(s + h)] = m, \forall s, \dots \quad (4)$$

em que m é a média dos valores amostrais e não depende do ponto s . Isso implica que considerando um vetor h de separação entre dois pontos, que apresenta módulo e direção, para qualquer h tem-se:

$$E[Z(s) - Z(s+h)] = 0, \forall s. \quad (5)$$

Como a diferença entre as duas variáveis aleatórias $\{Z(s) - Z(s+h)\}$ é uma variável aleatória, o valor esperado da diferença entre os valores observados de s e $s+h$ é nulo, ou seja, o primeiro momento desta variável aleatória é igual a zero.

O fenômeno é dito estacionário de segunda ordem, quando além de atender à estacionaridade de primeira ordem, a função aleatória apresenta a característica de que, para cada par de valores $\{Z(s), Z(s+h)\}$, a covariância existe e depende apenas de h . Matematicamente, a estacionaridade de segunda ordem pode ser expressa por:

$$E[Z(s)] = m, \forall s \quad (6)$$

$$C(h) = E[Z(s)Z(s+h)] - E[Z(s)]E[Z(s+h)] = E[Z(s)Z(s+h)] - m^2, \forall s. \quad (7)$$

A existência de estacionaridade possibilita repetir um experimento, mesmo que as amostras sejam coletadas em pontos diferentes em relação ao experimento inicial, porque todas as amostras fazem parte de populações com mesmos momentos estatísticos.

A estacionaridade de segunda ordem implica na existência da covariância e, portanto uma variância finita de valores medidos.

$$\text{Var}[Z(s)] = E[Z^2(s)] - \{E[Z(s)]\}^2 = E[Z^2(s)] - m^2 = C(0). \quad (8)$$

Essa hipótese pode não ser satisfeita para alguns fenômenos físicos. Para tais situações, uma hipótese menos restritiva, a hipótese intrínseca, pode ser

aplicável (VIEIRA, 2000). Existem muitos fenômenos com uma capacidade infinita para dispersão, que não possuem uma variância a priori nem uma covariância. Então, assume-se a existência de uma função variância como uma hipótese mais fraca do que a existência da covariância. Segundo Guimarães (2004), uma hipótese mais fraca (mais abrangente) é a hipótese intrínseca.

A hipótese intrínseca é verificada se para todo vetor h , o incremento $[Z(s) - Z(s+h)]$ admitir esperança matemática e variância finita independentes da posição s , , mas dependendo apenas do valor de h . Ou seja, a hipótese intrínseca estabelece que sejam os acréscimos $[Z(s) - Z(s+h)]$ da variável regionalizada que obedecem a estacionaridade de segunda ordem, verificando as seguintes relações:

$$E[Z(s) - Z(s+h)] = 0, \forall s \text{ e } h \quad (9)$$

$$\text{Var}[Z(s) - Z(s+h)] = E[Z(s) - Z(s+h)]^2 = 2\gamma(h), \forall s, h \quad (10)$$

em que $2\gamma(h)$ é a função variância.

A função variância é uma função aleatória intrínseca, pois para qualquer deslocamento h , os dois primeiros momentos do incremento $[Z(s) - Z(s+h)]$ são independentes da localização de s e função apenas de h . Então, a função semivariância $\gamma(h)$ é definida por:

$$\gamma(h) = \frac{1}{2} E[Z(s) - Z(s+h)]^2 = C(0) - C(h). \quad (11)$$

Na hipótese de estacionaridade de segunda ordem, a covariância e a semivariância são ferramentas equivalentes para caracterizar a auto-correlação

entre duas variáveis $Z(s) - Z(s+h)$, separadas pela distância h (JOURNEL; HUIBREGTS, 1978). A primeira expressa a similaridade dos valores e a segunda o afastamento relativo destes.

A hipótese intrínseca é a hipótese mais frequentemente usada em Geoestatística, por ser menos restritiva e, portanto, o semivariograma é a ferramenta mais difundida na geoestatística porque exige apenas a hipótese intrínseca (GUIMARÃES, 2004).

2.2.3 Semivariograma

Segundo Ribeiro Júnior (1995), o semivariograma é a parte central dos estudos geoestatísticos, sendo capaz de descrever tanto qualitativamente quanto quantitativamente a variação espacial.

O gráfico de $\gamma(h)$ versus h (ver Figura 1) representa o semivariograma, que permite obter a estimativa do valor de semivariância para as diferentes combinações de pares de pontos e assim analisar o grau de dependência espacial da variável estudada e definir os parâmetros necessários para a estimativa de suas características em locais não amostrados (SOUZA, 1999).

Segundo Guerra (1988), existem três tipos de semivariogramas, o verdadeiro (real, mas desconhecido), o observado ou o experimental (obtido a partir das amostras coletadas no campo), gerado por meio da semivariância em função da distância h e o teórico (utilizado para o ajuste de modelos ao semivariograma experimental)

O gráfico de $\gamma(h)$ em função dos valores correspondentes de h , chamado semivariograma experimental, é uma função do vetor h e, portanto, depende de ambos, magnitude e direção de h (VIEIRA, 1995).

Se o semivariograma é idêntico para qualquer direção de h , então a dependência espacial é a mesma em todas as direções e o fenômeno é chamado

isotrópico. O semivariograma omnidirecional considera todas as direções e deve ser sempre comparado com o semivariograma em determinada direção, pois não apresentando diferenças significativas entre eles, o fenômeno é dito isotrópico

A anisotropia é uma característica muito frequente nos fenômenos da natureza, isto é, a variabilidade, ou a distribuição espacial das variáveis que quantificam estes fenômenos, ocorre mais intensamente numa direção e menos intensamente em outra direção (CAMARGO; FUCKS; CÂMARA, 2004).

A função de semivariância $\gamma(h)$ é uma função matemática intrínseca, definida como a metade da variância de diferenças entre todos os pares de pontos presentes na área estudada, separados por uma distância h conforme a equação (11). Portanto, essa função representa o grau de dependência espacial entre duas observações da função aleatória $Z(s)$, que pode ser estimada pelo estimador “clássico” de Matheron que, por definição, é dado por

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(s_i) - z(s_i + h)]^2, \dots \quad (12)$$

em que $\hat{\gamma}(h)$ é o estimador da semivariância para uma distância h ; $N(h)$ representa o número de pares de valores medidos $z(s_i)$, $z(s_i + h)$ separados por um vetor de distância h .

O estimador de Matheron (estimador dos momentos) é o mais utilizado para a semivariância $\gamma(h)$. Para Braga (1990), se $Z(s)$ é função aleatória estacionária, então esse estimador, sob a hipótese intrínseca é não tendencioso. Ribeiro Júnior (1995) acrescenta ainda que esse estimador de Matheron seja não viesado para $\gamma(h)$ quando $Z(s)$ representa um processo estocástico intrinsecamente estacionário, mas é pouco resistente e muito afetado por

observações atípicas (*outliers*), sobretudo aquelas que produzem distribuições de caudas longas.

Nesses casos, surge a necessidade de estimador robusto, que não seja afetado por valores discrepantes que não possam ser eliminados ou substituídos.

Cressie e Hawkins (1980) apresentam o estimador robusto de semivariância dado por

$$\hat{\gamma}(h) = \frac{1}{2} \frac{\left[N(h)^{-1} \sum_{i=1}^{N(h)} \sqrt{|z(s_i + h) - z(s_i)|} \right]^4}{0,457 + \frac{0,494}{N(h)}}. \quad (13)$$

A justificativa para o estimador robusto é que para variáveis normais, $[z(s_i + h) - z(s_i)]^2$ é uma variável aleatória com distribuição qui-quadrado com 1 grau de liberdade. A transformação que a torna mais próxima a uma distribuição normal é a raiz quarta, ou seja, $\sqrt[4]{|z(s_i + h) - z(s_i)|}$, também conhecida como a raiz quadrada da diferença absoluta. A inclusão do denominador visa corrigir o viés da distribuição (CRESSIE; HAWKINS, 1980).

Na Figura 1 está representado um semivariograma, próximo ao ideal, usado para representar o que se espera de dados amostrais de superfície contínua, ou seja, a medida que h aumenta a semivariância $\hat{\gamma}(h)$ também aumenta, pois é esperado que quanto mais próximos estiverem os pontos amostrados, maior será a semelhança entre eles e, portanto menor a semivariância.

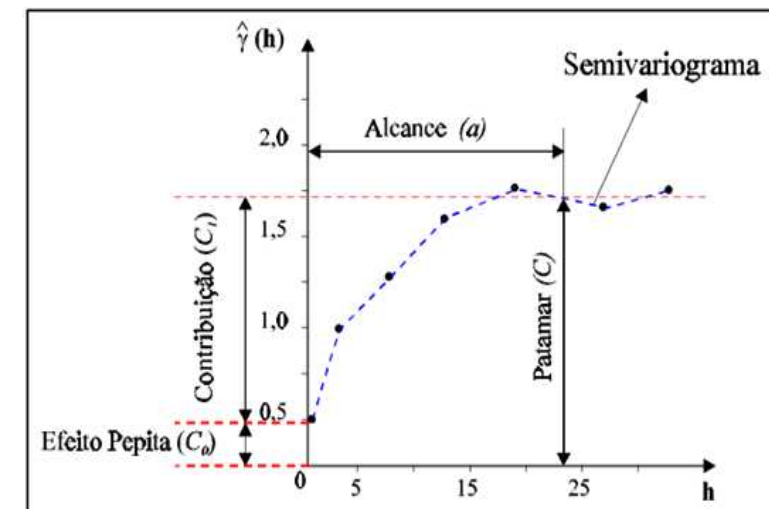


Figura 1 Esquema do semivariograma experimental

Fonte: Camargo, Fucks e Câmara (2004)

A interpretação dos parâmetros da Figura 1 é apresentada a seguir.

Alcance (a): o alcance da dependência espacial (a) corresponde ao valor da distância h , a partir da qual a semivariância $\gamma(h)$ se torna aproximadamente constante. É a distância a partir da qual as amostras passam a ser independentes, sem correlações espaciais, ou seja, a variação média entre duas observações não é mais função da distância entre elas. Então, o alcance do variograma separa o campo estruturado (amostras correlacionadas) do campo aleatório (amostras independentes).

Efeito Pepita (C_0): é o valor que representa a descontinuidade na origem do semivariograma, que é justificada por vários fatores como erros de amostragem, erros de medição, erros de análises de laboratório etc. Segundo Guimarães (2004), a utilização de dados amostrais na estimativa da

semivariância e na construção do semivariograma, revela que, frequentemente, para $h=0$, a semivariância $\gamma(0)$ difere de zero, embora na teoria, $\gamma(0)$ seja igual a zero. Para h ser igual a zero, duas amostras seriam coletadas no mesmo local, ou seja, seria realizada uma reamostragem exatamente sobre um ponto já amostrado. Na prática isso não é possível e quando a distância entre dois pontos diminui gradativamente tendendo a zero, $\gamma(h)$ se aproxima de um valor positivo chamado efeito pepita (C_0), que expressa a descontinuidade do semivariograma para distâncias muito próximas de zero, ou seja, para distâncias menores que a menor distância entre as observações.

Patamar (C): representa o nível de variabilidade, onde o variograma se estabiliza, ou seja, é o valor constante da semivariância $\gamma(h)$ que corresponde ao seu alcance. Teoricamente, esse patamar deve ser igual à variância amostral. Segundo Silva (1988), é um parâmetro importante, pois permite a determinação da distância limite entre dependência e independência das amostras. Se o semivariograma experimental for constante e igual ao patamar para qualquer valor de h , tem-se o efeito pepita puro e, nesse caso, tem-se a ausência total de dependência espacial. Assim, a dependência espacial, se existir, será manifestada à distância menor do que o menor espaçamento entre as amostras (GUIMARÃES, 2004).

Contribuição (C_1): é a diferença entre o patamar e o efeito pepita, correspondendo à variação espacial dos dados.

O grau de dependência espacial pode ser medido pela relação entre o efeito pepita e o patamar, chamado coeficiente de efeito pepita, dado por:

$$\varepsilon = \frac{C_0}{C_0 + C_1} \quad (14)$$

Segundo os valores encontrados para o coeficiente dado em (14), pode-se classificar o grau de dependência espacial do atributo em estudo, como forte, se for menor que 0,25; moderado, se os valores estiverem entre 0,25 e 0,75; e fraco dependência espacial, se o valor for maior que 0,75 (CAMBARDELLA et al. 1994).

O grau de dependência espacial (*GDE*) também pode ser avaliado a partir da relação entre a variação estruturada (contribuição) e o patamar (BIONDI; MYERS; AVERY, 1994). Sendo $GDE > 75\%$, o grau de dependência espacial é considerado alto; $35 \leq GDE \leq 75\%$ médio grau de dependência espacial e $< 35\%$, fraca estrutura de dependência espacial.

2.2.4 Modelos teóricos de semivariogramas

Modelos teóricos de semivariogramas são superpostos à sequência de pontos obtidos no semivariograma experimental, de modo que a curva que melhor se ajusta aos pontos obtidos represente a magnitude, o alcance e a intensidade da variabilidade espacial da variável estudada (ZIMBACK, 2003).

Segundo Guimarães (2004), a escolha do modelo de semivariograma que será utilizado é um dos aspectos mais importantes da geoestatística. Todos os cálculos da geoestatística dependem do modelo de semivariograma ajustado e, conseqüentemente, se o modelo ajustado não for apropriado, todos os cálculos seguintes conterão erros que poderão afetar as inferências. A seguir, apresentam-se alguns modelos teóricos de semivariogramas de uso mais frequente.

Modelo esférico

Segundo Andriotti (2009), o modelo esférico é o mais comumente utilizado nas variáveis de geociências. Ele apresenta um crescimento rápido na

origem, alcança um patamar a uma distância h finita e é representado por (VIEIRA, 1995):

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[\frac{3}{2} \left(\frac{|h|}{a} \right) - \frac{1}{2} \left(\frac{|h|}{a} \right)^3 \right], & \text{se } 0 < |h| \leq a; \\ C_0 + C_1, & \text{se } |h| > a. \end{cases} \quad (15)$$

Modelo exponencial

O modelo exponencial apresenta um comportamento linear na origem, alcança seu patamar apenas assintoticamente e é representado por (VIEIRA, 1995):

$$\gamma(h) = \begin{cases} C_0 + C_1 \left(1 - e^{-\frac{3|h|}{a}} \right), & \text{se } 0 < |h| \leq a; \\ C_0 + C_1, & \text{se } |h| > a. \end{cases} \quad (16)$$

Modelo Gaussiano

O que caracteriza esse modelo é o seu comportamento parabólico próximo à origem e é o único modelo que apresenta em sua forma um ponto de inflexão (ISAAKS; SRIVASTAVA, 1989). É dado por:

$$\gamma(h) = C_0 + C_1 \left[1 - e^{-3 \left(\frac{|h|}{a} \right)^2} \right] \text{ se } 0 < |h| \leq a. \quad (17)$$

Modelo potência ou linear

Este modelo é representado por:

$$\gamma(h) = C_0 + Ah^B, \text{ para } 0 < B < 2. \quad (18)$$

Os parâmetros A e B são constantes e segundo Guimarães (2004), este modelo implica uma dispersão ilimitada, entretanto, não tem patamar. O modelo linear é um caso particular do modelo potência e recebe esse nome quando $B = 1$.

Modelo efeito pepita puro

O modelo representa a variabilidade não espacial e é dado por:

$$\gamma(h) = C_0. \quad (19)$$

Nesse modelo, C_0 é um valor constante, que é o próprio efeito pepita.

2.2.5 Métodos de estimação de parâmetros

Segundo McBratney e Webster (1987), um procedimento baseado em critérios estatísticos deve ser preferível a um ajuste visual, na estimação do vetor de parâmetros de um determinado modelo teórico de semivariograma. A escolha de um modelo ajustado não é um procedimento automático. Requer um bom julgamento baseado na experiência e compreensão das limitações das funções.

Os métodos de estimação de parâmetros encontrados na literatura são: mínimos quadrados ordinários, mínimos quadrados ponderados, máxima verossimilhança e métodos bayesianos. Descreve-se a seguir o método que faz parte deste trabalho, ou seja, o método dos quadrados mínimos ordinários.

Este estimador é um dos mais utilizados da estatística e se baseia no princípio dos quadrados mínimos introduzido por Gauss em 1809 (BUSSAB; MORETTIN, 2010).

Seja $\theta = (C_0, C_1, a)^T$ o vetor de parâmetros desconhecidos a serem estimados. Segundo o método dos quadrados mínimos ordinários, o vetor θ a ser escolhido é o que minimiza a expressão:

$$\sum_{j=1}^k [\hat{\gamma}(h_j) - \gamma(h_j; \theta)]^2, \quad (20)$$

em que $\hat{\gamma}(h_j)$ é o valor estimado da semivariância, k corresponde ao número de distâncias (*lags*) do semivariograma experimental e $\gamma(h_j; \theta)$ é a semivariância h_j calculada pelo modelo teórico escolhido a ser ajustado, dependente dos parâmetros estimados por θ .

2.2.6 Krigagem

A krigagem compreende um conjunto de técnicas de estimação e predição de superfícies baseada na modelagem da estrutura de correlação espacial (CÂMARA et al., 2002). É um método de interpolação de dados que utiliza um estimador estatístico não tendencioso e eficiente, para estimar cada ponto da superfície a partir da interpolação das amostras mais próximas.

Segundo Andriotti (2009), o estimador de krigagem é do tipo BLUE (best linear unbiased estimator) e é dito linear por ser formado por uma combinação linear dos dados, ou seja:

$$Z^*(k) = \sum_{i=1}^n \lambda_i \cdot z(x_i) \quad (21)$$

em que λ_i são os ponderadores, $z(x_i)$ são os dados experimentais, n o número total desses dados e $Z^*(k)$ o estimador de krigagem. Este é o melhor estimador por apresentar a variância de estimação mínima e ser não tendencioso por assegurar que o somatório dos pesos é igual à unidade.

2.3 Análise multivariada de agrupamento

A estatística multivariada consiste em métodos e técnicas estatísticas aplicados no estudo de fenômenos em que são analisadas diversas variáveis simultaneamente em cada unidade amostral.

Segundo Mingoti (2005), a estatística multivariada se divide em dois grupos: um primeiro, consistindo em técnicas exploratórias de sintetização (ou simplificação) da estrutura de variabilidade dos dados, e um segundo,

consistindo em técnicas de inferência estatística. Fazem parte do primeiro grupo métodos como a análise de componentes principais, análise fatorial, análise de correlações canônicas, análise de agrupamento, análise discriminante e análise de correspondência. No segundo grupo, encontram-se os métodos de estimação de parâmetros, testes de hipóteses, análise de variância, de covariância e de regressão multivariadas.

Análise de agrupamento (originalmente *Cluster Analysis*) é um conjunto de técnicas que tem por objetivo identificar padrões ao formar grupos homogêneos (os mais semelhantes pertencem a um mesmo grupo) a partir de n observações ou elementos existentes. A construção dos grupos é feita de modo que as observações de um mesmo grupo pareçam-se mais entre si do que com as observações existentes nos outros grupos formados (BUSSAB; MIAZAKI; ANDRADE, 1990).

A análise agrupa os n objetos (indivíduos, itens, cidades, genótipos, etc.) em um número de grupos desconhecidos k . Pode-se também agrupar as p variáveis ao invés de agrupar objetos (FERREIRA, 2008).

Utilizando a notação de vetor aleatório para cada unidade amostral, cada vetor tem p componentes que representam as variáveis aleatórias p -variadas e indica o elemento do vetor por Y_{jk} , referindo-se a j -ésima unidade amostral na k -ésima variável mensurada, para $j=1, 2, \dots, n$ e $k=1, 2, \dots, p$.

Na análise multivariada representaremos as n unidades amostrais e $p > 1$ variáveis aleatórias medidas pela matriz $Y_{n \times p}$.

$$Y = \begin{bmatrix} Y_{11} & Y_{12} & \dots & Y_{1k} & \dots & Y_{1p} \\ Y_{21} & Y_{22} & \dots & Y_{2k} & \dots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{j1} & Y_{j2} & \dots & Y_{jk} & \dots & Y_{jp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \dots & Y_{nk} & \dots & Y_{np} \end{bmatrix},$$

em que, Y_j^t corresponde ao vetor linha p -dimensional de observações do j -ésima unidade amostral (objeto) e Y_k o vetor coluna n -dimensional de observações da k -ésima variável.

A análise de agrupamento, em geral, tem cinco etapas: (1) seleção de indivíduos ou de uma amostra de indivíduos a serem agrupados; (2) definição de um conjunto de variáveis a partir das quais serão obtidas informações necessárias ao agrupamento dos indivíduos; (3) definição de uma medida de semelhança ou distância entre os indivíduos; (4) escolha de um algoritmo de partição/classificação; (5) validação dos resultados encontrados.

2.3.1 Medidas de similaridade e dissimilaridade

Os métodos de agrupamento representam os objetos de acordo com suas similaridades ou dissimilaridades, pelas medidas de proximidade entre dois objetos. Com base nessa medida os indivíduos similares são agrupados e os demais são colocados em grupos separados.

Segundo Ferreira (2008), há uma necessidade de especificar um coeficiente que indique a proximidade entre os indivíduos. Para todos os casos, verifica-se a natureza das variáveis se são discretas, contínuas ou binárias e a escala de medida se é nominal, ordinal, real ou razão.

Os coeficientes de parecnça (similaridade /dissimilaridade) mais usuais podem medir: a distância entre pares de pontos, a correlação entre pares de valores ou a associação entre pares de caracteres qualitativos. Quando as variáveis em estudo são medidas em escalas de razão ou intervalares, ou seja, possuem propriedades métricas, a medida de distância que é uma medida de dissimilaridade é utilizada para indicar a “proximidade” entre dois objetos. A medida de dissimilaridade é adequada para variáveis quantitativas e quanto menor o valor, maior a semelhança entre os objetos.

A maioria dos métodos de análise de agrupamento requer uma medida de similaridade (ou dissimilaridade) entre os elementos a serem agrupados que é, normalmente, expressa como uma função distância ou métrica.

Seja B um conjunto. Uma métrica em B é uma função $d: B \times B \rightarrow R$, tal que para quaisquer $x, y, z \in B$, satisfaz as seguintes propriedades:

- a) $d(x, y) = d(y, x)$ (simétrica)
- b) $d(x, y) > 0, \forall x \neq y$
- c) $d(x, y) = 0$, se e somente se, $x = y$
- d) $d(x, y) < d(x, z) + d(y, z), \forall x, y, z$ (desigualdade triangular)

Os coeficientes de similaridade mais utilizados para atributos quantitativos são apresentados a seguir.

Distância quadrática: Sejam X e Y vetores p -dimensionais de observações dos objetos x e y para $x, y = 1, 2, \dots, n$. A distância quadrática entre x e y é definida por:

$$d^2(x, y) = (x - y)^t B(x - y). \quad (22)$$

em que $B_{p \times p}$ é uma matriz métrica positiva definida com as seguintes propriedades:

- a) Se B for a matriz identidade (I), a distância generalizada é a distância euclidiana dada por:

$$d^2(x, y) = (x - y)^t I (x - y) = \sum_{k=1}^p (x_k - y_k)^2. \quad (23)$$

A distância euclidiana é a mais conhecida para indicar a proximidade entre dois objetos e a mais adequada para variáveis que tenham escalas similares e não correlacionadas.

- b) Se a matriz $B = D^{-1} = \text{diag} \frac{1}{S_{kk}}$, então a distância euclidiana

padronizada (distância estatística de Karl Pearson) é dada por:

$$d^2(x, y) = (x - y)^t D^{-1} (x - y) = \sum_{k=1}^p \frac{1}{S_{kk}} (x_k - y_k)^2. \quad (24)$$

em que S_{kk} é a variância amostral da k -ésima variável aleatória, para $k = 1, 2, \dots, p$.

Nesse caso, a matriz B considera na ponderação apenas a diferença da variabilidade entre as variáveis e por isso a distância Euclidiana padronizada é indicada para variáveis não correlacionadas e que estão em diferentes escalas.

- c) Se a matriz $B = S^{-1}$ tem-se a distância de Mahalanobis expressa por:

$$d^2(x, y) = (x - y)' S^{-1} (x - y) \quad (25)$$

Na distância dada em (25), a matriz B considera na ponderação quaisquer padrões de covariância entre as diferentes variáveis, ou seja, contempla as possíveis diferenças de variâncias e as relações lineares entre as variáveis.

Segundo Bussab, Miazaki e Andrade (1990), os coeficientes de parença são criados com o intuito de moldar situações de interesse do pesquisador, e por isso depara-se com uma série bem ampla de tais medidas, em que muitas delas aparecem em publicações de áreas específicas em que é grande a utilização de técnicas de análise de agrupamento. Alguns coeficientes de uso frequente:

Valor absoluto: a distância baseada no valor absoluto entre x e y , $x \neq y$ é definida por:

$$d(x, y) = \sum_{i=1}^p w_i |x_i - y_i| \quad (26)$$

em que w_i 's são os pesos de ponderação para as variáveis. Os valores comumente usados são $w_i = 1$ (equiponderação) ou $w_i = 1/p$ (da média).

Distância de Minkowski: a distância de Minkowski entre x e y , $x \neq y$ é definida por:

$$d(x, y) = \left[\sum_{i=1}^p w_i |x_i - y_i|^\lambda \right]^{1/\lambda} \quad (27)$$

O peso dado às pequenas e grandes distâncias é definido pela escolha do valor escalar real λ . Para $p=1$ e $\lambda =1$, a distância é identificada como Manhattan ou *city-block* que geralmente é pouco afetada pela presença de valores discrepantes (*outliers*) e para $p=2$ e $\lambda =2$, tem-se a distância Euclidiana.

Coefficiente de correlação: o coeficiente de correlação (ex. Pearson e Spearman) é uma medida de covariância que é também um tipo de proximidade, mas não necessariamente de similaridade. Segundo Ferreira (2008), se dois objetos forem perfeitamente correlacionados, não significa que sejam similares, mas que os elementos de seus vetores são linearmente relacionados.

Ao usar coeficientes de similaridades ou distâncias obtidas a partir do coeficiente de correlação entre observações de dois objetos quaisquer, recomenda-se uma transformação sobre o conjunto original dos dados (matriz de dados) de modo que cada variável apresente média zero e variância igual a um. Assim, seja a matriz de dados com p variáveis ($k =1, 2, \dots, p$) e n objetos ($j= 1, 2, \dots, n$). O valor padronizado (Z_{jk}) do j -ésimo objeto e k -ésima variável (Y_{jk}) é:

$$Z_{jk} = \frac{Y_{jk} - \bar{Y}_k}{S_k} \quad (28)$$

Se o valor numérico entre as variáveis diferem acentuadamente, de modo que a comparação direta entre elas levaria a uma ponderação maior para as variáveis com maior valor numérico ou se as variáveis foram medidas em unidades diferentes, então recomenda-se antes de aglomerar as amostras efetuar a padronização dos dados originais. Portanto, a finalidade da padronização é expressar cada observação em termos de variações inerentes ao sistema, permitindo que os resultados da análise possam ser interpretados

apropriadamente de modo que as características contribuam igualmente na avaliação da similaridade entre os indivíduos.

2.3.2 Técnicas de agrupamento

As técnicas de agrupamento podem ser classificadas em dois tipos: hierárquicas e não hierárquicas. Para aplicar as técnicas não hierárquicas é preciso que o número de grupos já tenha sido pré- definido pelo pesquisador. As técnicas hierárquicas são as mais utilizadas e, geralmente, são aplicadas em análises exploratórias de dados com o objetivo de identificar os agrupamentos possíveis e o número de grupos formados.

Segundo Bussab, Miazaki e Andrade (1990), as técnicas hierárquicas podem ser subdivididas em dois tipos: aglomerativas, em que por meio de fusões sucessivas dos n objetos, vão sendo obtidos $n - 1$, $n - 2$, ... grupos, até reunir todos os objetos num único grupo; divisivos, partem de um único grupo, e por divisões sucessivas vão sendo obtidos 2, 3,... grupos. O que caracteriza estes processos é que a reunião de dois agrupamentos numa certa etapa produz um dos agrupamentos da etapa superior, caracterizando o processo hierárquico.

Segundo Mingoti (2005), as principais etapas para aplicar as técnicas hierárquicas aglomerativas são resumidas da seguinte forma: (1) no estágio inicial, cada elemento amostral constitui um cluster de tamanho 1. Portanto, tem-se n clusters, ou seja, n grupos com um objeto cada; (2) em cada estágio do algoritmo de agrupamento, os pares de conglomerados mais “similares” são combinados e passam a constituir um único conglomerado. Apenas um novo conglomerado pode ser formado em cada passo. Dessa forma, em cada estágio do processo, o número de conglomerados vai sendo diminuído; (3) propriedade da hierarquia. Em cada estágio do algoritmo, cada novo conglomerado formado é um agrupamento de conglomerados formados nos estágios anteriores. Se dois

elementos amostrais aparecem juntos no mesmo *cluster* em algum estágio do processo de agrupamento, eles aparecerão juntos em todos os estágios subsequentes, ou seja, uma vez unidos estes elementos não poderão ser separados; (4) construir o gráfico que representa o agrupamento (dendrograma).

Em termos de variabilidade, no estágio inicial, tem-se a partição com a menor dispersão interna possível, já que todos os conglomerados têm um único elemento e, logo, a variância de cada um deles é igual a zero. No estágio final, tem-se a maior dispersão interna possível, já que todos os elementos amostrais estão num único cluster.

Segundo Landim (2011), a análise de agrupamento é utilizada quando se deseja explorar as similaridades entre os indivíduos (modo Q) ou entre as variáveis (modo R), definindo-se grupos que consideram simultaneamente, no primeiro caso, todas as variáveis observadas em cada indivíduo e, no segundo, todos os indivíduos nos quais foram feitas as medidas.

2.3.3 Métodos de agrupamento hierárquicos

Na literatura já foram propostos vários métodos de agrupamento, sendo os hierárquicos os mais importantes (BUSSAB; MIAZAKI; ANDRADE, 1990). A seguir, apresentam-se alguns métodos de agrupamento hierárquicos de uso mais frequente.

2.3.3.1 Método de ligação simples

Nesse método, a similaridade entre dois grupos é definida pelos dois elementos mais parecidos entre si, que se baseando na distância mínima, regra do vizinho mais próximo. A, os dois primeiros objetos agrupados são os que apresentam menor distância entre si. Identificando a menor distância o terceiro

objeto é agrupado com os dois primeiros, ou forma-se um novo grupo de dois objetos. Em cada estágio a distância entre dois grupos é definida como a distância entre seus dois pontos mais próximos e o processo continua até que todos os objetos estejam em um único grupo.

Segundo Lattin, Carroll e Green (2011), o agrupamento de ligação simples é eficiente do ponto de vista computacional e mais eficiente para dados escassos, não exigindo dados métricos. Como desvantagem, esse método também conhecido como método do vizinho mais próximo, possui uma tendência a produzir agrupamentos longos e encadeados.

Encadeamento é um termo que descreve a situação em que há um primeiro grupo de um ou mais elementos que passa a incorporar, a cada interação, um grupo de apenas um elemento. Assim, é formada uma longa cadeia, em que se torna difícil definir um nível de corte para classificar os elementos em grupos.

2.3.3.2 Método de ligação completa

Nesse método, a similaridade entre dois grupos é definida pelos objetos de cada grupo que menos se parecem, ou seja, baseia-se na distância máxima entre objetos do grupo. Em cada estágio do processo de agrupamento, a distância é calculada para todos os pares de grupos, sendo então, combinados num único aqueles que apresentarem o menor valor da distância, isto é, o menor valor de máximo (MINGOTI, 2005). Assim, a fusão é feita com os grupos mais parecidos (BUSSAB; MIAZAKI; ANDRADE, 1990).

Esse método também denominado método do vizinho mais distante tende a formar grupos compactos. Porém a ligação completa pode ser altamente sensível a discrepância nos dados (LATTIN; CARROLL; GREEN, 2011).

2.3.3.3 Método de ligação média

Esse método, conhecido também como método da distância média, define a distância entre dois grupos como sendo a média das distâncias entre todos os pares de elementos, sendo um em cada grupo.

Os grupos são reunidos em um novo grupo quando a média das distâncias entre seus elementos é mínima. Um grupo passa a ser definido como um conjunto de indivíduos no qual cada um tem mais semelhanças, em média, com todos os membros do mesmo grupo do que com todos os elementos de qualquer outro grupo (REIS, 1997).

Esse método tem como vantagem evitar valores extremos e considerar toda a informação dos grupos, tendendo a formar grupos com número de elementos similares.

No caso de dados não métricos, a média pode ser substituída pela mediana e neste caso o método é chamado de ligação mediana.

2.3.3.4 Método do centroide

Segundo Lattin, Carroll e Green (2011), para calcular a distância entre dois agrupamentos, determina-se a média dos objetos em cada agrupamento (ou seja, calcula os centroides do agrupamento) e define a distância entre os dois centroides. Portanto, nesse método, a distância entre dois grupos é definida como a distância euclidiana quadrática entre os vetores de médias, centroides, dos dois grupos (FERREIRA, 2008).

O método centroide é robusto para as discrepâncias, mas pode ser superado pela ligação média (LATTIN; CARROLL; GREEN, 2011). Uma desvantagem desse método é que, em termos de dimensão, se os dois grupos forem muito diferentes, o centroide do novo agrupamento estará mais próximo

daquele grupo com maior peso (mais objetos), tendendo a perder as características do grupo menor.

2.3.3.5 Método de Ward

Nos algoritmos de agrupamentos anteriores, com exceção do método do centroide, a variação entre grupos diminui e a variação dentro dos grupos aumenta. Considerando essa “mudança de variação” entre os grupos, Ward (1963) apresentou um método fundamentado na análise de variância, que busca juntar dois grupos cuja fusão minimiza a soma de quadrados dentro dos grupos, ou seja, a soma de quadrados dos erros, isto é, a variância mínima dentro do grupo.

No método de Ward, também denominado “Mínima Variância”, a formação dos grupos se dá pela maximização da homogeneidade dentro dos grupos e as somas de quadrados são distâncias quadráticas, usadas como medidas de homogeneidade.

Segundo Ferreira (2008), o método de Ward norteia-se pelas seguintes ideias: inicialmente, têm-se tantos grupos quanto objetos, ou seja, temos n grupos com 1 objeto cada. Devem-se agrupar, no primeiro ciclo do algoritmo, os dois grupos, de um único objeto cada, que minimizam o incremento na soma de quadrados dos erros em relação ao estágio anterior. É conveniente ter a noção clara que ao aglomerar dois grupos, a soma de quadrados dos erros irá aumentar. O que busca-se é o par de grupos que provocará o menor incremento possível na soma de quadrados dos erros entre todos os possíveis pares. Ao aglomerar os grupos que atendem à função objetivo, minimizar o incremento na soma de quadrados dos erros deve-se atualizar a matriz de dissimilaridades e então repetir o processo.

Segundo Rencher (2002), se AB é um grupo obtido pela combinação dos grupos A e B , então a soma das distâncias dentro do grupo (dos elementos a partir do vetor de média dos grupos) são SSE_A (a soma de quadrados relativa ao grupo A), SSE_B (a soma de quadrados relativa ao grupo B) e SSE_{AB} (a soma de quadrados relativa ao grupo AB), em que:

$$SSE_A = \sum_{i=1}^{n_A} (y_{iA} - \bar{y}_A)' (y_{iA} - \bar{y}_A) \quad (29)$$

sendo $y_{iA} = (y_{i1A}, y_{i2A}, \dots, y_{ikA})'$ o vetor de observações do elemento amostral i pertencente ao grupo A , $\bar{y}_A = (\bar{y}_{1A}, \bar{y}_{2A}, \dots, \bar{y}_{kA})'$ o vetor de médias amostrais do grupo A , $i = 1, 2, \dots, k$, sendo k o número de variáveis e n_A o número de elementos amostrais do grupo A ;

$$SSE_B = \sum_{i=1}^{n_B} (y_{iB} - \bar{y}_B)' (y_{iB} - \bar{y}_B) \quad (30)$$

sendo $y_{iB} = (y_{i1B}, y_{i2B}, \dots, y_{ikB})'$ o vetor de observações do elemento amostral i pertencente ao grupo B , $\bar{y}_B = (\bar{y}_{1B}, \bar{y}_{2B}, \dots, \bar{y}_{kB})'$ o vetor e médias amostrais do grupo B , $i = 1, 2, \dots, k$, sendo k o número de variáveis e n_B o número de elementos amostrais do grupo B .

$$SSE_{AB} = \sum_{i=1}^{n_{AB}} (y_{iAB} - \bar{y}_{AB})' (y_{iAB} - \bar{y}_{AB}) , \quad (31)$$

sendo $y_{iAB} = (y_{i1AB}, y_{i2AB}, \dots, y_{ikAB})'$ o vetor de observações do elemento amostral i pertencente ao grupo AB , $\bar{y}_{AB} = (\bar{y}_{1AB}, \bar{y}_{2AB}, \dots, \bar{y}_{kAB})'$ o vetor e médias amostrais do grupo AB , $i = 1, 2, \dots, k$, sendo k o número de variáveis e n_{AB} o número de elementos amostrais do grupo AB . Sendo,

$$\bar{y}_{AB} = \frac{(n_A \bar{y}_A + n_B \bar{y}_B)}{n_A n_B}, \quad (32)$$

em que, n_A, n_B e $n_{AB} = n_A + n_B$ são os números de pontos em A , B e AB , respectivamente. Então, as somas das distâncias são equivalentes à soma de quadrados dentro dos grupos, denotadas por SSE_A, SSE_B e SSE_{AB} .

O método de Ward une dois grupos A e B que minimizam o acréscimo em SSE , definido como:

$$I_{AB} = SSE_{AB} - (SSE_A + SSE_B). \quad (33)$$

Conforme Rencher (2002) pode-se demonstrar que I_{AB} em (10) possui duas formas equivalentes:

$$I_{AB} = n_A (\bar{y}_A - \bar{y}_{AB})' (\bar{y}_A - \bar{y}_{AB}) + n_B (\bar{y}_B - \bar{y}_{AB})' (\bar{y}_B - \bar{y}_{AB}) \quad (34)$$

$$\text{e } I_{AB} = \frac{n_A n_B}{n_A + n_B} (\bar{y}_A - \bar{y}_B)' (\bar{y}_A - \bar{y}_B). \quad (35)$$

Assim, por (35) minimizar o aumento na SSE é equivalente a minimizar a distância entre grupos. Se A e B possuem apenas um elemento y_i e y_j , respectivamente, então $SSE_A = 0$ e $SSE_B = 0$, e (11) e (12) reduz a:

$$I_{ij} = I_{AB} = SSE_{AB} = \frac{1}{2}(\bar{y}_A - \bar{y}_B)'(\bar{y}_A - \bar{y}_B) = \frac{1}{2}d^2(y_i, y_j). \quad (36)$$

O método de Ward é um método não sensível a *outliers*, com tendência a combinar grupos com poucos elementos e a produzir agrupamentos de tamanhos aproximadamente iguais, ou seja, agrupamentos com aproximadamente o mesmo número de observações em cada, devido a sua minimização de variação interna. Nesses aspectos, muitos estudos mostraram que, de maneira geral, o método de Ward possui melhor desempenho (RENCHER, 2002).

2.3.4 Método de agrupamento não hierárquico

Existem muitos métodos não hierárquicos baseados em misturas de distribuição, estimação de densidades e partição (FERRERRA, 2008). A seguir, apresenta-se o método não hierárquico mais conhecido e utilizado em problemas práticos.

2.3.4.1 Método das k -Médias

O método das k -Médias é um método de partição baseado na redução da soma de quadrados dos erros que tem como objetivo a redução da heterogeneidade entre os grupos.

Segundo Mingoti (2005), basicamente, cada elemento amostral é alocado àquele *cluster* cujo centroide (vetor das médias amostral) é o mais próximo vetor de valores observados para o respectivo elemento e, originalmente, o método é composto de quatro passos: (1) primeiramente escolhe-se k centroides, chamados de “sementes” para inicializar o processo de partição; (2) cada elemento do conjunto de dados é, então, comparado com cada centroide inicial, por meio de uma medida de distância que, em geral, é a distância Euclidiana. O elemento é alocado ao grupo cuja distância é menor; (3) depois de aplicar o passo 2 para cada um dos n elementos amostrais, recalcula-se os valores dos centroides para cada novo grupo formado, e repete-se o passo 2, considerando os centroides desses novos grupos; (4) os passos 2 e 3 devem ser repetidos até que todos os elementos amostrais estejam bem “alocados” em seus grupos, isto é, até que nenhuma realocação de elementos seja necessária.

2.3.5 Dendrograma

Devido à propriedade de hierarquia, os grupos são geralmente representados por um diagrama (gráfico) bi-dimensional chamado de dendrograma (ou diagrama de árvore). Neste diagrama, cada ramo representa um elemento, enquanto a raiz representa o agrupamento de todos os elementos. No eixo horizontal, são marcados os elementos amostrais numa ordem conveniente relacionada à história do agrupamento. As linhas verticais, partindo dos elementos amostrais agrupados, têm altura correspondente ao nível em que os elementos foram considerados semelhantes, isto é, a distância do agrupamento ou o nível de similaridade (MINGOTI, 2005).

2.3.6 Determinação do número de grupos

Uma questão importante é como determinar o número final de grupos, o qual define a partição do conjunto de dados analisados, pois não há uma resposta exata.

O número de grupos nos métodos hierárquicos pode ser determinado pelo ponto de corte na escala de distâncias, no dendrograma. Segundo Mojena (1977), o número de grupos que aperfeiçoa a qualidade do ajuste do agrupamento de dados pode ser determinado pela maior amplitude nas distâncias de junção dos grupos formados. Mingoti (2005) relaciona outros critérios a serem adotados: análise da soma de quadrados entre grupos (coeficiente R^2); estatística pseudo F; correlação semiparcial.

2.3.7 Validação do agrupamento

Validar o agrupamento significa certificar-se que os grupos realmente diferem uns dos outros (FERREIRA, 2008). Uma medida de validação utilizada, principalmente, nos métodos de agrupamentos hierárquicos é o coeficiente de correlação cofenética. Este índice compara as distâncias efetivamente observadas entre os objetos e as distâncias previstas a partir do processo de agrupamento. Portanto, o coeficiente de correlação cofenética mede o grau de ajuste entre a matriz de dissimilaridade original e a matriz de dissimilaridade produzida após a construção do dendrograma (matriz cofenética).

Um método é melhor do que o outro, quando o dendrograma fornece uma imagem menos distorcida da realidade. O coeficiente de correlação cofenética permite avaliar o grau de deformação provocado pela construção do dendrograma e o menor grau de distorção será definido pelo maior valor do

coeficiente cofenética. Segundo Bussab, Miazaki e Andrade (1990), a correlação cofenética é dada por:

$$r_{cof} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n (c_{ij} - \bar{c})(d_{ij} - \bar{d})}{\sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^n (c_{ij} - \bar{c})^2} \sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^n (d_{ij} - \bar{d})^2}}, \quad (37)$$

em que c_{ij} = valor de dissimilaridade entre os objetos i e j , obtidos pela matriz cofenética, d_{ij} = valor de dissimilaridade entre os objetos i e j , obtidos pela matriz de dissimilaridade original

$$\bar{c} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_{ij}, \quad e \quad (38)$$

$$\bar{d} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n d_{ij}. \quad (39)$$

2.4 Métodos estatísticos na análise de foraminíferos

Pode-se observar que na maioria dos trabalhos científicos sobre associação de foraminíferos são utilizadas estatísticas descritivas tais como distribuições de frequência, medidas de tendência central e dispersão e índices de riqueza e diversidade (ARAÚJO, 2009; LIMAS; BONETTI; BONETTI, 2011; VIEIRA, 2011).

Métodos multivariados começam a ser utilizados na análise dos dados de foraminíferos. Alperin, Bernasconi e Cusminsky (2008) caracterizaram a plataforma continental da Argentina aplicando a análise de componentes principais e a análise de agrupamento aos dados da microfauna de foraminíferos presentes nos sedimentos da área em estudo.

Braga, Machado e Moraes (2008) utilizaram a análise de componentes principais e análise de agrupamento para estudar a fauna de foraminíferos da ordem Rotaliida da parte da borda do recife de Itacimirim no litoral da Bahia.

Santos et al. (2007) caracterizaram a baía de Guanabara, como um ambiente altamente poluído, utilizando a análise de agrupamento entre dados dos foraminíferos bentônicos.

A análise fatorial foi utilizada para identificar as espécies de foraminíferos na costa da Índia (GANDHI; SOLAI, 2010; JAYALAKSHMY; RAO, 2003).

Buzas (1970) aplicou análise de variância multivariada e análise canônica para estudar a homogeneidade das espécies de populações de foraminíferos.

Machado e Araújo (2012) aplicaram a análise de agrupamento entre espécies (modo R) e entre amostras (modo Q) e relacionaram a microfauna de foraminíferos e a granulometria no recife em Abrolhos (BA).

Verificam-se na literatura poucos estudos que utilizam métodos de estatística espacial para análise de dados de foraminíferos. Legendre (1993) e Legendre e Fortin (1989) afirmam que a autocorrelação espacial, que provoca perda de independência, é uma propriedade da maioria dos dados ecológicos. Telford e Birks (2005) argumentaram que muitos resultados de trabalhos publicados com dados ecológicos podem estar comprometidos, por não considerarem nas análises estatísticas o componente espacial. Telford e Birks

(2009) mostraram que há estrutura espacial nos dados ecológicos de assembleias de foraminíferos encontradas nos oceanos Atlântico, Ártico e Pacífico.

Parizotto et al. (2008) utilizaram dados do sedimento de foraminíferos e krigagem para obter mapas dos perfis de salinidade e turbidez da zona estuarina do Rio Biguaçu (SC).

Limas, Bonetti e Bonetti (2011) identificaram a condição ecológica dos subambientes de sedimentação que ocorrem na lagoa da Conceição (SC), por meio da integração dos resultados da análise de agrupamento e da espacialização dos dados sedimentológicos.

Neste capítulo procurou-se fazer uma revisão bibliográfica detalhada da base teórica que será utilizada nos dois próximos capítulos (artigos), uma vez que nos artigos essa teoria não é apresentada de maneira detalhada. Com isso, espera-se que o leitor não familiarizado com a geoestatística e análise de agrupamento possa voltar a este capítulo sempre que algum método aplicado não estiver detalhado o suficiente para o entendimento.

Procurou-se também, fazer uma breve apresentação do estado da arte sobre o uso de métodos estatísticos na análise de dados da microfauna de foraminíferos. Com isso, espera-se que o leitor possa se situar no atual desenvolvimento da área e passe a ter uma ideia da importância desta tese.

REFERÊNCIAS

ALPERIN, M.; BERNASCONI, E.; CUSMINSKY, G. C. C. Asociaciones de foraminíferos bentónicos recientes de la Plataforma Continental Argentina (39°-43° S y 58°-65° O) analizadas con métodos estadísticos para datos composicionales. **Ameghiniana**, Buenos Aires, v. 45, n. 2, abr./jun. 2008. Disponível em: <http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S0002-70142008000200015>. Acesso em: 10 fev. 2013.

ANDRIOTTI, J. L. S. **Fundamentos de estatística e geoestatística**. São Leopoldo: UNISINOS, 2009. 165 p.

ARAUJO, H. A. B. **Assembleias de foraminíferos indicadores de mudanças ambientais no complexo recifal de Abrolhos, Bahia**. 2009. 209 p. Tese (Doutorado em Geologia) - Universidade Federal da Bahia, Salvador, 2009.

ARAÚJO, T. M. F. **Estudo da microfauna de foraminíferos do sedimento da superfície e da subsuperfície da plataforma e do talude continentais da região norte do Estado da Bahia, Salvador à Barra do Itariri**. 2004. 528 p. Tese (Doutorado em Geologia) - Universidade Federal da Bahia, Salvador, 2004.

ARAÚJO, T. M. F.; ARAÚJO, H. A. B. Assembleias de foraminíferos encontradas nos sedimentos superficiais da plataforma continental e talude superior do norte da Bahia. **Revista de Geologia**, Fortaleza, v. 23, n. 1, p. 115-134, 2010.

BIONDI, F.; MYERS, D. E.; AVERY, C. C. Geostatistically modelling stem size and increment in an old growth forest. **Canadian Journal of Forest Research**, Ottawa, v. 24, n. 7, p. 1354-1368, July 1994.

BRAGA, L. P. V. Geoestatística e aplicações. In: SIMPÓSIO BRASILEIRO DE PROBABILIDADE E ESTATÍSTICA, 9., 1990, São Paulo. **Anais...** São Paulo: IME/USP, 1990. p. 36.

BRAGAA, Y. S.; MACHADO, A. J.; MORAES, S. S. A fauna de Rotallida (Foraminiferida) da borda do recife de Itacimirim, Bahia, Brasil: um estudo comparativo do sedimento e das algas como substrato. **Revista de Geologia**, Fortaleza, v. 21, n. 2, p. 147-157, 2008.

BRUSCA, R. C.; BRUSCA, G. J. **Invertebrates**. Sunderland: Sinauer Associates, 1990. 922 p.

BUSSAB, W. O.; MIAZAKI, E. S.; ANDRADE, D. F. **Análise de agrupamentos**. São Paulo: IME-USP, 1990. 87 p.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. São Paulo: Saraiva, 2010. 540 p.

BUZAS, M. A. Spatial homogeneity: statistical analyses of unispecies and multispecies populations of foraminifera. **Ecology**, Durham, v. 51, n. 5, p. 874-879, 1970.

CÂMARA, G. et al. **Análise espacial e geoprocessamento**. São José dos Campos: INPE, 2002. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro>>. Acesso em: 3 ago. 2011.

CAMARGO, E. C. G.; FUCKS, S. D.; CÂMARA, G. Análise espacial de superfícies. In: CAMARGO, E. C. G. (Ed.). **Análise espacial de dados geográficos**. Brasília: EMBRAPA, 2004. p. 1-26.

CAMBARDELLA, C. A. et al. Field: scale variability of soil properties in Central Iowa Soils. **Soil Science Society America Journal**, Madison, v. 58, n. 5, p. 1501-1511, 1994.

CARVALHO, I. S. **Paleontologia**. Rio de Janeiro: Interciência, 2000. 628 p.

CRESSIE, N. **Statistics for spatial data**. New York: J. Wiley, 1993. 900 p.

CRESSIE, N. A. C.; HAWKINS, D. M. Robust estimation of the variogram. **Journal of the International Association for Mathematical Geology**, New York, v. 12, p. 115-125, 1980.

ERNST, S. R. et al. Benthic foraminiferal response to experimentally induced Erika oil pollution. **Marine Micropalaeontology**, New York, v. 61, n. 1/3, p. 76-93, Oct. 2006.

FERREIRA, D. F. **Estatística multivariada**. Lavras: UFLA, 2008. 662 p.

GANDHI, M. S.; SOLAI, A. Statistical studies and ecology of benthic foraminifera from the depositional environment: a case study between Mandapam and Tuticorin, South East Coast of India. **International Journal of Research and Reviews in Applied Sciences**, Islamabad, v. 5, n. 1, p. 86-94, Oct. 2010.

GUERRA, P. A. G. **Geoestatística operacional**. Brasília: MME/DNPM, 1988. 145 p.

GUIMARÃES, E. C. **Geoestatística básica e aplicada**. Uberlândia: UFU, 2004. 77 p.

HALLOCK, P. et al. A new disease in reef-dwelling foraminifera: implications for coastal sedimentation. **Journal of Foraminiferal Research**, Washington, v. 25, n. 3, p. 280-286, 1995.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University, 1989. 561 p.

JAYALAKSHMY, K. V.; RAO, K. K. Multivariate statistical study with a factor analysis of foraminiferal fauna from the Chilka Lake, India. **Journal of the Marine Biological Association of India**, Tamil Nadu, v. 45, n. 2, p. 133-143, 2003.

JORISSEN, F. J. Benthic foraminiferal successions across Late Quaternary Mediterranean sapropels. **Marine Geology**, Amsterdam, v. 153, n. 1/4, p. 91-101, Jan. 1999.

JOURNAL, A. G.; HUIJBREGTS, C. J. **Mining geostatistics**. London: Academic, 1978. 600 p.

KOUTSOUKOS, E. A. M. et al. Upper aptian-lower coniacian carbonate sequences in the Sergipi Basin, Northeastern Brazil. In: SIMO, J. A. T.; SCOTT, R. W.; MASSE, J. P. (Ed.). **Cretaceous carbonate platforms**. Tulsa: The American Association of Petroleum Geologists, 1993. p. 127-144.

LANDIM, P. M. B. **Análise estatística de dados geológicos multivariados**. São Paulo: Oficina de Textos, 2011. 208 p.

LATTIN, J.; CARROLL, J. D.; GREEN, P. E. **Análise de dados multivariados**. São Paulo: Cengage Learning, 2011. 455 p.

LEGENDRE, P. Spatial autocorrelation: trouble or new paradigm? **Ecology**, Durham, v. 74, n. 6, p. 1659-1673, Sept. 1993.

LEGENDRE, P.; FORTIN, M. J. Spatial pattern and ecological analysis. **Vegetatio**, The Hague, v. 80, p. 107-138, 1989.

LIMAS, J. F.; BONETTI, C.; BONETTI, J. Relações espaciais e morfossedimentares dos foraminíferos no subsistema sul da lagoa da Conceição, Florianópolis, SC. In: CONGRESSO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS DO QUATERNÁRIO ABEQUA, 13., 2011, Niterói. **Anais...** Niterói: ABEQUA, 2011. p. 1-5.

MACHADO, A.; ARAÚJO, H. A. B. Relação entre a microfauna de foraminíferos e a granulometria do sedimento do Complexo Recifal de Abrolhos, Bahia, a partir de análises multivariadas. **Revista Brasileira de Geociências**, São Paulo, v. 42, n. 3, p. 547-562, set. 2012.

MCBRATNEY, A. G.; WEBSTER, A. G. Choosing functions for semi-variograms and fitting them to sampling estimates. **Journal of Soil Science**, Oxford, v. 37, p. 617-639, 1987.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2005. 297 p.

MOJENA, R. Hierarchical grouping methods and stopping rules: an evaluation. **Computer Journal**, London, v. 20, p. 359-363, 1977.

PARIZOTTO, B. A. D. M. et al. Caracterização da zona estuarina do Rio Biguaçu (SC) de parâmetros físico-químicos da água, associados a foraminíferos e tecamebas bentônicos. In: _____. **Oceanografia e mudanças globais**. São Paulo: Instituto Oceanográfico da USP, 2008. p. 561-572.

REIS, E. **Estatística multivariada aplicada**. Lisboa: Silabo, 1997. 342 p.

RENCHE, C. A. **Methods of multivariate analysis**. New York: J. Wiley, 2002. 708 p.

RIBEIRO JÚNIOR, P. J. **Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo**. 1995. 99 f. Dissertação (Mestrado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 1995.

SANTOS, P. T. et al. Análise multivariada de dados ecológicos da baía da Guanabara, RJ. **Anuário do Instituto de Geociências**, Rio de Janeiro, v. 30, n. 1, p. 101-108, 2007.

SILVA, A. P. **Variabilidade espacial de atributos físicos do solo**. 1988. 105 f. Tese (Doutorado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz, Piracicaba, 1988.

SOUZA, L. C. de. **Variabilidade espacial da salinidade de um solo aluvial no semi-árido paraibano**. 1999. 77 p. Dissertação (Mestrado em Agronomia) - Universidade Federal da Paraíba, Campina Grande, 1999.

TELFORD, R. J.; BIRKS, H. J. B. Evaluation of transfer functions in spatially structured environments. **Quaternary Science Reviews**, Oxford, v. 28, n. 13/14, p. 1309-1316, 2009.

_____. The secret assumptions of transfer functions. **Quaternary Science Reviews**, Oxford, v. 24, n. 20/21, p. 2173-2179, 2005.

TOBLER, W. R. Cellular geography. In: GALE, S.; OLSSON, G. (Ed.). **Philosophy in geography**. Dordrecht: D. Reidel, 1979. p. 379-386.

TRANGMAR, B. B.; YOST, R. S.; UEHARA, G. Application of geostatistics to spatial studies of soil properties. **Advances in Agronomy**, San Diego, v. 38, p. 45-94, 1985.

VICALVI, M. A.; FERREIRA, C. S. A ocorrência de um macroforaminífero (Soritidae) na formação (N-NE) do Brasil. **Anais da Academia Brasileira de Ciências**, Rio de Janeiro, v. 64, n. 1, p. 99-100, 1992.

VIEIRA, F. S. **Caracterização microbiofaciológica e ambiental com base em foraminíferos bentônicos na plataforma continental da bacia de Campos, Rio de Janeiro**. 2011. 137 p. Tese (Doutorado em Geologia) - Universidade Federal da Bahia, Salvador, 2011.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R. F. de; ALVAREZ, V. H.; SCHAEFER, C. E. G. R. (Ed.). **Tópicos em ciências do solo**. Viçosa, MG: SBCS, 2000. p. 1-54.

_____. Uso da geoestatística em estudos de variabilidade espacial. In: INSTITUTO AGRONÔMICO. **Curso de atualização em conservação do solo**. Campinas, 1995. p. 61.

WARD, J. H. Hierarchical grouping to optimize an objective function. **Journal of the American Statistical Association**, Alexandria, v. 58, p. 236, 1963.

WETMORE, K. L. Chamber repair and growth in foraminifera-evidence for differences in developmental controls. **Museum of Paleontology**, Berkeley, n. 38, p. 4720-4780, 1998.

YANKO, V.; KRONFELD, J.; FLEXER, A. Response of benthic foraminifera to various pollution source: implications for pollution monitoring. **Journal of Foraminiferal Research**, Washington, v. 24, n. 1, p. 1-17, 1994.

ZIMBACK, C. R. L. **Geoestatística**. Botucatu: UNESP, 2003. 25 p.

SEGUNDA PARTE - ARTIGOS

**ARTIGO 1 Variabilidade espacial de foraminíferos nos sedimentos
da plataforma e talude continentais do estado da Bahia**

Artigo normalizado de acordo com a NBR 6022 (ABNT, 2003).

A ser submetido para a *Revista Pesquisas em Geociências*.

**Variabilidade espacial de foraminíferos nos sedimentos da
plataforma e talude continentais do Estado da Bahia**

Ângela Cristina da Fonseca MIRANTE¹

João Domingos SCALON²

Tânia Maria Fonseca ARAÚJO³

Enio Júnior SEIDEL⁴

RESUMO

Objetivou-se com este trabalho avaliar a presença de variabilidade espacial no número de indivíduos das espécies foraminíferos encontradas nos sedimentos superficiais da plataforma e talude continentais do litoral norte da Bahia e realizar predições por meio de krigagem. Foram utilizados os estimadores de Matheron e de Cressie & Hawkins para obter os semivariogramas das variáveis em estudo. Os resultados mostraram que as espécies *Amphistegina lessonii*, *Peneroplis carinatus*, *Quinqueloculina bicarinata* e *Reusella spinulosa* apresentaram estruturas de dependência espacial descrita pelos modelos teóricos esférico e exponencial. Esses resultados permitiram construir mapas de krigagem da distribuição espacial dessas espécies na região de estudo.

Palavras-chave: Ecossistemas. Geoestatística. Semivariância. Krigagem.

-
- 1 Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA), Departamento de Ciências Aplicadas (DCA), Rua Emídio dos Santos, s/nº - Barbalho, CEP: 40300-010, Salvador, Bahia, Brasil. angelamirante@ifba.edu.br.
 - 2 Universidade Federal de Lavras (UFLA), Departamento de Exatas (DEX), Campus da UFLA, Caixa Postal 3037, CEP 37200-000. Lavras, Minas Gerais, Brasil. scalon@dex.ufla.br;
 - 3 Universidade Federal da Bahia (UFBA), Grupo de Estudos de Foraminíferos, Instituto de Geociências (IGEO), Rua Caetano Moura, 123 – Federação, CEP 40210-340, Salvador, Bahia, Brasil. tfaraujo@ufba.br.
 - 4 Universidade Federal do Pampa (UNIPAMPA), Rua Luiz Joaquim de Sá Britto, s/nº- Bairro Promorar. CEP 97650-000, Itaqui, Rio Grande do Sul, Brasil. enioseidel@unipampa.edu.br.

1 INTRODUÇÃO

Estudos revelaram que existe uma grande fauna de foraminíferos presentes nos sedimentos de superfície e subsuperfícies da plataforma e no talude continental. No Litoral Norte do Estado da Bahia, Araújo (2004) identificou assembléias de foraminíferos que podem ser utilizadas como bioindicadoras ambientais, paleoclimáticas e paleobatimétricas.

Para descrição biológica das associações de foraminíferos, geralmente, é utilizada uma metodologia clássica de análises estatísticas descritivas e medidas tais como número de indivíduos, número de espécies, índice de riqueza de Margalef, índice de diversidade de Shannon-Wiener e índice de equitatividade de Pielou (PERRY, 1998). Observa-se também, um grande uso de técnicas estatísticas multivariadas tais como: análise de variância multivariada, análise fatorial e análise de agrupamentos (LIMAS; BONETTI; BONETTI, 2011).

A análise da dependência espacial é defendida por vários autores como de fundamental importância na análise de assembleias de foraminíferos (LEGENDRE, 1993). Essa importância é devida não somente porque a dependência espacial é uma condição biológica natural dessas espécies, mas também porque a presença de dependência espacial entre as amostras coletadas pode acarretar na quebra da suposição de independência que é de fundamental importância para o uso da maioria dos métodos estatísticos (LEGENDRE, 1993; LEGENDRE; FORTIM, 1989). Entretanto, o que se observa é que muitas das análises para caracterizar assembleias de foraminíferos não levam em conta o estudo da variabilidade (dependência) espacial (TELFORD; BIRKS, 2005).

Apesar de existir alguns estudos que analisam a variabilidade espacial de foraminíferos em outros países (TELFORD; BIRKS, 2009), esse tipo de estudo ainda não foi realizado no Brasil. Em geral, o que se observa é a aplicação de métodos geoestatísticos em dados de sedimentos de foraminíferos para obter mapa de perfis de salinidade e turbidez de zona estuarina (PARIZOTTO et al., 2008).

Diante do exposto, este trabalho tem como objetivo utilizar a análise geoestatística para avaliar a estrutura de dependência espacial e construir um mapa com as estimativas do número de indivíduos das espécies *Amphistegina lessonii*, *Cibicides pseudoungerianus*, *Peneroplis bradyi*, *Peneroplis carinatus*, *Quinqueloculina bicarinata*, *Quinqueloculina disparilis curta* e *Reussella spinulosa* nos sedimentos presentes na plataforma e no talude continentais do norte do Estado da Bahia.

2 MATERIAL E MÉTODOS

2.1 Localização da área e das amostras

A área de estudo deste trabalho refere-se a um trecho da plataforma continental norte do Estado da Bahia, situado entre a praia da Barra, no município de Salvador, e a foz do rio Itariri, no município de Mata de São João, entre as coordenadas longitudinais 37°30'W e 38°30'W e latitudinais 12°S e 13°8'S. Essa área tem uma extensão costeira de aproximadamente 131 km, abrangendo as plataformas interna, média e externa e o talude continental superior, com área de 2.227 km².

As coletas dos dados foram realizadas em trinta e oito estações de coleta. Cada ponto de coleta foi georeferenciado e diversos dados biológicos foram coletados tais como: número de espécies, contagem de indivíduos por espécie, profundidade e informações sedimentológicas da área de estudo.

Esses dados amostrais foram coletados durante a operação GEOMAR XXV em 1985 e estão arquivadas no Instituto de Geociências da Universidade Federal da Bahia.

Para expressar a fisiografia (estreita) e a batimetria (rasa) da área em estudo, as amostras foram agrupadas de acordo com a profundidade em que foram aleatoriamente coletadas, em subambientes delimitados da seguinte maneira: a plataforma interna de 0 a 20 m, contendo cinco amostras; a plataforma média de > 20 a 40 m, contendo 16 amostras; a plataforma externa de > 40 a 60 m, contendo 11 amostras; e o talude superior de 80 a 300 m, com seis amostras.

A área com profundidade de 60 a 80 m corresponde a quebra da plataforma continental, não sendo possível a coleta de amostras por ser de difícil acesso.

Na Figura 1 mostra-se a região de estudo com as localizações espaciais dos pontos de coleta, enquanto na Tabela 1 têm-se as localizações geográficas e marinha e profundidade das amostras.

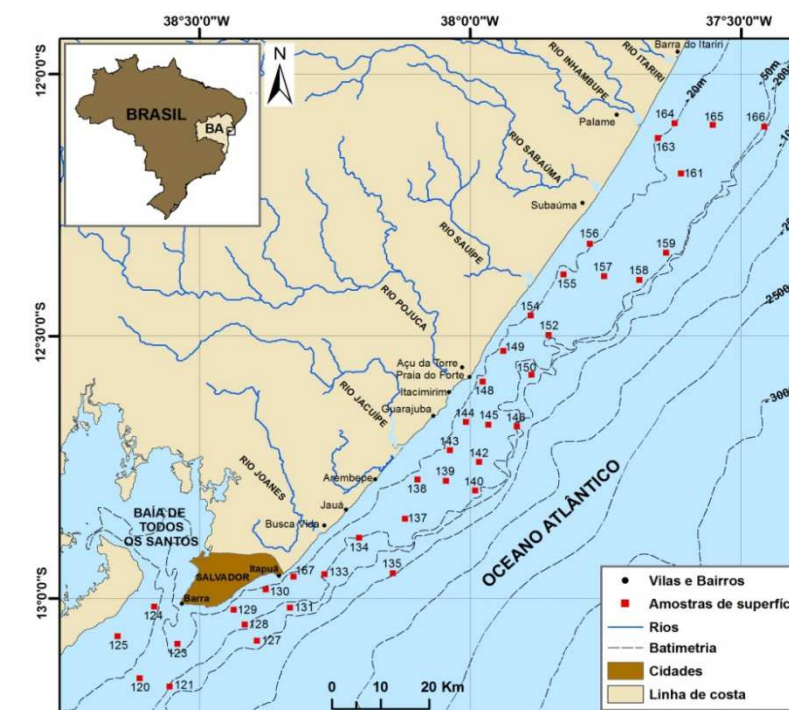


Figura 1 Localização da área de estudo e das estações de coleta das amostras
Fonte Araújo e Araújo (2010)

Tabela 1 Profundidade, localização e posição geográfica das 38 amostras na região de estudo

Amostras	Prof. (m)	Local	Latitude (S)	Longitude (W)
123	14	Plataforma interna	13° 05' 31"	38° 32' 02"
124	15	Plataforma interna	13° 01' 36"	38° 34' 53"
125	17,7	Plataforma interna	13° 03' 26"	38° 23' 50"
156	19	Plataforma interna	12° 19' 56"	37° 48' 08"
164	20	Plataforma interna	12° 03' 24"	37° 36' 57"
155	21,5	Plataforma média	12° 22' 46"	37° 49' 50"
143	22,1	Plataforma média	12° 43' 26"	38° 03' 37"
154	23	Plataforma média	12° 27' 32"	37° 53' 15"
138	26	Plataforma média	12° 47' 39"	38° 07' 03"
144	26,4	Plataforma média	12° 40' 13"	38° 00' 44"
133	27	Plataforma média	12° 58' 25"	38° 18' 10"
161	28	Plataforma média	12° 11' 42"	37° 36' 53"
163	28	Plataforma média	12° 06' 13"	37° 38' 50"
148	29	Plataforma média	12° 35' 58"	37° 58' 57"
149	30	Plataforma média	12° 32' 01"	37° 56' 41"
165	32	Plataforma média	12° 04' 16"	37° 32' 25"
130	32,5	Plataforma média	12° 59' 37"	38° 24' 32"
120	36,3	Plataforma média	13° 08' 19"	38° 34' 58"
129	37	Plataforma média	13° 02' 45"	38° 28' 04"
157	39	Plataforma média	12° 23' 10"	37° 44' 38"
134	40	Plataforma média	12° 52' 56"	38° 14' 13"
142	43	Plataforma externa	12° 45' 27"	37° 00' 28"
139	44	Plataforma externa	12° 48' 07"	38° 04' 40"
145	44	Plataforma externa	12° 41' 53"	38° 58' 37"
158	44	Plataforma externa	12° 23' 28"	37° 40' 44"
159	47	Plataforma externa	12° 20' 17"	37° 37' 42"
140	48	Plataforma externa	12° 49' 56"	38° 01' 24"
167	49	Plataforma externa	12° 59' 15"	37° 21' 05"
137	50	Plataforma externa	12° 51' 21"	38° 09' 24"
152	51	Plataforma externa	12° 29' 50"	37° 50' 46"
150	57	Plataforma externa	12° 34' 57"	37° 52' 35"
128	60	Plataforma externa	13° 04' 56"	38° 26' 58"
166	80	Talude	12° 05' 15"	37° 27' 14"
146	120	Talude	12° 43' 30"	37° 55' 35"
131	150	Talude	13° 02' 20"	38° 20' 46"
127	190	Talude	13° 07' 07"	38° 26' 07"
121	218,5	Talude	13° 11' 04"	38° 33' 15"
135	300	Talude	12° 58' 22"	38° 11' 35"

2.2 Dados da fauna de foraminíferos

De cada amostra de superfície do fundo foram retirados, 10 cm³ de material que foi lavado, em água destilada por meio de uma peneira com espaçamento de malha 0,062 mm para eliminação dos sais e separação da fração lama. O material foi colocado para secar em estufa a 50°C, e, posteriormente, foi pesado e quarteado. Em seguida foram separadas, aleatoriamente, as 300 primeiras testas de foraminíferos, para identificação das espécies, por meio de lupa binocular. Esta metodologia descrita por Drooger e Kaasschietter (1958) é considerada adequada quando se utiliza grande número de amostras e se deseja realizar tratamento estatístico (ARAÚJO; ARAÚJO, 2010). Os espécimes foram identificados segundo Loeblich Junior e Tappan (1988) e publicações de foraminíferos recentes tais como Araújo, T. e Machado (2008a, 2008b, 2008c) e Eichler et al. (2012).

A quantidade de indivíduos por espécie foi calculada considerando os 300 espécimes identificados em cada amostra. A partir desses dados, foi calculada a frequência de ocorrência (FO) que é expressa por $FO = \frac{p \times 100}{P}$, em que p é o número de amostras que contêm a espécie e P é o número total de amostras analisadas. de abundância absoluta das espécies (AB´SABER et al., 1997). De acordo com a classificação de Dajoz (1983), as espécies foram agrupadas em três categorias, sendo classificadas como constantes (as espécies presentes em mais de 50 % das amostras); acessórias (que ocorrem entre 25 % a 49 % das amostras) e acidentais (presentes em menos de 25 % das amostras).

Também se obteve a frequência relativa (abundância das espécies) que consiste na razão entre o número total de indivíduos de uma categoria obtidos na amostra (n) e o total de indivíduos de todas as categorias (T) na amostra, sendo expressa, em percentagem, por $FR = \frac{n \times 100}{T}$ (AB´SABER et al., 1997). Este critério determina as espécies abundantes em cada amostra. Os resultados permitiram classificar as espécies, de acordo com critérios de classificação de Dajoz (1983), em principais (frequência acima de 5 %); acessória (entre 5 e 1 %) e traço (inferior a 1 %).

2.3 Análise estatística

Inicialmente foi realizada uma análise exploratória das contagens do número de indivíduos das espécies principais *Amphistegina lessonii*, *Cibicides pseudoungerianus*, *Peneroplis bradyi*, *Peneroplis carinatus*, *Quinqueloculina bicarinata*, *Quinqueloculina disparilis curta* e *Reussella spinulosa*, encontradas na área de estudo, utilizando medidas da estatística descritiva tais como média, mediana, valor mínimo, valor máximo, desvio padrão, assimetria, curtose e coeficiente de variação para obter informações sobre a distribuição das contagens das espécies e a presença de outliers.

Para realizar a análise espacial foi efetuada a conversão das coordenadas geográficas das amostras para UTM (Universal Transversal of Mercator) que é um sistema referencial de localização terrestre baseado em coordenadas métricas.

A estimativa da dependência espacial, entre amostras vizinhas, foi realizada utilizando o semivariograma. Para estimar as semivariâncias, foram utilizados os estimadores de Cressie e Hawkins (1980) e Matheron (1963).

O estimador “clássico” de Matheron (1963), por definição, é dado por

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(s_i) - z(s_i + h)]^2, \quad (1)$$

em que $\hat{\gamma}(h)$ é a semivariância estimada para uma distância h ; $N(h)$ representa o número de pares de valores medidos $z(s_i)$, $z(s_i + h)$ separados por um vetor de distância h . Esse estimador foi utilizado quando as observações não apresentaram *outliers* (valores atípicos).

Nos casos em que se verificou a presença de *outliers*, foi usado o estimador “robusto” de Cressie e Hawkins (1980). Esse estimador é dado por

$$\hat{\gamma}(h) = \frac{1}{2} \frac{\left[N(h)^{-1} \sum_{i=1}^{N(h)} \sqrt{|z(s_i + h) - z(s_i)|} \right]^4}{0,457 + \frac{0,494}{N(h)}}. \quad (2)$$

Nas espécies que apresentaram estrutura de dependência espacial, foi necessário o ajuste de um modelo teórico para expressar tal dependência. Neste trabalho foram testados os modelos esférico,

exponencial e gaussiano descritos em Isaaks e Srivastava (1989) e Vieira (2000).

O grau de dependência espacial (*GDE*) foi usado para expressar a magnitude da dependência espacial e foi medido pela relação entre a variação estruturada (contribuição) e o patamar (BIONDI; MYERS; AVERY, 1994) que é dado por

$$GDE = \frac{C_1}{C_0 + C_1}. \quad (3)$$

Segundo os valores encontrados para o *GDE* dado em (5) pode-se classificar o grau de dependência espacial do atributo em estudo como alto, se $GDE > 75\%$; médio se $35 \leq GDE \leq 75\%$; e fraco se $GDE < 35\%$.

A escolha do modelo teórico de semivariograma foi realizada pela validação cruzada e também pelo critério de menor erro quadrático médio (EQM).

Segundo Isaaks e Srivastava (1989), a validação cruzada é uma técnica de avaliação de erros de estimativas que permite comparar os valores previstos com os amostrados. O valor da amostra, em certa localização, $z(s_i)$, é temporariamente descartado do conjunto de dados e, então, é feita uma previsão por krigagem no local $\hat{z}(s_{(i)})$, usando-se as amostras restantes. Assim, o Erro Médio por validação cruzada (EM) é obtido da equação:

$$EM = \frac{1}{n} \sum_{i=1}^n (z(s_i) - \hat{z}(s_{(i)})). \quad (4)$$

A diferença entre o valor real e a previsão do valor dá origem ao erro de previsão. O erro quadrático médio (*EQM*) também pode ser usado como uma medida do erro de previsão e é expresso pela seguinte equação:

$$EQM = \frac{1}{n} \sum_{i=1}^n [(z(s_i) - \hat{z}(s_{(i)}))]^2, \quad (5)$$

em (6) e (7), n é o número de dados; $z(s_i)$, valor observado no ponto s_i e $\hat{z}(s_{(i)})$, valor predito por krigagem ordinário no ponto s_i , sem considerar a observação $z(s_i)$.

Segundo Vieira (2000), conhecido o modelo teórico do semivariograma e havendo dependência espacial entre as amostras podem-se interpolar valores em qualquer posição no campo de estudo, sem tendência e com variância mínima utilizando-se a krigagem. De acordo com Andriotti (2009), o estimador de krigagem é do tipo BLUE (best linear unbiased estimator) e é dito linear por ser formado por uma combinação linear dos dados. O estimador é dado por

$$Z^*(k) = \sum_{i=1}^n \lambda_i \cdot z(x_i), \quad (6)$$

em que λ_i são os ponderadores obtidos a partir da estrutura dependência espacial dados pelos modelos teóricos de semivariâncias, $z(x_i)$ são os dados experimentais e n o número total desses dados.

Toda a análise estatística foi realizada utilizando-se de funções disponíveis e/ou construídas no *software* R (R DEVELOPMENT CORE TEAM, 2012) e a biblioteca *geoR* (RIBEIRO JÚNIOR; DIGGLE, 2001) do mesmo *software*.

3 RESULTADOS

Na Tabela 2 apresentam-se os resultados referentes à análise exploratória das contagens das espécies foraminíferos *Amphistegina lessonii* (Ampl), *Cibicides pseudoungerianus* (Cibpseu), *Peneroplis bradyi* (Penbra), *P. carinatus* (Pencar), *Quinqueloculina bicarinata* (Quibicar), *Q. disparilis curta* (Quidiscur) e *Reusella spinulosa* (Reuspin).

Os coeficientes de variação apresentados na Tabela 2 permitiram afirmar que há elevada variabilidade dos dados em relação à média do número de indivíduos das espécies foraminíferos, mostrando que os dados têm alta dispersão.

Na Tabela 2 também é possível observar a presença de assimetria à direita. Esses resultados podem indicar a presença de *outliers* e/ou ausência de normalidade nas variáveis analisadas. Entretanto, a normalidade dos dados não é uma exigência para estimar o semivariograma empírico, mas é conveniente que a distribuição não apresente caudas muito longas, como é o caso, de forma a não comprometer as análises (CRESSIE, 1991).

Tabela 2 Medidas descritivas da contagem do número de indivíduos por espécie

Espécies	\bar{X}	md	X_1	X_n	S	$\sqrt{b_1}$	b_2	CV
<i>Amplex</i>	25,26	16,5	1	112	25,62	1,6	3,21	101,42
<i>Cibpseu</i>	6,08	4	0	46	7,8	3,65	18,97	128,36
<i>Penbra</i>	12,71	10,5	0	47	10,8	1,39	2,67	85,01
<i>Pencar</i>	31,55	22	0	142	33,55	1,49	2,6	106,36
<i>Quibicar</i>	6,58	6	0	18	4,43	0,63	0,25	67,39
<i>Quidiscur</i>	6,29	6	0	17	4,84	0,54	-0,7	76,91
<i>Reuspin</i>	5,4	5	0	15	4,21	0,58	-0,46	78,12

Média (\bar{X}), mediana (md), valor mínimo (X_1), valor máximo (X_n), desvio-padrão (s), assimetria ($\sqrt{b_1}$), curtose (b_2) e coeficiente de variação (CV).

Uma maneira mais eficiente de verificar a presença de *outliers* é por meio da análise dos *box plot* apresentados nas Figuras 2 e 3.

Uma inspeção visual realizada nos gráficos *box plot*, apresentados na Figura 2, não mostra a presença de *outliers* nas espécies *Quinqueloculina bicarinata*, *Q. disparilis curta* e *Reusella spinulosa*. Assim, para essas espécies, pode-se obter as semivariâncias empíricas utilizando o estimador clássico de Matheron (equação 1).

A Figura 3 mostra que as espécies *Amphistegina lessonii*, *Peneroplis carinatus*, *P.bradyi* e *Cibicides pseudoungerianus* apresentam *outliers* e, portanto, para essas espécies será utilizado o estimador robusto de Cressie e Hawkins (1980) (equação 2).

As semivariâncias empíricas e os respectivos modelos teóricos ajustados para as espécies de foraminíferos são apresentados nas Figuras de 4 a 6.

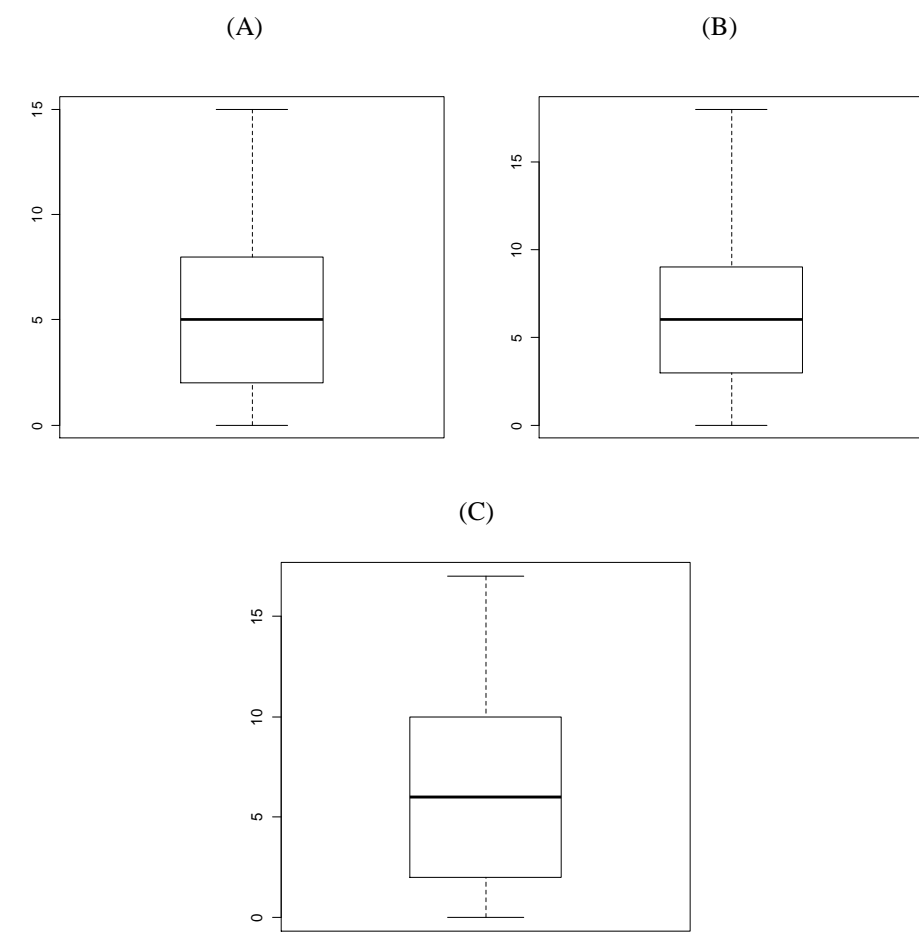


Figura 2 Gráficos boxplot para as espécies *Reusella spinulosa* (A); *Quinqueloculina bicarinata* (B) e *Quinqueloculina disparilis curta* (C)

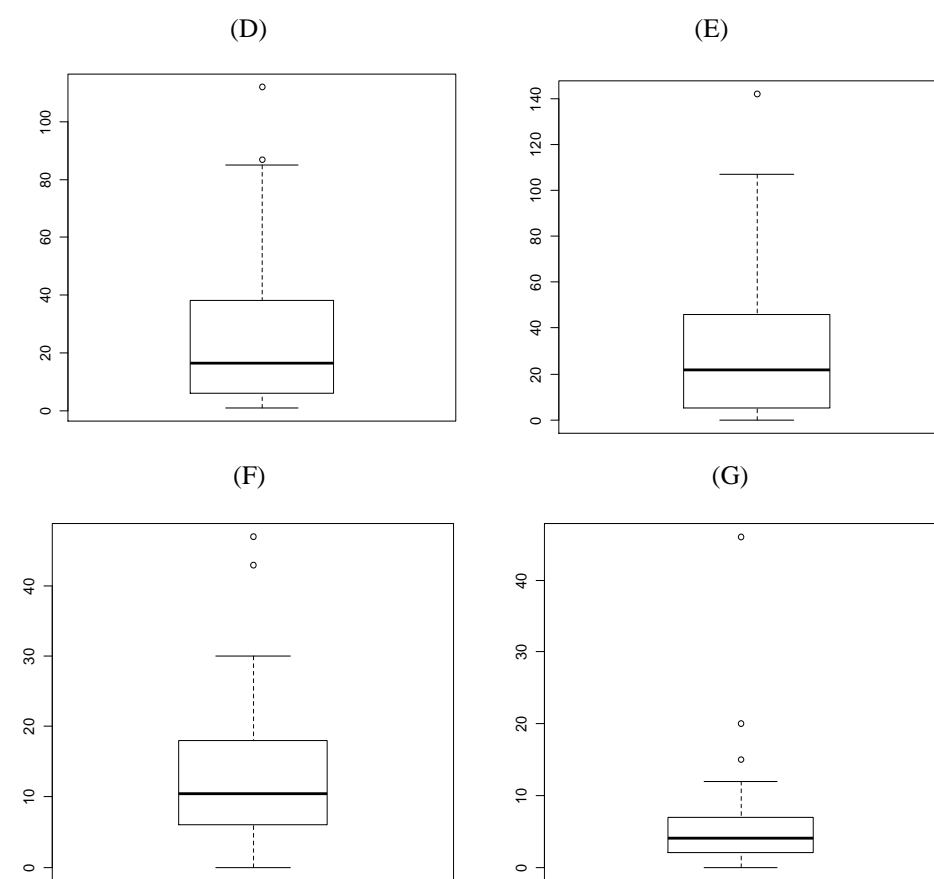


Figura 3 Gráficos boxplot para as espécies *Amphistegina lessonni* (D); *Peneroplis carinatus* (E); *Peneroplis bradyi* (F) e *Cibicides pseudoungerianus* (G)

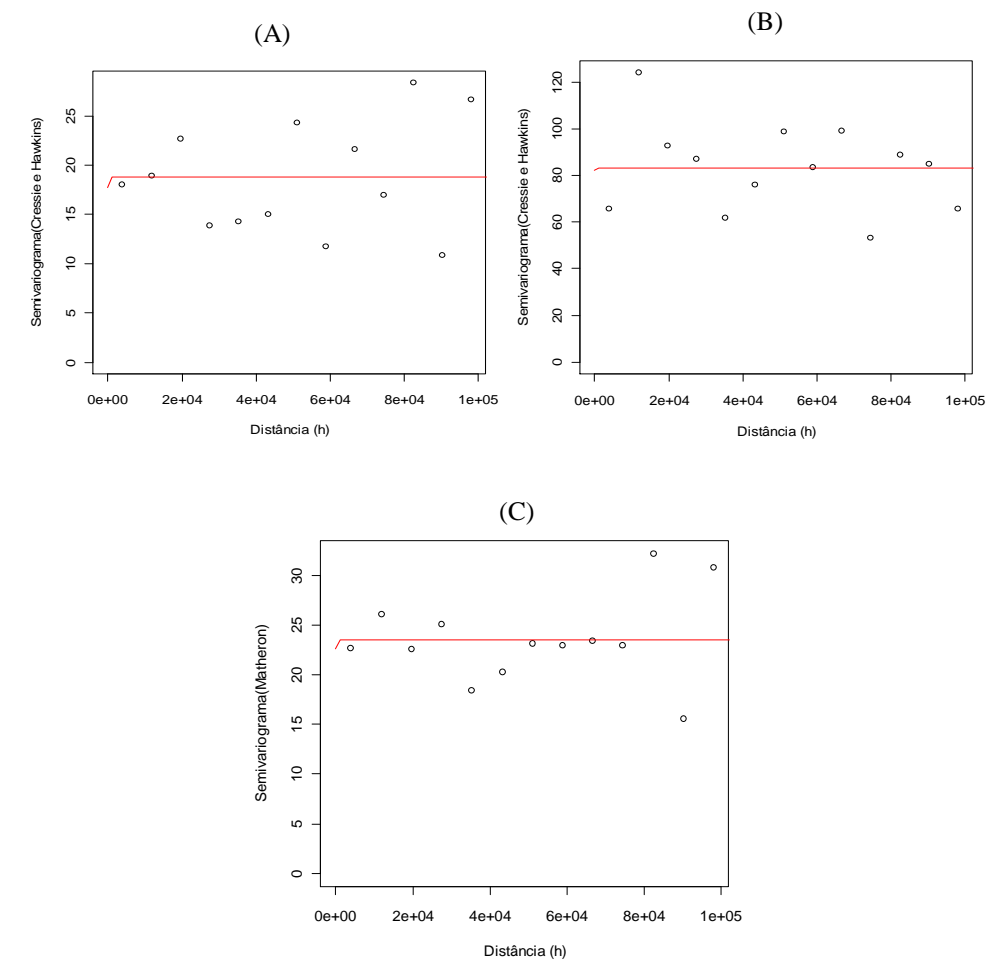


Figura 4 Modelo efeito pepita puro (linha contínua) ajustado ao semivariograma experimental (pontos) para as espécies *Cibicides pseudoungerianus* (A); *Peneroplis bradyi* (B) e *Quinqueloculina disparilis curta* (C)

Os semivariogramas apresentados na Figura 4 mostram a presença de efeito pepita puro e, portanto, indicam que não é possível identificar a estrutura de dependência espacial quanto ao número de indivíduos das

espécies *Cibicides pseudoungerianus*, *Peneroplis bradyi* e *Quinqueloculina disparilis curta*. Dessa forma, pode-se assumir que a distribuição é completamente ao acaso, há independência entre amostras e os métodos da estatística clássica podem ser aplicados. Vieira (2000) adverte que, no entanto, esses resultados não indicam, necessariamente, que não exista dependência espacial nessas espécies. Pode ocorrer, nesses casos, que a dependência espacial ocorre em uma distância menor do que a menor distância entre os pontos de amostragem utilizada.

Foi observado que o número de indivíduos da espécie *Peneroplis carinatus* apresentou dependência espacial que pode ser modelada pelo modelo exponencial (Figura 5), enquanto as espécies *Amphistegina lessonii*, *Quinqueloculina bicarinata* e *Reusella spinulosa* apresentaram dependência espacial que pode ser modelada pelo modelo esférico (Figura 6).

Os modelos exponencial e esférico foram escolhidos segundo os critérios do menor erro quadrático médio e o da validação cruzada.

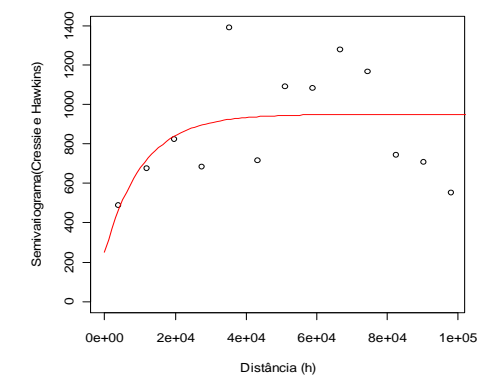


Figura 5 Modelo exponencial (linha contínua) ajustado ao semivariograma experimental (pontos) para a espécie *Peneroplis carinatus*

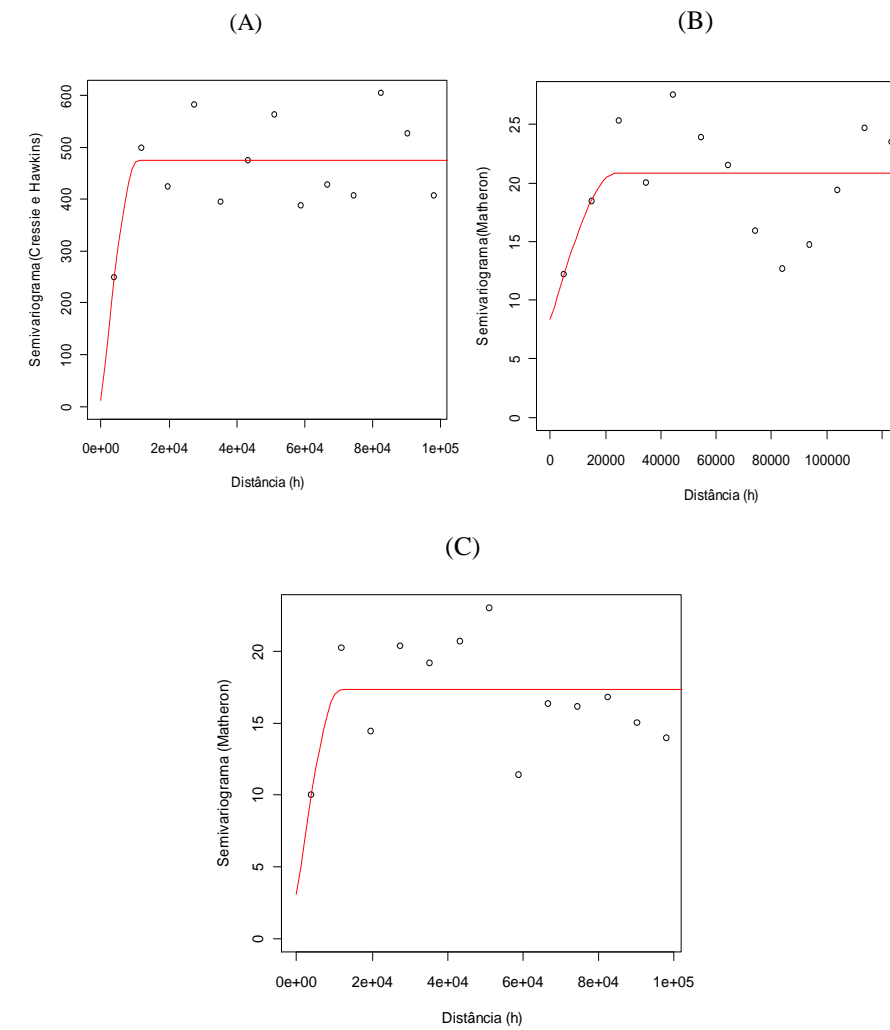


Figura 6 Modelo esférico (linha contínua) ajustado ao semivariograma experimental (pontos) para as espécies *Amphistegina lessonii* (A); *Quinqueloculina bicarinata* (B) e *Reusella spinulosa* (C)

Os parâmetros ajustados indicam que, até o valor do alcance estimado (em metros), existe forte dependência espacial no número de indivíduos para as espécies *Amphistegina lessonii*, *Reusella spinulosa* e

Peneroplis carinatus e moderada dependência para a espécie *Quinqueloculina bicarinata* (Tabela 3).

Tabela 3 Parâmetros estimados no semivariograma

Variável	Efeito pepita	Patamar	Alcance	Grau de dependência espacial (GDE) (%)*
<i>Amphistegina lessonii</i>	10,86	475,46	10985,91	98
<i>Peneroplis carinatus</i>	250,07	950,86	32076,1	74
<i>Quinqueloculina bicarinata</i>	8,38	20,85	23962,03	60
<i>Reusella spinulosa</i>	3,1	17,32	11605,7	82

*GDE=(Contribuição/Patamar)*100

As Figuras de 7 a 10 representam os mapas de krigagem obtidos com base na estrutura de dependência espacial descrita pelos modelos teóricos, indicando valores estimados em posições não amostradas na área de estudo.

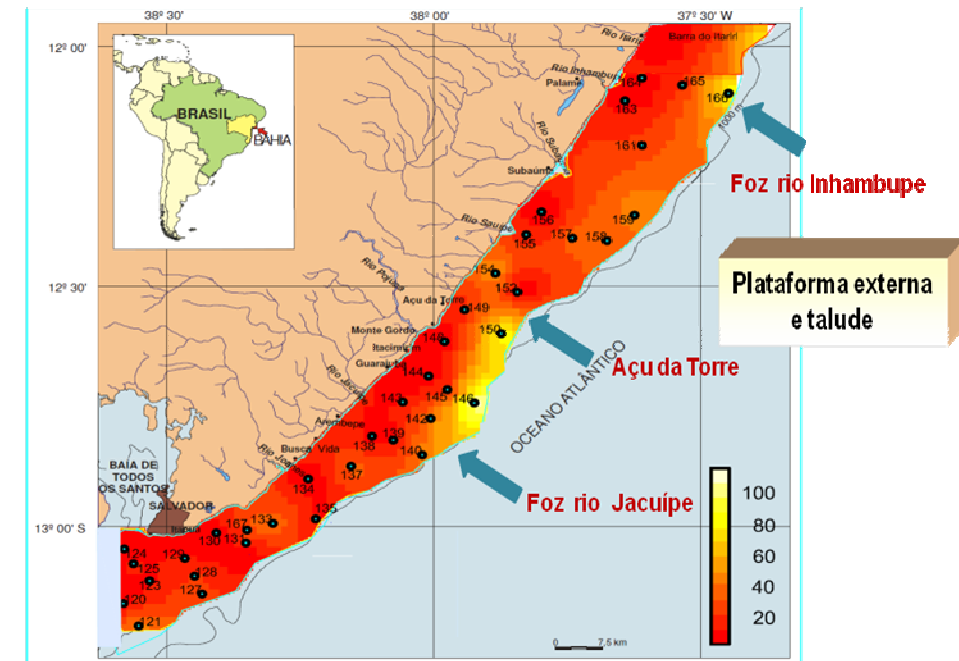


Figura 7 Mapa de krigagem do número de indivíduos da espécie *Amphistegina lessonii*

Na Figura 7 observa-se que a região de predição com os valores mais altos de indivíduos da espécie *Amphistegina lessonii* foram encontradas no talude (80 a 100 espécimes) e na plataforma externa (em torno de 80 espécimes), abrangendo a área em frente a foz do Rio Jacuípe até Açú da Torre e em frente a foz do Rio Inhambupe.

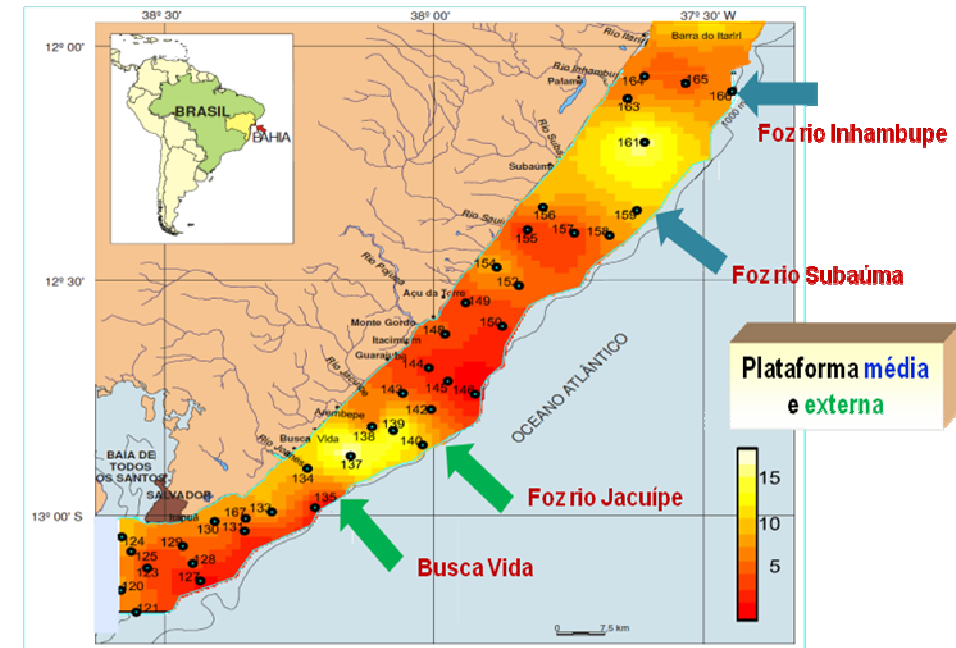


Figura 8 Mapa de krigagem do número de indivíduos da espécie *Quinqueloculina bicarinata*.

Na Figura 8 observa-se que os valores mais altos estimados para o número de indivíduos da espécie *Quinqueloculina Bicarinata* encontram-se na plataforma externa (>15 espécimes) abrangendo as áreas de Busca Vida até o Rio Jacuípe e na plataforma média (10 a 15 espécimes) entre a foz do Rio Subaúma e a foz do Rio Inhambupe.

Na Figura 9 observa-se que a área estimada com o maior número de indivíduos da espécie *Reusella spinulosa* abrange a plataforma média (10 a 15 espécimes) nas áreas de Lauro de Freitas, Praia do Forte, Rio

Sauípe e entre os Rios Subaúma e Inhambupe e na plataforma externa (>14 espécimes) na área em frente a Jauá.

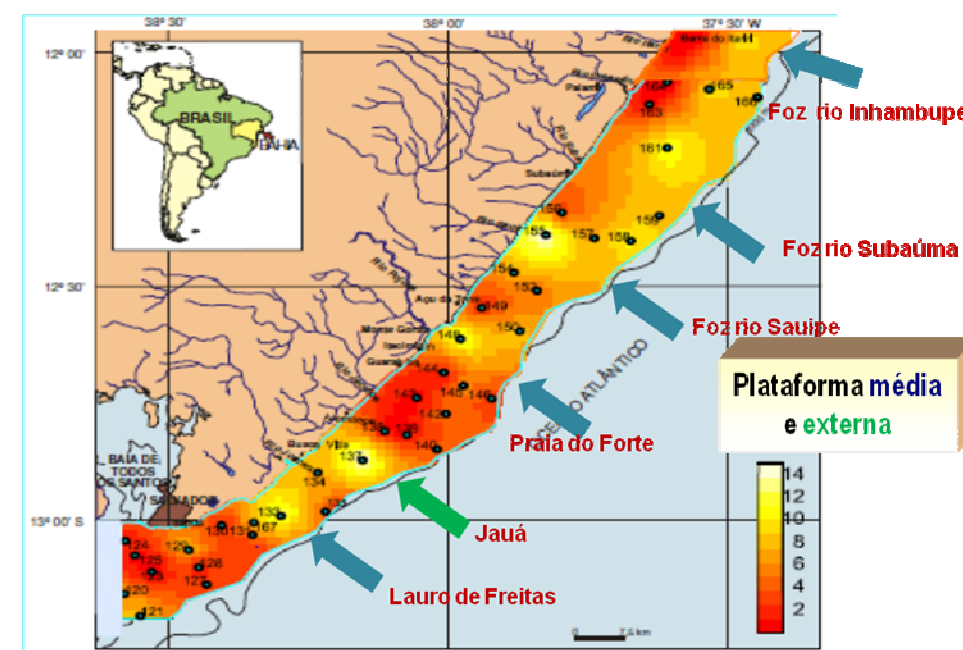


Figura 9 Mapa de krigagem do número de indivíduos da espécie *Reusella spinulosa*.

A espécie *Peneroplis carinatus* apresentou maior número de indivíduos na plataforma média (>120 espécimes) nas áreas de Arembepe e Itacimirim e na plataforma interna (> 100 espécimes) entre os Rios Sauípe e Subaúma (Figura 10).

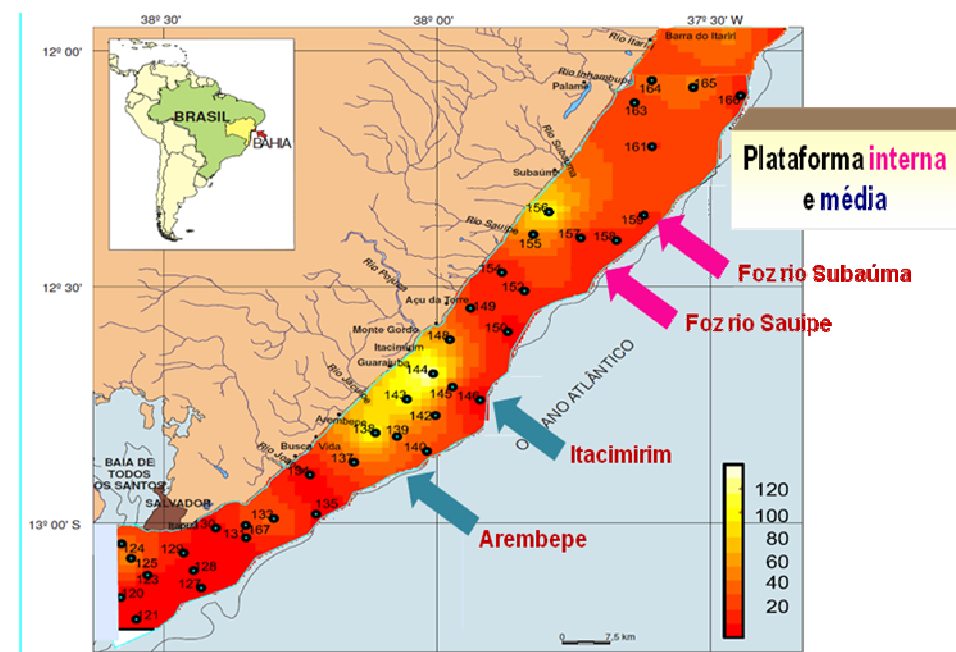


Figura 10 Mapa de krigagem do número de indivíduos da espécie *Peneoplis carinatus*

4 DISCUSSÃO

A análise da dependência espacial é considerada, por vários autores, como de extrema importância em estudos ecológicos (LEGENDRE, 1993; LEGENDRE; FORTIN, 1989). Telford e Birks (2005) afirmam que esse tipo de análise deveria ser realizada antes de ser aplicada qualquer outro tipo de análise estatística que requeira a independência entre as amostras. Além disso, a análise da dependência espacial é absolutamente necessária para um melhor entendimento da distribuição espacial de espécies. Entretanto, o que se observa na literatura é a presença de poucos estudos que utilizam métodos de estatística espacial para análise de dados de assembleias de espécies de foraminíferos (PARIZOTTO et al., 2008; TELFORD; BIRKS, 2009).

A espécie *Amphistegina lessonii* foi encontrada na plataforma média por Machado, Andrade e Araújo (2006) na margem costeira até 45m de profundidade no litoral norte do Estado da Bahia entre Itapuã e Praia do Forte. Figueiredo et al. (2011) também constatou expressiva abundância de *Amphistegina lessonii* em plataforma média a externa do município de Conde (BA).

No talude, Leipnitz e Leipnitz (1996) relatam a presença de formas vivas da espécie *Amphistegina lessonii* com batimetria máxima de 129m. *Amphistegina lessonii* é encontrada nas praias de Salvador, sendo comum em plataformas e áreas recifais do litoral norte da Bahia (ARAÚJO; ARAÚJO, 2010; MACHADO; ANDRADE; ARAÚJO, 2006). Araújo, H. e Machado (2008) assinalam que a maior abundância dessa espécie encontra-se no talude em frente a recifes da praia de

Guarajuba, no litoral norte da Bahia. Em Porto Rico ela é relatada como uma das espécies mais abundantes entre as assembleias recifais (SEIGLIE, 1968). A espécie *Amphistegina lessonii* ocorre também no Indo-Pacífico e Mediterrâneo do sul, alcançando 0-90m de profundidade (MURRAY, 2006) e no Hawaii é abundante em áreas com menos de 10m de profundidade (HALLOCK, 1984).

A maior abundância de *Peneroplis carinatus* em plataforma interna foi observada por Figueiredo et al. (2011) em estudo de análise do zoneamento batimétrico das assembleias de foraminíferos da plataforma e do talude continentais do município de Conde (BA). Bruno, Araújo e Machado (2009) também encontraram as espécies *Amphistegina lessonii* e *Peneroplis carinatus* em locais cuja batimetria estava entre 40-50m na plataforma de Ilhéus na Bahia.

Na plataforma média a espécie *Reusella spinulosa* foi encontrada por Machado, Andrade e Araújo (2006) na margem costeira até 45m de profundidade no litoral norte do Estado da Bahia entre Itapuã e Praia do Forte. Macedo (1994) também observou a espécie *Reusella spinulosa* nas praias do litoral norte do Estado da Bahia e Silva (2004) registrou essa espécie entre as praias de Barra e Piatã, na profundidade de 40m a 60 m. A espécie *Reusella spinulosa* foi encontrada por Moraes (2006) com expressiva abundância na plataforma externa (42m a 44m) ao norte da ilha de Itanharé e na plataforma média a uma profundidade de 30m em Morro de São Paulo. Bruno, Araújo e Machado (2009) encontraram o gênero *Reusella* com abundância, entre 30m a 100m de profundidade.

Moraes (2006) analisando a distribuição de foraminíferos nos sedimentos da foz do Rio Jequiriçá à Ponta dos Castelhanos, na Costa do

Dendê, encontrou a espécie *Quinqueloculina bicarinata* em amostras com 18 e 52 m de profundidades. Silva (2004) registrou a presença de *Q. bicarinata* em sedimentos da plataforma continental na costa atlântica de Salvador entre as praias do farol da Barra até a praia de Piatã, em profundidades de 10 m a 100m, predominantemente nas plataformas média e externa (entre 30 e 60m). Araújo e Machado (2008) realizaram coletas de sedimentos para análise da fauna de foraminíferos associada aos recifes do extremo Sul do Estado da Bahia (Corumbau a Nova Viçosa) e encontrou *Q. bicarinata* nas amostras coletadas na plataforma interna

Finalizando, pode-se afirmar que os resultados desse trabalho evidenciaram que as espécies *Quinqueloculina bicarinata*, e *Reussella spinulosa* predominam nas plataformas média e externa. A espécie *Amphistegina lessonii*, é encontrada em abundância nas áreas do talude e plataforma externa e a espécie *Peneroplis carinatus* predomina nas plataformas interna e média. Os resultados obtidos também mostraram que existe estrutura de dependência espacial nas contagens do número de indivíduos das espécies foraminíferos *Amphistegina lessonii*, *Peneroplis carinatus*, *Quinqueloculina bicarinata*, e *Reussella spinulosa* presentes na plataforma e no talude continentais da parte norte do Estado da Bahia. Pode-se sugerir que a razão da inexistência de dependência espacial observada no número de indivíduos das espécies *Cibicides pseudoungerianus*, *Peneroplis bradyi* e *Quinqueloculina. disparilis curta*, seja devido às grandes distâncias entre as amostras que não permitiram identificar a dependência espacial. Portanto, os resultados obtidos neste trabalho corroboram os resultados de Telford e Birks (2005, 2009) que

afirmam existir dependência espacial no número de indivíduos de diferentes espécies de foraminíferos.

5 CONCLUSÃO

A aplicação dos métodos geostatísticos possibilitou detectar que existe dependência espacial, quanto ao número de indivíduos, em quatro espécies de foraminíferos. Além disso, possibilitou com apenas trinta e oito amostras, a construção de mapas que mostram a distribuição espacial dessas espécies nos sedimentos de superfície da plataforma e do talude continental na região do Litoral Norte do Estado da Bahia.

**Spatial variability of foraminifera in sediments of the continental
shelf and slope of the State of Bahia**

ABSTRACT

This work aimed at evaluating the presence of spatial variability in the number of individuals of the foraminifera species found in superficial sediments of the continental shelf and slope of the northern coast of the state of Bahia, Brazil, and make predictions using kriging. We used the Matheron and Cressie & Hawkins estimators in order to obtain the experimental semivariograms of the studied variables. The results showed that the species' *Amphistegina lessonii*, *Peneroplis carinatus*, *Quinqueloculina bicarinata* and *Reusella spinulosa* presented spatial dependence structures described by spherical and exponential theoretical models. These results allowed us to build kriging maps of the spatial distribution of these species in the study area.

Keywords: ecosystems, geostatistics, semivariance, kriging.

REFERÊNCIAS

AB'SABER, A. N. et al. **Glossário de ecologia**. 2. ed. São Paulo: ACIESP, 1997. 352 p.

ANDRIOTTI, J. L. S. **Fundamentos de estatística e geoestatística**. São Leopoldo: UNISINOS, 2009. 165 p.

ARAÚJO, H. A. B.; MACHADO, A. J. Benthic foraminifera associated with the south Bahia coral reefs, Brazil. **The Journal of Foraminiferal Research**, Washington, v. 38, n. 1, p. 23-38, 2008.

ARAÚJO, T. M. F. **Estudo da microfauna de foraminíferos do sedimento da superfície e da subsuperfície da plataforma e do talude continentais da região norte do Estado da Bahia, Salvador à Barra do Itariri**. 2004. 528 p. Tese (Doutorado em Geologia) - Universidade Federal da Bahia, Salvador, 2004.

ARAÚJO, T. M. F.; ARAÚJO, H. A. B. Assembleias de foraminíferos encontradas nos sedimentos superficiais da plataforma continental e talude superior do norte da Bahia. **Revista de Geologia**, Fortaleza, v. 23, n. 1, p. 115-134, 2010.

ARAÚJO, T. M. F.; MACHADO, A. J. M. Análise sedimentar e micropaleontológica (Foraminíferos) de Seções Quaternárias do Talude Continental Superior do Norte da Bahia, Brasil. **Revista Pesquisas em Geociências**, Porto Alegre, v. 35, n. 2, p. 97-113, 2008a.

_____. Eventos bioestratigráficos, paleoclimáticos e paleobatimétricos do talude continental superior da Bahia, evidenciados do estudo da fauna dos Foraminíferos. **Gravel**, Porto Alegre, v. 6, n. 1, p. 27-45, 2008b.

_____. Foraminíferos da superfície do talude continental superior do norte da Bahia, Brasil. **Revista de Geologia**, Fortaleza, v. 21, n. 1, p. 49-77, 2008c.

BIONDI, F.; MYERS, D. E.; AVERY, C. C. Geostatistically modelling stem size and increment in an old growth forest. **Canadian Journal of Forest Research**, Ottawa, v. 24, n. 7, p. 1354-1368, July 1994.

BRUNO, R. L. M.; ARAÚJO, H. A. B.; MACHADO, A. J. Análise das assembleias de foraminíferos no sedimento superficial do Recife de Fora, região sul da Bahia. **Revista Brasileira de Geociências**, São Paulo, v. 39, p. 599-607, dez. 2009.

CRESSIE, N. **Statistics for spatial data**. New York: J. Wiley, 1991. 900 p.

CRESSIE, N. A. C.; HAWKINS, D. M. Robust estimation of the variogram. **Journal of the International Association for Mathematical Geology**, New York, v. 12, p. 115-125, 1980.

DAJOZ, R. **Ecologia geral**. 4. ed. Petrópolis: Vozes, 1983. 472 p.

DROOGER, C. W.; KAASSCHIETTER, J. P. H. Foraminifera of orenoco-trinidad-paria shelf expedition. **Verhandlungen der Kon. Nedel Akademie Wetensch Afd. Natnurk**, Brussels, v. 4, n. 22, p. 108, 1958.

EICHLER, P. P. B. et al. Foraminifera as indicator of marine pollutant contamination on inner continental shelf of southern Brazil. **Marine Pollution Bulletin**, Oxford, v. 64, n. 1, p. 22-30, Jan. 2012.

FIGUEIREDO, S. M. C. et al. Zoneamento batimétrico das assembléias de foraminíferos da plataforma e talude continentais do município de Conde, Bahia. **Gravel**, Porto Alegre, v. 9, n. 1, p. 1-10, 2011.

HALLOCK, P. Distribution of selected species of living algal symbiont-bearing foraminifera on two Pacific coral reefs. **Journal of Foraminiferal Research**, Washington, v. 23, n. 3, p. 162-179, 1984.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University, 1989. 561 p.

LEGENDRE, P. Spatial autocorrelation: trouble or new paradigm? **Ecology**, Durham, v. 74, n. 6, p. 1659-1673, Sept. 1993.

LEGENDRE, P.; FORTIN, M. J. Spatial pattern and ecological analysis. **Vegetatio**, The Hague, v. 80, p. 107-138, 1989.

LEIPNITZ, I. I.; LEIPNITZ, B. Nova proposta de subdivisão biogeográfica com base nos foraminíferos para as porções norte e nordeste da Plataforma Continental Brasileira. In: CONGRESSO BRASILEIRO DE GEOLOGIA, 39., 1996, Salvador. **Resumos Expandidos...** Salvador: UFBA, 1996. p. 270-273.

LIMAS, J. F.; BONETTI, C.; BONETTI, J. Relações espaciais e morfossedimentares dos foraminíferos no subsistema sul da lagoa da Conceição, Florianópolis, SC. In: CONGRESSO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS DO QUATERNÁRIO ABEQUA, 13., 2011, Niterói. **Anais...** Niterói: ABEQUA, 2011. p. 1-5.

LOEBLICH JUNIOR, A. R.; TAPPAN, H. **Foraminiferal genera and their classification**. New York: V. N. Reinhold, 1988. 970 p.

MACEDO, D. L. S. **Análise das assembleias de foraminíferos do sedimento de zona costeira dos municípios de Salvador e Camaçari estado da Bahia**. 1994. 91 p. Dissertação (Mestrado em Geociências) - Universidade Federal da Bahia, Salvador, 1994.

MACHADO, A. J.; ANDRADE, E. D.; ARAÚJO, H. A. B. Fauna de foraminíferos do litoral norte do Estado da Bahia. **Revista de Geologia**, Fortaleza, v. 19, n. 2, p. 147-154, 2006.

MATHERON, G. Principles of geostatistics. **Economic Geology**, Lancaster, v. 58, p. 1246-1266, 1963.

MORAES, S. S. **Distribuição especial e tafonomia de foraminíferos na plataforma continental da região norte da costa de dendê (foz do Rio Jequiricá à Ponta dos Castelhanos), Bahia.** 2006. 102 p. Tese (Doutorado em Geociências) - Universidade Federal da Bahia, Salvador, 2006.

MURRAY, J. W. **Ecology and applications of Benthic Foraminifera.** Cambridge: Cambridge University, 2006. 426 p.

PARIZOTTO, B. A. D. M. et al. Caracterização da zona estuarina do Rio Biguaçu (SC) através de parâmetros físico-químicos da água, associados a foraminíferos e tecamebas bentônicos. In: _____. **Oceanografia e mudanças globais.** São Paulo: Instituto Oceanográfico da USP, 2008. p. 561-572.

PERRY, J. N. Measures of spatial pattern for counts. **Ecology**, Durham, v. 79, n. 3, p. 1008-1017, 1998.

R DEVELOPMENT CORE TEAM. **R:** a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2012. Disponível em: <<http://www.R-project.org/>>. Acesso em: 10 dez. 2012.

RIBEIRO JUNIOR, P. J.; DIGGLE, P. J. GeoR: a package for geostatistical analysis. **R-News**, New York, v. 1, n. 2, p. 15-18, June 2001.

SEIGLIE, G. A. Relationship between the distribution of *Amphistegina* and the submerged Pleistocene reefs off western Puerto Rico. **Tulane Studies in Geology**, New Orleans, v. 6, p. 138-147, 1968.

SILVA, S. S. F. **Caracterização da fauna de foraminíferos presente no sedimento superficial do fundo da plataforma continental da costa atlântica de Salvador, Bahia.** 2004. 82 p. Dissertação (Mestrado em Geociências) - Universidade Federal da Bahia, Salvador, 2004.

TELFORD, R. J.; BIRKS, H. J. B. Evaluation of transfer functions in spatially structured environments. **Quaternary Science Reviews**, Oxford, v. 28, n. 13/14, p. 1309-1316, 2009.

_____. The secret assumptions of transfer functions. **Quaternary Science Reviews**, Oxford, v. 24, n. 20/21, p. 2173-2179, 2005.

VIEIRA, S. R. Geoestatística em estudos de variabilidade especial do solo. In: NOVAIS, R. F. de; ALVAREZ, V. H.; SCHAEFER, C. E. G. R. (Ed.). **Tópicos em ciências do solo**. Viçosa, MG: SBCS, 2000. p. 1-54.

ARTIGO 2 **Caracterização da plataforma e do talude continentais do Estado da Bahia, a partir da análise de agrupamento dos dados sedimentológicos e da microfauna de foraminíferos**

Artigo redigido conforme as normas da *Revista Pesquisas em Geociências* na qual foi submetido, aceito e deve ser publicado no próximo número da revista.

Caracterização da plataforma e do talude continentais do Estado da Bahia, a partir da análise de agrupamento dos dados sedimentológicos e da microfauna de foraminíferos

Characterization of the continental shelf and slope of the State of Bahia, from the cluster analysis of sedimentological data and the foraminifera microfauna

Ângela Cristina da Fonseca MIRANTE¹, João Domingos SCALON², Tânia Maria Fonseca ARAÚJO^{3*} & Helisângela Acris Borges de ARAÚJO⁴

1. Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA), Departamento de Ciências Aplicadas (DCA), Rua Emídio dos Santos, s/nº - Barbalho, CEP: 40300-010, Salvador, Bahia, Brasil. angelamirante@ifba.edu.br

2. Universidade Federal de Lavras (UFLA), Departamento de Exatas (DEX), Campus da UFLA, Caixa Postal 3037, CEP 37200-000. Lavras, Minas Gerais, Brasil. scalon@dex.ufla.br

3. Universidade Federal da Bahia (UFBA), Grupo de Estudos de Foraminíferos, Instituto de Geociências (IGEO), Rua Caetano Moura, 123 – Federação, CEP 40210-340, Salvador, Bahia, Brasil. tfaraujo@ufba.br

4. Universidade Federal da Bahia (UFBA), Grupo de Estudos de Foraminíferos, Instituto de Geociências (IGEO), Rua Caetano Moura, 123 – Federação, CEP 40210-340, Salvador, Bahia, Brasil. hacris@gmail.com.

Resumo

Esse trabalho teve como objetivo caracterizar a plataforma e o talude continentais do Estado da Bahia, a partir da análise multivariada aplicada aos dados de sedimento e da microfauna de foraminíferos. Trinta e oito amostras de sedimento da superfície do fundo foram analisadas, as quais foram processadas de acordo com a metodologia padrão para estudos sedimentológicos (granulometria e composição) e microfaunísticos. Na fração destinada à análise da microfauna de foraminíferos foram triadas e identificadas as primeiras trezentas testas. Para a realização da análise multivariada foram utilizadas as espécies que possuíam, ao mesmo tempo, frequência de ocorrência maior que 50 % (constantes) e frequência relativa maior que 5% (principais) em pelo menos duas amostras. A metodologia consistiu na aplicação da técnica de agrupamento em que foi definida a distância euclidiana ao quadrado como métrica de formação dos grupos, utilizando os métodos Ward e k-Médias. Foi realizada uma análise de agrupamento das espécies (modo R), a partir dos dados de frequência absoluta, para os diferentes intervalos batimétricos. Em seguida, foi aplicada uma análise de agrupamento entre as amostras (modo Q), a partir dos teores de sedimento presentes em cada amostra. A integração dos dados, a partir dos dendrogramas gerados pela análise multivariada, revelou que nas plataformas interna e média predomina areia siliciclástica; na plataforma externa areia carbonática cascalhosa e no talude areia carbonática lamosa. Com relação às espécies *Archaias angulatus* e *Peneroplis proteus*, elas apresentaram distribuição estreitamente relacionada com areia siliciclástica, enquanto a distribuição de *Discorbis mira* estava associada com areia carbonática e as espécies *Cassidulina laevigata* e *Cibicides pseudoungerianus* foram associadas com areia carbonática lamosa. As espécies *Amphistegina lessonii*, *Amphistegina gibbosa* e, principalmente, *Peneroplis carinatus* apresentaram distribuição

relacionada com sedimentação em área de transição entre areia siliciclástica e areia carbonática. A espécie *Globigerinoide ruber*, planctônica, foi encontrada principalmente em sedimento do tipo areia carbonática lamosa, localizado mais distante da costa.

Palavras-chaves: Análise multivariada, método Ward, método k-Médias, foraminíferos.

Abstract

This study aimed to characterize the continental shelf and slope of the State of Bahia, based on the multivariate analysis applied to sediments data and the foraminifera' microfauna. Thirty-eight samples from the bottom sediment surface were analyzed, which were processed in accordance with the standard methodology for sediment studies (grain size and composition) and microfaunal analysis. From the sediment fraction used for the foraminifer's microfauna analysis, the firsts 300 tests were randomly sorted and identified. For the multivariate analysis, it was used species that have, at the same time, frequency of occurrence higher than 50% (constant) and relative frequency greater than 5% (major) in, at least, two samples. The methodology consisted in the application of the technique of multivariate grouping in which it was set the squared Euclidean distance defined as a metric formation of groups, applying the methods Ward's and k-Means. A cluster analysis of species (R mode) was performed based on the absolute frequency data for each different bathymetric range. Then, a cluster analysis was applied between the samples (Q mode) based on the percent of sediment present in each sample. The dendograms data integration, generated by the multivariate analysis, revealed that siliciclastic

sands predominate in the inner and middle continental shelves, carbonate gravelly sands dominated in the outer shelf and carbonate muddy sands were the major sediment type found in the continental slope. The distribution of the foraminifera species *Archaias angulatus* and *Peneroplis proteus* were closely related to the occurrence of siliciclastic sands, while the *Discorbis mira* distribution were associated with carbonate sands, and the species *Cassidulina laevigata* and *Cibicides pseudoungerianus* were associated to carbonate muddy sands. The distribution of the species *Amphistegina lessonii*, *Amphistegina gibbosa* and especially *Peneroplis carinatus* were associated with the sediments from the transition between siliciclastic and carbonate sands. The planktonic *Globigerinoide ruber* were particularly found in carbonate muddy sands located farther from the coast.

Keywords: Multivariate analysis, Ward's method, k-Means's method, foraminifera.

1. Introdução

Nas margens continentais ocorre a transição fisiográfica entre os continentes (crosta continental) e as bacias oceânicas (crosta oceânica). Apesar de representarem apenas 20% da superfície coberta pelos oceanos, as margens continentais concentram 90% do potencial econômico do ecossistema marinho, representado por recursos bióticos, minerais e energéticos (Asmus, 1984). Nesta área, o Brasil prospecta mais de 80% de seu petróleo e mais de 50% de seu gás natural.

Em linhas gerais, a localização geográfica favorável à captação e distribuição de recursos terrestres e marinhos e a diversidade de ecossistemas associados, destacando-se principalmente praias, recifes, plataforma e talude continentais, ajudam a explicar a acentuada exploração de recursos da margem continental. Também alterações físicas diretas, resultantes de atividades como a construção de barragens ao longo de bacias hidrográficas, dragagem de portos, aterros, despejo de resíduos sólidos, derrubada de florestas litorâneas, além de atividades de turismo e recreação contribuem para o impacto ambiental nesta área (PNUMA, 2004).

Neste contexto, o uso de organismos como indicadores ecológicos têm sido amplamente empregados. Entre os organismos da microfauna marinha utilizada em estudos oceanográficos, os foraminíferos ocupam uma posição bastante significativa na biota oceânica e são reconhecidamente excelentes indicadores ecológicos e paleoecológicos. As testas desses organismos revelam algumas propriedades físico-químicas dos oceanos tornando-os potenciais indicadores de mudanças climáticas, ocorrência de agentes poluidores e na dinâmica dos oceanos (Bergue & Coimbra, 2008).

Devido ao tamanho microscópico dos foraminíferos, em pequenas amostras de sedimento é encontrada uma grande quantidade de indivíduos e de

espécies, o que permite trabalhar com dados estatísticos expressivos (Araújo & Araújo, 2010). Associado a isso, estudos revelaram que existe uma grande microfauna de foraminíferos presentes nos sedimentos de superfície e subsuperfícies da plataforma e do talude continentais (Passos *et al.*, 2001; Araújo & Machado, 2008b).

Nas últimas décadas, devido ao avanço tecnológico, intensificou-se a obtenção de dados geológicos e biológicos quantitativos. Entretanto, a maioria das informações coletadas não é analisada estatisticamente (Landim, 2011). Em estudos para descrição biológica das associações de foraminíferos, geralmente, é utilizada uma metodologia clássica de análises estatísticas descritivas e medidas tais como: número de indivíduos, frequência relativa, número de espécie ou riqueza e o índice de diversidade de Shannon-Wiener (Vieira, 2011). Observa-se também, o uso de técnicas estatísticas multivariadas, como análises de agrupamentos (Clarke & Warwick, 2001; Pawlowski & Holzmann, 2008; Araujo, 2009).

Em estudos oceanográficos, as assembléias de foraminíferos obtidas de diferentes localidades requerem análise e interpretação de dados provenientes de um grande número de variáveis. Nesse caso, a análise estatística multivariada se faz necessária, pois os procedimentos desenvolvidos por estas técnicas analisam os dados através da combinação das inúmeras variáveis, facilitando a compreensão da variabilidade total dos dados.

Com a aplicação de análise estatística multivariada será possível relacionar as variáveis envolvidas no processo e, mais precisamente, avaliar a relação tanto entre as espécies da microfauna de foraminíferos, como entre as espécies e os parâmetros abióticos, a exemplo da profundidade e do sedimento. Portanto, se o interesse na utilização destas técnicas é verificar como as amostras se relacionam, ou seja, o quanto estas são semelhantes segundo as variáveis utilizadas no trabalho, o método a ser utilizado será a análise por agrupamento.

A análise de agrupamento aplicada à exploração dos dados é uma ferramenta estatística que visa agrupar objetos semelhantes segundo suas características (variáveis) ou agrupar variáveis segundo os valores obtidos pelos objetos (Bussab *et al.*, 1990). A ideia é maximizar a homogeneidade de objetos dentro de grupos, ao mesmo tempo em que se maximiza a heterogeneidade entre os grupos (Hair *et al.*, 2005).

Esse trabalho tem como objetivo caracterizar a plataforma e talude continentais do Estado da Bahia, a partir da análise multivariada aplicada aos dados dos sedimentos e da microfauna de foraminíferos. Essa análise permite avaliar as condições sedimentológica e hidrodinâmica atuais, além de servirem de subsídio para futuros estudos de biomonitoramento da margem continental em estudo.

2. Área, material e métodos

2.1. Localização da área

Este trabalho apresenta como área de estudo o trecho da plataforma continental norte do Estado da Bahia, situado entre a praia da Barra, no município de Salvador, e a foz do rio Itariri, no município de Mata de São João, entre as coordenadas longitudinais 37°30'W e 38°30'W e latitudinais 12°S e 13°8'S. Essa área tem uma extensão costeira de aproximadamente 131 km, abrangendo as plataformas interna, média e externa e o talude continental superior, com área de 2.227 km². A figura 1 mostra a localização da área de estudo e das estações de coleta das amostras.

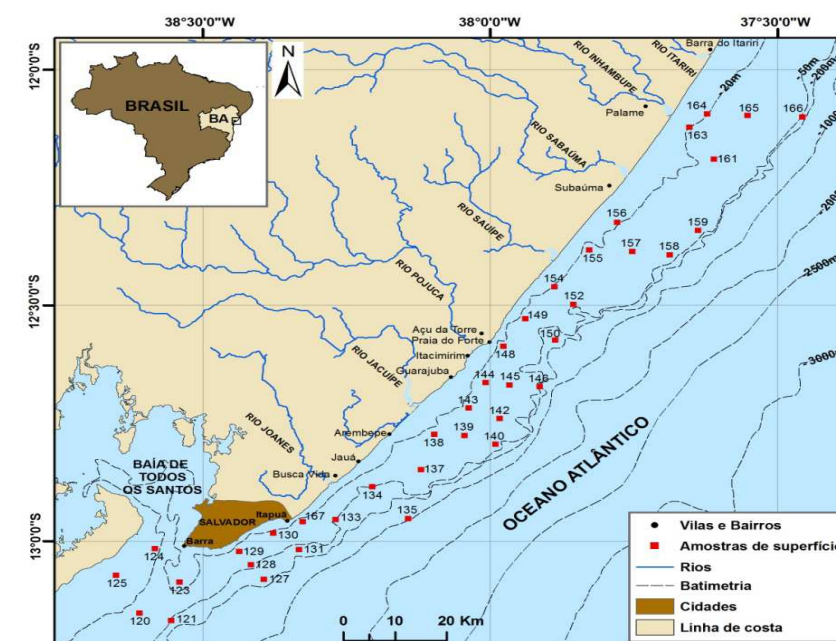


Figura 1 Localização da área de estudo e das estações de coleta das amostras
Fonte Araújo & Araújo (2010)

2.2. Localização e preparação das amostras

As amostras foram coletadas durante a operação GEOMAR XXV, em 1985, e estão arquivadas no Instituto de Geociências da Universidade Federal da Bahia. A operação GEOMAR XXV foi executada a bordo do navio oceanográfico Almirante Câmara, da Marinha Brasileira, Diretoria de Hidrografia e Navegação (DHN), cobrindo a área da plataforma continental e talude superior dos Estados de Alagoas, Sergipe e Bahia. Foi utilizada a draga tipo Gibbs para coletas das amostras de superfície.

Neste trabalho foram analisadas trinta e oito amostras do sedimento da superfície do fundo e as posições das estações de coleta estão apresentadas em

coordenadas geográficas na tabela 1 e sua distribuição espacial encontra-se representada na figura 1. As amostras foram agrupadas de acordo com a profundidade em que foram aleatoriamente coletadas, com intuito de examinar a variação na distribuição da microfauna de foraminíferos nas zonas batimétricas da plataforma e do talude continental. Dessa forma, para expressar a fisiografia (estreita) e a batimetria (rasa) da área em estudo, as amostras foram agrupadas em subambientes delimitados da seguinte maneira: a plataforma interna de 0 a 20 m, contendo cinco amostras; a plataforma média de > 20 a 40 m, contendo 16 amostras; a plataforma externa de > 40 a 60 m, contendo 11 amostras; e o talude superior de 80 a 300 m, com seis amostras (Tab.1).

Esses contornos batimétricos com intervalos de 20 m sobre a plataforma e intervalos de 400 m sobre o talude, seguiram subdivisão adotada por França (1979) no Projeto Reconhecimento Global da Margem Continental Brasileira (REMAC).

A área com profundidade de 60 a 80 m corresponde a quebra da plataforma continental, não sendo possível a coleta de amostras por ser de difícil acesso.

Tabela 1 Profundidade e coordenadas das amostras de fundo estudadas

Amostras	Prof. (m)	Local	Latitude (S)	Longitude (W)
123	14	Plataforma interna	13° 05' 31"	38° 32' 02"
124	15	Plataforma interna	13° 01' 36"	38° 34' 53"
125	17,7	Plataforma interna	13° 03' 26"	38° 23' 50"
156	19	Plataforma interna	12° 19' 56"	37° 48' 08"
164	20	Plataforma interna	12° 03' 24"	37° 36' 57"
155	21,5	Plataforma média	12° 22' 46"	37° 49' 50"
143	22,1	Plataforma média	12° 43' 26"	38° 03' 37"
154	23	Plataforma média	12° 27' 32"	37° 53' 15"
138	26	Plataforma média	12° 47' 39"	38° 07' 03"
144	26,4	Plataforma média	12° 40' 13"	38° 00' 44"
133	27	Plataforma média	12° 58' 25"	38° 18' 10"
161	28	Plataforma média	12° 11' 42"	37° 36' 53"
163	28	Plataforma média	12° 06' 13"	37° 38' 50"
148	29	Plataforma média	12° 35' 58"	37° 58' 57"
149	30	Plataforma média	12° 32' 01"	37° 56' 41"
165	32	Plataforma média	12° 04' 16"	37° 32' 25"
130	32,5	Plataforma média	12° 59' 37"	38° 24' 32"
120	36,3	Plataforma média	13° 08' 19"	38° 34' 58"
129	37	Plataforma média	13° 02' 45"	38° 28' 04"
157	39	Plataforma média	12° 23' 10"	37° 44' 38"
134	40	Plataforma média	12° 52' 56"	38° 14' 13"
142	43	Plataforma externa	12° 45' 27"	37° 00' 28"
139	44	Plataforma externa	12° 48' 07"	38° 04' 40"
145	44	Plataforma externa	12° 41' 53"	38° 58' 37"
158	44	Plataforma externa	12° 23' 28"	37° 40' 44"
159	47	Plataforma externa	12° 20' 17"	37° 37' 42"
140	48	Plataforma externa	12° 49' 56"	38° 01' 24"
167	49	Plataforma externa	12° 59' 15"	37° 21' 05"
137	50	Plataforma externa	12° 51' 21"	38° 09' 24"
152	51	Plataforma externa	12° 29' 50"	37° 50' 46"
150	57	Plataforma externa	12° 34' 57"	37° 52' 35"
128	60	Plataforma externa	13° 04' 56"	38° 26' 58"
166	80	Talude	12° 05' 15"	37° 27' 14"
146	120	Talude	12° 43' 30"	37° 55' 35"
131	150	Talude	13° 02' 20"	38° 20' 46"
127	190	Talude	13° 07' 07"	38° 26' 07"
121	218,5	Talude	13° 11' 04"	38° 33' 15"
135	300	Talude	12° 58' 22"	38° 11' 35"

2.3. Análises sedimentológicas

As amostras foram secas, pesadas, homogeneizadas com água destilada e posteriormente passadas através de uma peneira com espaçamento de malha de 0,062 mm para eliminação dos sais e separação da fração lama. Em seguida, a fração retida e secada a 60°C, foi peneirada segundo a escala granulométrica de Wentworth (1962), com intervalo de um phi. A fração lama, quando presente, foi colocada para secar a 40°C e pesada. O tipo de sedimento de cada amostra foi determinado segundo a nomenclatura proposta por Shepard (1954), que apresenta um esquema classificativo puramente descritivo, não existindo preocupações em viabilizar deduções de características hidrodinâmicas do ambiente de deposição. Nesse esquema existem três classes correspondentes à componente dominante, ou seja, componente com percentagem superior a 75%. A designação da classe corresponde à componente dominante: areia, silte e argila. Nas outras seis classes existem misturas de dois componentes e recebem a denominação da componente dominante adjetivada pelo da componente secundária: areia siltosa, argila arenosa, argila siltosa, silte argiloso, silte arenoso e areia argilosa. Schlee (1973) modificou o esquema de Shepard (1954) complementando-o com um segundo diagrama ternário em que a componente cascalho é valorizada, passando a poder ser aplicado a todos os tipos de sedimentos.

Para a análise da composição do sedimento, 300 grãos de cada amostra foram identificados com o auxílio de microscópio estereoscópio (lupa binocular) e multiplicados pelo peso da amostra para obtenção dos percentuais de cada componente. A classificação do sedimento segundo a composição carbonática e/ou siliciclástica foi de acordo com a classificação apresentada por Testa (1996) onde: sedimento siliciclástico possui teores de grãos carbonáticos inferiores a 40

%; sedimento misto apresenta teores de grãos carbonáticos entre 40 e 60 %; e sedimento carbonático apresenta teores de grãos carbonáticos superiores a 60 %.

2.4. Análise da fauna de foraminíferos

Foram retirados de cada amostra de superfície do fundo, 10 cm³ de material que foi lavado, em água destilada através de uma peneira com espaçamento de malha 0,062 mm para eliminação dos sais e separação da fração lama. Em seguida o material foi colocado para secar em estufa a 50°C, e, posteriormente, foi pesado e quarteado. Em seguida foram separadas, aleatoriamente, as 300 primeiras testas de foraminíferos, para identificação das espécies, através de lupa binocular.

Esta metodologia descrita por Drooger & Kaasschietter (1958) é considerada adequada quando se utiliza grande número de amostras e se deseja realizar tratamento estatístico (Araújo & Araújo, 2010). Os espécimes foram identificados segundo Loeblich & Tappan (1988) e publicações de foraminíferos recentes: Eichler *et al.* 2012; Araújo & Machado, 2008a,b,c.

Exemplares de foraminíferos, que apresentaram melhor grau de conservação foram fotografados em microscópio eletrônico de varredura (MEV). As fotomicrografias desses exemplares estão expostas nas figuras 7 e 8.

2.5. Tratamento estatístico

Considerando os 300 espécimes identificados em cada amostra, foi calculada a quantidade de indivíduos por espécie. A partir desses dados de abundância absoluta das espécies, foram confeccionadas matrizes básicas de dados e aplicando a metodologia proposta por Dajoz (1983), Clarke & Warwick (1994), Ab'Saber *et al.* (1997).

2.5.1. Frequência de ocorrência (constância das espécies)

Consiste no número de amostras que contêm a espécie (p) em relação ao número total de amostras analisadas (P) (Ab'Saber *et al.*, 1997). Esta frequência foi calculada pela fórmula:

$$FO = \frac{p \times 100}{P} \quad (1)$$

Após estes cálculos, as espécies foram agrupadas em três categorias de acordo com a classificação de Dajoz (1983), sendo classificadas como constantes (as espécies presentes em mais de 50 % das amostras); acessórias (que ocorrem entre 25 % a 49 % das amostras) e acidentais (presentes em menos de 25 % das amostras).

2.5.2. Frequência relativa (abundância das espécies)

Consiste na razão entre o número total de indivíduos de uma categoria obtidos na amostra (n) e o total de indivíduos de todas as categorias (T) na amostra, expresso em percentagem (Ab'Saber *et al.*, 1997). É dada por:

$$FR = \frac{n \times 100}{T} \quad (2)$$

Este critério determina as espécies abundantes em cada amostra. Os resultados permitiram classificar as espécies, de acordo com critérios de classificação de Dajoz (1983), em principais (frequência acima de 5 %); acessória (entre 5 e 1 %) e traço (inferior a 1 %).

2.5.3. *Análise de agrupamento*

Devido à grande quantidade de espécies encontradas na área de estudo, para aplicar a análise multivariada foi realizado um levantamento das espécies que possuem, ao mesmo tempo, frequência de ocorrência maior que 50 % (constantes) e abundância relativa maior que 5% (principais) em pelo menos duas amostras. Estes critérios foram anteriormente utilizados por Araújo (2004), Araújo & Araújo (2010) e Bruno (2008). Dessa forma, as espécies raras e as espécies que não têm representação gráfica foram eliminadas.

Foi aplicada a análise de agrupamento visando adequar uma técnica multivariada para estudos envolvendo grande número de amostras e espécies, com o objetivo de medir o grau de similaridade entre os objetos estudados e reuni-los em um mesmo grupo. Os dados originais foram padronizados para assegurar que as variáveis sejam igualmente importantes na determinação do grau de similaridade entre os objetos, permitindo que os resultados obtidos apresentassem maior confiabilidade.

A técnica de agrupamento hierárquico interliga as amostras ou espécies por suas associações, produzindo um dendrograma (gráfico). A suposição básica de sua interpretação é que quanto menor a distância entre os pontos, maior a semelhança entre as amostras.

Os dados foram submetidos a diferentes métodos de agrupamento hierárquicos para decidir o mais adequado para o desenvolvimento do trabalho. Para os diversos métodos de agrupamento utilizados foram obtidas, após a construção dos dendrogramas, as respectivas matrizes cofenéticas resultantes da simplificação proporcionada pelo método.

Para avaliar a consistência dos agrupamentos obtidos, ou seja, verificar a capacidade do dendrograma em reproduzir a matriz de dissimilaridade, calculou-se o coeficiente de correlação cofenética. Esse índice compara as distâncias

efetivamente observadas entre os objetos e as distâncias previstas a partir do processo de agrupamento, medindo o grau de ajuste entre a matriz de dissimilaridade original e a matriz de dissimilaridade produzida após a construção do dendrograma (matriz cofenética).

O coeficiente de correlação cofenética permite avaliar o grau de deformação provocado pela construção do dendrograma e o menor grau de distorção será definido pelo maior valor desse coeficiente. Segundo Bussab *et al.* (1990), quanto mais próximo da unidade estiver esse índice de correlação cofenética melhor será a representação do dendrograma. Segundo este autor, a correlação cofenética é dada por:

$$r_{cof} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n (c_{ij} - \bar{c})(d_{ij} - \bar{d})}{\sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^n (c_{ij} - \bar{c})^2} \sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^n (d_{ij} - \bar{d})^2}} \quad (3)$$

em que:

c_{ij} = valor de dissimilaridade entre os objetos i e j , obtidos pela matriz cofenética.

d_{ij} = valor de dissimilaridade entre os objetos i e j , obtidos pela matriz de dissimilaridade original.

$$\bar{c} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_{ij}, \quad (4)$$

$$\bar{d} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n d_{ij}, \quad (5)$$

Portanto, um método é melhor do que o outro, quando o dendrograma fornecer uma imagem menos distorcida da realidade.

O método de Ward foi aplicado nesse trabalho porque obteve o maior coeficiente de correlação cofenética, apresentando o melhor resultado em relação aos outros métodos utilizados. Muitos estudos mostraram que, de maneira geral, o método de Ward possui melhor desempenho (Rencher, 2002).

Esse método hierárquico, às vezes chamado de método de variância mínima, está fundamentado na análise de variância e busca juntar dois grupos cuja fusão minimiza a soma de quadrados dentro dos grupos, ou seja, a soma de quadrados dos erros (isto é, a variância mínima dentro do grupo) (Lattin *et al.*, 2011). A cada passo do processo de análise, as amostras ou espécies que apresentaram padrões similares são agrupadas de acordo com uma medida de dissimilaridade, definida nesse trabalho pela distância euclidiana ao quadrado.

Ferreira (2008) recomenda utilizar a distância euclidiana quadrática ao aplicar o método de Ward, que é norteado pelas seguintes ideias: Inicialmente, têm-se tantos grupos quantos objetos, ou seja, temos n grupos com 1 objeto cada. Agrupam-se, no primeiro ciclo do algoritmo, os dois grupos, de um único objeto cada, que minimizam o incremento na soma de quadrados dos erros em relação ao estágio anterior. É conveniente ter a noção clara que ao aglomerar dois grupos, a soma de quadrados dos erros irá aumentar. O que irá se buscar é o par de grupos que provocará o menor incremento possível na soma de quadrados dos erros entre todos os possíveis pares. Ao aglomerar os grupos que atendem à função objetivo, minimizar o incremento na soma de quadrados dos erros, deve-se atualizar a matriz de dissimilaridades e então repetir o processo.

Segundo Rencher (2002), se AB é um grupo obtido pela combinação dos grupos A e B , então a soma das distâncias dentro do grupo (dos elementos a partir do vetor de média dos grupos) são SSE_A (a soma de quadrados relativa ao grupo A), SSE_B (a soma de quadrados relativa ao grupo B) e SSE_{AB} (a soma de quadrados relativa ao grupo AB), em que:

$$SSE_A = \sum_{i=1}^{n_A} (y_{iA} - \bar{y}_A)' (y_{iA} - \bar{y}_A) \quad (6)$$

sendo $y_{iA} = (y_{i1A}, y_{i2A}, \dots, y_{ikA})'$ o vetor de observações do elemento amostral i pertencente ao grupo A , $\bar{y}_A = (\bar{y}_{1A}, \bar{y}_{2A}, \dots, \bar{y}_{kA})'$ o vetor de médias amostrais do grupo A , $i = 1, 2, \dots, k$, sendo k o número de variáveis e n_A o número de elementos amostrais do grupo A .

$$SSE_B = \sum_{i=1}^{n_B} (y_{iB} - \bar{y}_B)' (y_{iB} - \bar{y}_B) \quad (7)$$

sendo $y_{iB} = (y_{i1B}, y_{i2B}, \dots, y_{ikB})'$ o vetor de observações do elemento amostral i pertencente ao grupo B , $\bar{y}_B = (\bar{y}_{1B}, \bar{y}_{2B}, \dots, \bar{y}_{kB})'$ o vetor e médias amostrais do grupo B , $i = 1, 2, \dots, k$, sendo k o número de variáveis e n_B o número de elementos amostrais do grupo B .

$$SSE_{AB} = \sum_{i=1}^{n_{AB}} (y_{iAB} - \bar{y}_{AB})' (y_{iAB} - \bar{y}_{AB}) \quad (8)$$

sendo $y_{iAB} = (y_{i1AB}, y_{i2AB}, \dots, y_{ikAB})'$ o vetor de observações do elemento amostral i pertencente ao grupo AB , $\bar{y}_{AB} = (\bar{y}_{1AB}, \bar{y}_{2AB}, \dots, \bar{y}_{kAB})'$ o vetor e médias amostrais do grupo AB , $i = 1, 2, \dots, k$, sendo k o número de variáveis e n_{AB} o número de elementos amostrais do grupo AB . Sendo,

$$\bar{y}_{AB} = \frac{(n_A \bar{y}_A + n_B \bar{y}_B)}{n_A n_B} \quad (9)$$

em que, n_A, n_B e $n_{AB} = n_A + n_B$ são o número de pontos em A, B e AB , respectivamente. Então, as somas das distâncias são equivalentes à soma de quadrados dentro dos grupos, denotadas por SSE_A, SSE_B e SSE_{AB} .

O método de Ward une dois grupos A e B que minimizam o acréscimo em SSE , definido como:

$$I_{AB} = SSE_{AB} - (SSE_A + SSE_B) \quad (10)$$

Conforme Rencher (2002) pode-se demonstrar que I_{AB} em (10) possui duas formas equivalentes:

$$I_{AB} = n_A (\bar{y}_A - \bar{y}_{AB})' (\bar{y}_A - \bar{y}_{AB}) + n_B (\bar{y}_B - \bar{y}_{AB})' (\bar{y}_B - \bar{y}_{AB}) \quad (11)$$

$$I_{AB} = \frac{n_A n_B}{n_A + n_B} (\bar{y}_A - \bar{y}_B)' (\bar{y}_A - \bar{y}_B) \quad (12)$$

Assim, por (12) minimizar o aumento na SSE é equivalente a minimizar a distância entre grupos. Se A e B possuem apenas um elemento y_i e y_j , respectivamente, então $SSE_A = 0$ e $SSE_B = 0$, e (11) e (12) reduz a:

$$I_{ij} = I_{AB} = SSE_{AB} = \frac{1}{2}(\bar{y}_A - \bar{y}_B)'(\bar{y}_A - \bar{y}_B) = \frac{1}{2}d^2(y_i, y_j) \quad (13)$$

O método de Ward é apropriado apenas para variáveis quantitativas e sua aplicação não depende de que os dados sejam provenientes ou não de uma distribuição normal multivariada (Mingoti, 2005).

Para validar os agrupamentos obtidos pelo método de Ward foi aplicado o método não hierárquico k-Médias, que utilizou como sementes iniciais os vetores de médias de cada grupo (centroides) formado pelo método de Ward. Segundo Bussab *et al.* (1990) o método de partição k-Médias fornece indicações mais precisas sobre o número de grupos a ser formado e talvez seja um dos mais utilizados quando se têm muitos objetos para agrupar, com pequenas variações. O critério de homogeneidade dentro do grupo e heterogeneidade entre grupos é a soma de quadrados residual baseado na análise de variância. Assim, quanto menor for este valor, mais homogêneos são os elementos dentro de cada grupo e melhor será a partição.

Para avaliar o grau de homogeneidade dentro de cada grupo, calculou-se o coeficiente R^2 da partição definido por:

$$R^2 = \frac{SSB}{SST_c} \quad (14)$$

em que:

SSB = a soma de quadrados total entre os grupos de partição.

SST_c = a soma de quadrados total corrigida para a média global de cada variável.

Quanto maior for o valor de R^2 , maior será a soma de quadrados entre grupos e menor será o valor da soma de quadrados residual. O coeficiente R^2 cujo valor máximo alcançado é igual a 1 pode ser utilizado como critério para determinar o número de grupos da partição final (Mingoti, 2005).

Foi realizada uma análise de agrupamento entre espécies (modo R) com padrões similares, segundo as amostras (localizações) que as contém. A variável corresponde ao número de indivíduos (frequência absoluta) de cada espécie nas diferentes localizações. Esse tipo de análise foi realizada para as plataformas interna, média, externa e o talude.

Na plataforma interna foi realizada uma análise da associação entre as frequências absolutas (FA) das 11 espécies de foraminíferos para cada uma das 5 localizações, obtendo uma matriz de informação do tipo 11 x 5. Na plataforma média foi aplicada uma análise de agrupamento entre as frequências absolutas (FA) das 18 espécies de foraminíferos para cada uma das 16 localizações (amostras), obtendo uma matriz de informação do tipo 18 x 16. Na plataforma externa realizou-se uma análise da associação entre as frequências absolutas (FA) das 13 espécies de foraminíferos para cada uma das 11 localizações (amostras), obtendo uma matriz de informação do tipo 13 x 11. No talude foi realizada uma análise da associação entre as frequências absolutas (FA) das 10 espécies de foraminíferos para cada uma das 6 localizações (amostras), obtendo uma matriz de informação do tipo 10 x 6.

Em seguida, foi aplicada uma análise de agrupamento entre as amostras (modo Q) segundo os teores sedimentológicos presentes em cada amostra. Nessa análise, foram consideradas 38 amostras e 5 variáveis: teor de carbonato (%), teor de siliciclástico (%), areia (%), lama (%) e cascalho (%), obtendo uma matriz de informação do tipo 38 x 5. Toda a análise estatística foi realizada utilizando-

se programas e bibliotecas que estão disponíveis no *software* R (R Development Core Team, 2012).

3. Resultados

O dendrograma apresentado na figura 2 representa a análise de agrupamento entre as 11 espécies (modo R), da plataforma interna, considerando as cinco amostras (localizações) que as contém. O coeficiente de correlação cofenética foi igual a 0,97. No dendrograma, o corte foi estabelecido na distância **12,7** e foram definidos quatro grupos: A, B, C, D.

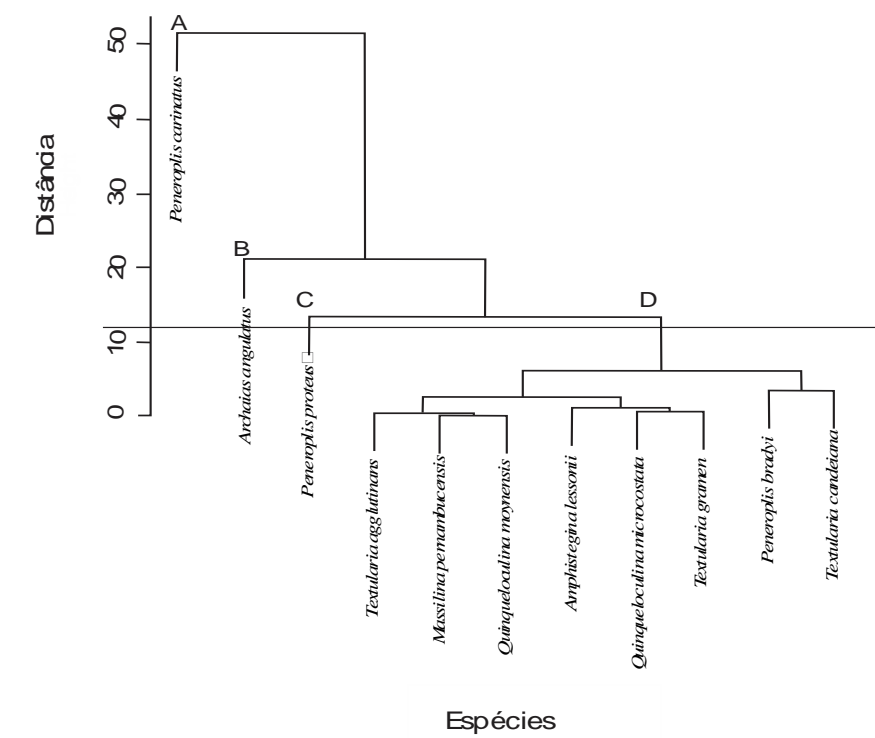


Figura 2 Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies da plataforma interna, a partir das amostras (localização)

No Grupo A foi encontrada a espécie *Peneroplis carinatus* individualizada que apresentou, em média, 49 indivíduos neste grupo e elevada abundância (frequência absoluta) nas diferentes localidades (amostras). A maior abundância desta espécie foi encontrada na amostra 156 (107 espécimes) e a menor abundância na amostra 125 (5 espécimes). O Grupo B isolou a espécie *Archaias angulatus* que obteve, em média, 20,4 indivíduos neste grupo. A maior abundância desta espécie foi observada na amostra 123 (75 indivíduos) e a menor abundância na amostra 124 (3 indivíduos). O Grupo C individualizou a espécie *Peneroplis proteus* que apresentou, em média, 21 indivíduos neste grupo. A abundância elevada foi evidenciada na amostra 125 (67 indivíduos) e abundância reduzida na amostra 123 (2 indivíduos). O Grupo D foi formado pelas demais espécies. A média de indivíduos neste grupo variou de 5,2 a 12,2.

O Grupo A foi caracterizado por apresentar maior abundância da espécie *Peneroplis carinatus* na plataforma interna (245 indivíduos). Os Grupos B e C apresentaram, respectivamente, as espécies *Archaias angulatus* e *Penerolopis proteus* como as mais abundantes, ambas apresentando mais de 100 espécimes nas amostras de plataforma interna. No Grupo D foram reunidas as espécies com o máximo de 61 e mínimo de 26 indivíduos (Tab. 2 e 3).

O método k-Médias foi aplicado utilizando-se como sementes iniciais os quatro grupos formados pelo método de Ward. Foi observado que as 11 espécies da plataforma interna foram classificadas nos mesmos agrupamentos pelos dois métodos. O coeficiente R^2 obtido foi 85,8%, indicando que os grupos encontrados são homogêneos internamente.

Tabela 2 Número de indivíduos das espécies da plataforma interna

Espécies / Amostras	123	124	125	156	164	Total
<i>Amphistegina lessonii</i>	10	5	15	2	5	37
<i>Archaias angulatus</i>	75	3	10	4	10	102
<i>Massilina pernambucensis</i>	1	8	4	5	18	36
<i>Peneroplis bradyi</i>	2	22	17	7	12	60
<i>Peneroplis carinatus</i>	5	36	55	107	42	245
<i>Peneroplis proteus</i>	2	6	67	19	11	105
<i>Quinqueloculina microcostata</i>	3	5	15	10	16	49
<i>Quinqueloculina moynensis</i>	0	6	3	0	17	26
<i>Textularia agglutinans</i>	1	7	5	17	19	49
<i>Textularia candeiana</i>	0	17	3	12	29	61
<i>Textularia gramen</i>	0	9	16	12	10	47

Tabela 3 Medidas descritivas dos agrupamentos de espécies encontrados na plataforma interna pelo método de Ward

Variáveis (FA)	Grupo A		Grupo B		Grupo C		Grupo D	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
<i>Amphistegina lessonii</i>							7,4	5,13
<i>Archaias angulatus</i>			20,4	30,70				
<i>Massilina pernambucensis</i>							7,2	6,53
<i>Peneroplis bradyi</i>							12,0	7,91
<i>Peneroplis carinatus</i>	49,0	37,26						
<i>Peneroplis proteus</i>					21	26,49		
<i>Quinqueloculina microcostata</i>							9,8	5,81
<i>Quinqueloculina moynensis</i>							5,2	7,05
<i>Textularia agglutinans</i>							9,8	7,82
<i>Textularia candeiana</i>							12,2	11,61
<i>Textularia gramen</i>							9,4	5,90

FA= frequência absoluta

A figura 3 apresenta o dendrograma obtido da análise de agrupamento entre as 18 espécies da plataforma média, considerando 16 amostras (localizações). O valor do índice de correlação cofenética obtido para essa associação de espécies correspondeu a 0,93. O corte no dendrograma foi estabelecido na distância 50 onde foram definidos cinco grupos A, B, C, D, E.

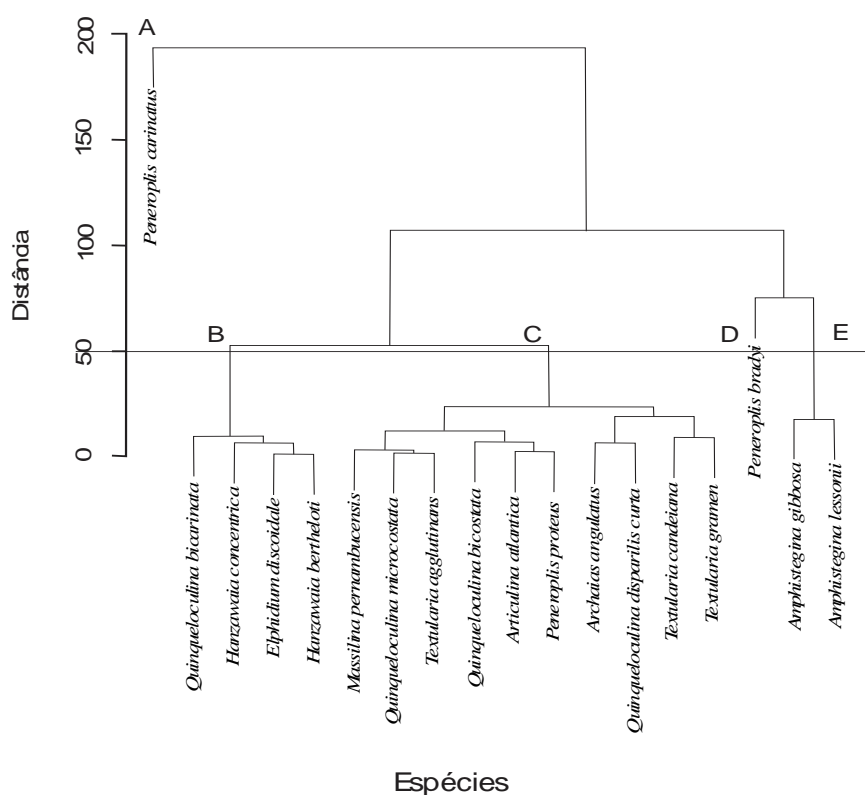


Figura 3 Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies da plataforma média, a partir das amostras (localização)

No Grupo A isolou a espécie *Peneroplis carinatus* que apresentou, em média, 41,1 indivíduos neste grupo e maior abundância na amostra 144 (142 espécimes). Esta espécie foi a mais frequente da plataforma média (658 indivíduos). O Grupo B reuniu as espécies com abundância absoluta entre 50 e 103 espécimes e a média de indivíduos neste grupo variou entre 3,1 e 6,4 (*Hanzawaia bertheloti* e *Quinqueloculina bicarinata*, respectivamente). O Grupo C agregou as demais espécies, que possuem abundância entre 41 e 149 indivíduos e a média de indivíduos neste grupo variou de 2,6 a 9,2. O Grupo D isolou a espécie *Peneroplis bradyi*, que possui abundância acima de 250 indivíduos que apresentou, em média, 41,1 indivíduos neste grupo. No Grupo E foram agrupadas duas espécies do mesmo gênero *Amphistegina gibbosa* (238 espécimes) e *A. lessonii* (273 espécimes). As médias de indivíduos encontradas neste grupo foram 14,9 e 17,1, respectivamente (Tab. 4 e 5).

Utilizando como sementes iniciais os cinco grupos formados pelo método de Ward foi aplicado o método k-Médias. Foi possível observar que as 18 espécies da plataforma média foram classificadas nos mesmos agrupamentos pelos dois métodos. O coeficiente R^2 obtido foi 80%, indicando que os grupos encontrados são homogêneos internamente.

Tabela 4 Número de indivíduos das espécies da plataforma média

Espécies / Amostras	120	129	130	133	134	138	143	144	148	149	154	155	157	161	163	165	Total
<i>Amphistegina gibbosa</i>	2	11	0	66	0	45	20	0	1	22	29	0	29	8	2	3	238
<i>Amphistegina lessonii</i>	7	13	3	40	3	9	14	6	5	18	50	15	38	25	6	21	273
<i>Archaias angulatus</i>	6	5	0	10	0	37	16	6	1	17	7	1	5	9	7	20	147
<i>Articulina atlântica</i>	2	3	2	1	0	3	3	24	12	4	0	5	0	2	8	10	79
<i>Elphidium discoidale</i>	6	7	10	1	18	0	0	1	2	1	2	3	3	1	2	5	62
<i>Hanzawaia bertheloti</i>	4	1	9	0	19	0	0	0	2	3	7	0	3	1	0	1	50
<i>Hanzawaia concêntrica</i>	20	7	6	2	17	0	0	0	3	2	2	2	0	1	0	0	62
<i>Massilina pernambucensis</i>	0	0	0	8	0	0	1	1	1	0	2	16	0	8	16	1	54
<i>Peneroplis bradyi</i>	30	47	18	9	5	6	12	29	19	21	5	2	12	10	14	17	256
<i>Peneroplis carinatus</i>	5	12	2	40	0	102	96	142	55	22	11	60	13	22	26	50	658
<i>Peneroplis proteus</i>	2	4	0	4	0	11	9	33	2	6	2	9	1	3	3	2	91
<i>Quinqueloculina bicarinata</i>	8	5	8	9	10	6	5	2	5	3	9	2	3	16	7	5	103
<i>Quinqueloculina bicostata</i>	2	4	2	3	2	2	0	0	1	20	2	1	0	1	0	1	41
<i>Quinqueloculina disparilis curta</i>	10	16	2	6	1	3	7	2	2	13	17	2	10	15	1	11	118
<i>Quinqueloculina microcostata</i>	1	1	0	1	1	5	15	7	3	2	4	6	1	2	16	7	72
<i>Textularia agglutinans</i>	1	2	0	0	1	2	3	1	6	2		5	0	0	25	1	49
<i>Textularia candeiana</i>	7	6	9	2	3	5	8	3	13	8	5	14	2	6	33	9	133
<i>Textularia gramen</i>	12	17	4	13	3	7	7	3	3	9	15	21	11	4	15	5	149

Tabela 5 Medidas descritivas dos agrupamentos de espécies encontrados na plataforma média pelo método de Ward

Variáveis (FA)	Grupo A		Grupo B		Grupo C		Grupo D		Grupo E	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
<i>Amphistegina gibbosa</i>									14,9	19,34
<i>Amphistegina lessonii</i>									17,1	14,40
<i>Archaias angulatus</i>					9,2	9,54				
<i>Articulina atlântica</i>					4,9	6,19				
<i>Elphidium discoidale</i>			3,6	4,67						
<i>Hanzawaia bertheloti</i>			3,1	5,00						
<i>Hanzawaia concêntrica</i>			3,9	6,11						
<i>Massilina pernambucensis</i>					3,4	5,57				
<i>Peneroplis bradyi</i>							16,0	11,62		
<i>Peneroplis carinatus</i>	41,1	41,28								
<i>Peneroplis proteus</i>					4,9	6,19				
<i>Quinqueloculina bicarinata</i>			6,4	3,60						
<i>Quinqueloculina bicostata</i>					2,6	4,79				
<i>Quinqueloculina disparilis curta</i>					7,4	5,78				
<i>Quinqueloculina microcostata</i>					4,5	4,86				
<i>Textularia agglutinans</i>					3,1	6,12				
<i>Textularia candeiana</i>					8,3	7,46				
<i>Textularia gramen</i>					9,3	5,71				

FA = Frequência absoluta

Na figura 4 observa-se o dendrograma correspondente à análise de agrupamento entre as 13 espécies da plataforma externa, considerando 11 amostras (localizações) que as encerram. O valor do índice de correlação cofenética obtido para essa associação de espécies correspondeu a 0,87. Estabelecendo o corte na distância de 18,3 foram definidos seis grupos: A, B, C, D, E e F.

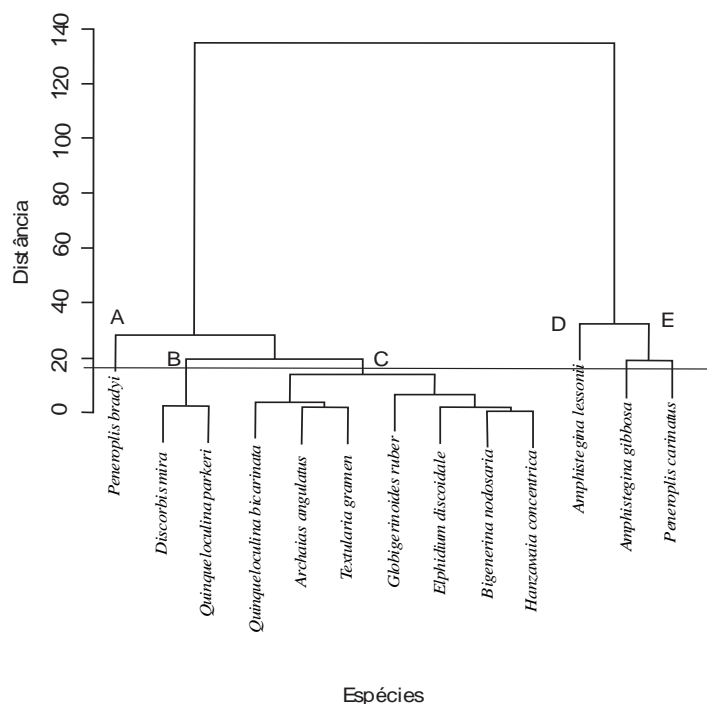


Figura 4 Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies da plataforma externa, a partir das amostras (localização)

No Grupo A foi individualizada a espécie *Penaeopsis bradyi*, com 147 indivíduos. Neste grupo, a média de indivíduos é 13,4. Na amostra 128 foi observada a maior abundância absoluta desta espécie (43 indivíduos) e na

amostra 158 a menor (1 indivíduo). O Grupo B reuniu as espécies *Quinqueloculina parkere* e *Discorbis mira*, com abundância entre 44 e 92 espécimes e as médias de indivíduos encontradas no grupo foram 4,9 e 9,2, respectivamente. O Grupo C foi constituído de 7 espécies: *Quinqueloculina bicarinata* (104 indivíduos), *Textularia gramem* (98 indivíduos), *Globigerinoides ruber* (96 indivíduos), *Archaias angulatus* (71 indivíduos), *Elphidium discoidale* (49 indivíduos), *Bigenerina nodosaria* (49 indivíduos) e *Hanzawaia concentrica* (32 indivíduos). A média de indivíduos neste grupo variou de 5,3 a 9,5. O Grupo D isolou a espécie *Amphistegina lessonii*, a qual apresentou maior abundância na plataforma externa, ou seja, 375 indivíduos. Neste grupo a maior abundância desta espécie (87 indivíduos) foi encontrada na amostra 150 e a menor abundância (1 indivíduo) na amostra 152. A média de indivíduos neste grupo foi 34,1. No Grupo E foi isolada a espécie *Amphistegina gibbosa* com 280 indivíduos. Neste grupo, a média de indivíduos é 25,5. Na amostra 142 foi observada a maior abundância desta espécie (51 indivíduos) e na amostra 158 a menor abundância (7 indivíduos). O Grupo F isolou a espécie *Peneroplis carinatus*, com 284 indivíduos. Neste grupo a média de indivíduos é 25,8. Observa-se a maior abundância desta espécie na amostra 142 (48 indivíduos) e a menor abundância na amostra 150 (4 indivíduos).

O método k-Médias foi aplicado utilizando-se como sementes iniciais os seis grupos formados pelo método de Ward. Foi evidenciado que as 13 espécies da plataforma externa foram classificadas nos mesmos agrupamentos pelos dois métodos. O coeficiente R^2 obtido foi 86,5%, indicando que os grupos encontrados são homogêneos internamente (Tab. 6 e 7).

Tabela 6 Número de indivíduos das espécies da plataforma externa

Espécies / Amostras	128	137	139	140	142	145	150	152	158	159	167	Total
<i>Amphistegina gibbosa</i>	15	8	39	50	51	27	39	0	7	22	22	280
<i>Amphistegina lessonii</i>	23	36	9	45	43	23	87	1	31	49	28	375
<i>Archaias angulatus</i>	9	6	16	5	11	5	3	4	3	2	7	71
<i>Bigenerina nodosaria</i>	10	5	3	2	6	5	1	16	0	0	1	49
<i>Discorbis mira</i>	2	2	7	9	0	4	15	0	27	19	7	92
<i>Elphidium discoidale</i>	5	7	2	0	0	0	2	17	3	3	10	49
<i>Globigerinoides ruber</i>	5	6	2	8	5	8	10	16	20	10	6	96
<i>Hanzawaia concêntrica</i>	9	1	0	0	0	1	2	16	0	0	3	32
<i>Peneroplis bradyi</i>	43	15	6	11	13	19	5	10	1	6	18	147
<i>Peneroplis carinatus</i>	8	35	46	31	48	25	4	22	20	22	23	284
<i>Quinqueloculina bicarinata</i>	5	18	15	10	11	14	3	6	7	10	5	104
<i>Quinqueloculina parkeri</i>	0	1	4	11	0	0	0	0	16	11	1	44
<i>Textularia gramen</i>	14	15	13	6	9	4	6	5	6	4	16	98

Tabela 7 Medidas descritivas dos agrupamentos de espécies encontrados na plataforma externa pelo método de Ward

Variáveis (FA)	Grupo A		Grupo B		Grupo C		Grupo D		Grupo E		Grupo F	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
<i>Amphistegina gibbosa</i>									25,5	17,47		
<i>Amphistegina lessonii</i>							34,1	22,89				
<i>Archaias angulatus</i>					6,5	4,16						
<i>Bigenarina nodosaria</i>					5,4	4,88						
<i>Discorbis mira</i>			9,2	8,64								
<i>Elphidium discoidale</i>					6,1	5,19						
<i>Globigerinoides ruber</i>					8,7	5,22						
<i>Hanzawaia concentrica</i>					5,3	6,02						
<i>Peneroplis bradyi</i>	13,4	11,32										
<i>Peneroplis carinatus</i>											25,8	13,70
<i>Quinqueloculina bicarinata</i>					9,5	4,76						
<i>Quinqueloculina parkeri</i>			4,9	6,13								
<i>Textularia gramen</i>					8,9	4,68						

FA= frequência absoluta

O dendrograma obtido na figura 5 representa a análise de agrupamento entre as 10 espécies do talude, considerando 6 amostras (localizações) que as contém. O valor do índice de correlação cofenética obtido para essa associação de espécies correspondeu a 0,89. O corte foi estabelecido na distância 12,5, sendo definidos quatro grupos: A, B, C, D.

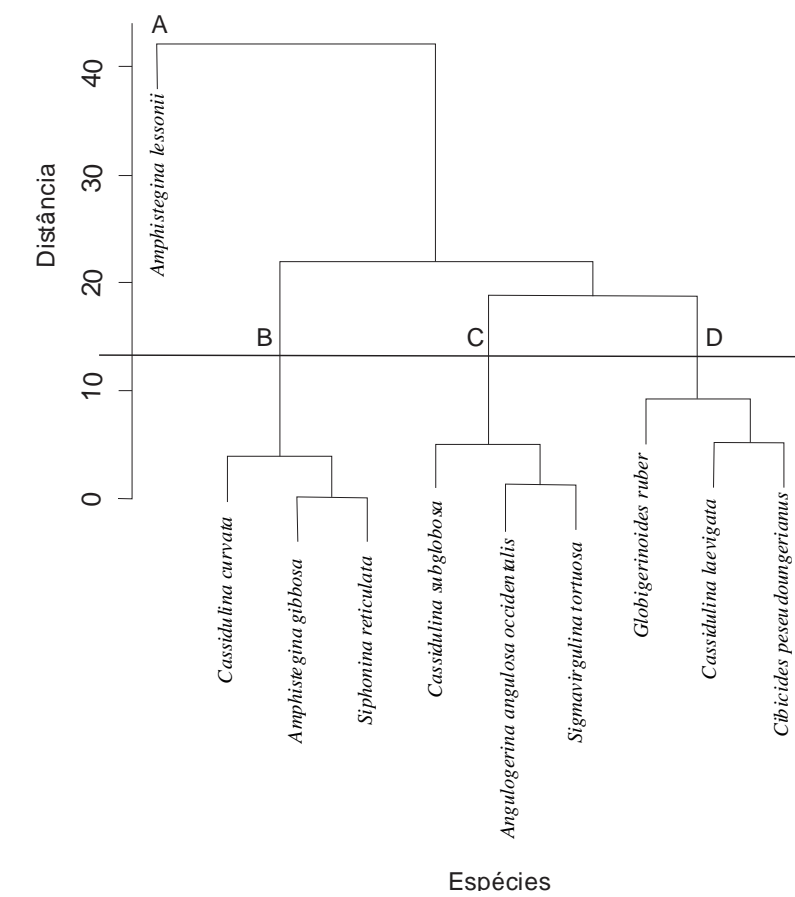


Figura 5 Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as espécies do talude, a partir das amostras (localização)

No Grupo A encontrou-se individualizada a espécie *Amphistegina lessonii*, com elevada abundância na amostra 146 (112 espécimes), e baixa abundância na amostra 131 (4 espécimes). Esta espécie foi a mais frequente do talude (275 indivíduos), apresentando, em média, 45,8 indivíduos neste grupo. No Grupo B foram reunidas as espécies *Amphistegina gibbosa* (28 indivíduos), *Cassidulina curvata* (33 indivíduos) e *Siphonina reticulata* (47 indivíduos). A média de indivíduos neste grupo variou de 4,7 a 7,8. No Grupo C foram encontradas as espécies *Sigmavirgulina tortuosa*, *Angulogerina angulosa occidentalis* e *Cassidulina subglobosa*, com abundância absoluta 61, 69 e 71 espécimes, respectivamente. A média de indivíduos neste grupo variou de 10,2 a 11,8. No Grupo D foram reunidas as espécies mais frequentes do talude após *Amphistegina lessonii*, sendo elas: *Cassidulina laevigata* (97 espécimes), *Cibicides pseudoungerianus* (108 espécimes) e *Globigerinoide ruber* (127 espécimes). A média de indivíduos neste grupo variou de 16,2 a 21,2. (Tab. 8 e 9).

Para aplicação do método k-Médias foram utilizadas como sementes iniciais os quatro grupos formados pelo método de Ward. Foi observado que as 10 espécies do talude foram classificadas nos mesmos agrupamentos pelos dois métodos. O coeficiente R^2 obtido foi 89%, indicando que os grupos encontrados são homogêneos internamente.

Tabela 8 Número de indivíduos das espécies do talude

Espécies / Amostras	121	127	131	135	146	166	Total
<i>Amphistegina gibbosa</i>	0	1	0	1	4	22	28
<i>Amphistegina lessonii</i>	44	25	4	5	112	85	275
<i>Angulogerina angulosa occidentalis</i>	1	26	34	0	7	1	69
<i>Cassidulina curvata</i>	0	0	1	25	1	6	33
<i>Cassidulina laevigata</i>	17	38	1	15	16	10	97
<i>Cassidulina subglobosa</i>	5	7	43	9	7	0	71
<i>Cibicides pseudoungerianus</i>	46	20	12	10	15	5	108
<i>Globigerinoides ruber</i>	20	20	17	46	16	8	127
<i>Sigmavirgulina tortuosa</i>	5	26	17	2	10	1	61
<i>Siphonina reticulata</i>	2	3	4	5	6	27	47

Tabela 9 Medidas descritivas dos agrupamentos de espécies encontrados no talude pelo método de Ward

Variáveis (FA)	Grupo A		Grupo B		Grupo C		Grupo D	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
<i>Amphistegina gibbosa</i>			4,7	8,62				
<i>Amphistegina lessonii</i>	45,8	44,20						
<i>Angulogerina angulosa occidentalis</i>					11,5	14,76		
<i>Cassidulina curvata</i>			5,5	9,81				
<i>Cassidulina laevigata</i>							16,2	12,22
<i>Cassidulina subglobosa</i>					11,8	15,57		
<i>Cibicides pseudoungerianus</i>							18,0	14,60
<i>Globigerinoides ruber</i>							21,2	12,94
<i>Sigmavirgulina tortuosa</i>					10,2	9,75		
<i>Siphonina reticulata</i>			7,8	9,50				

FA= frequência absoluta.

O dendrograma apresentado na figura 6 corresponde a análise de agrupamento entre as 38 amostras (modo Q), segundo os teores sedimentológicos, representados pelas cinco variáveis (teores de carbonato, siliciclasto, areia, cascalho e lama) presentes em cada amostra.

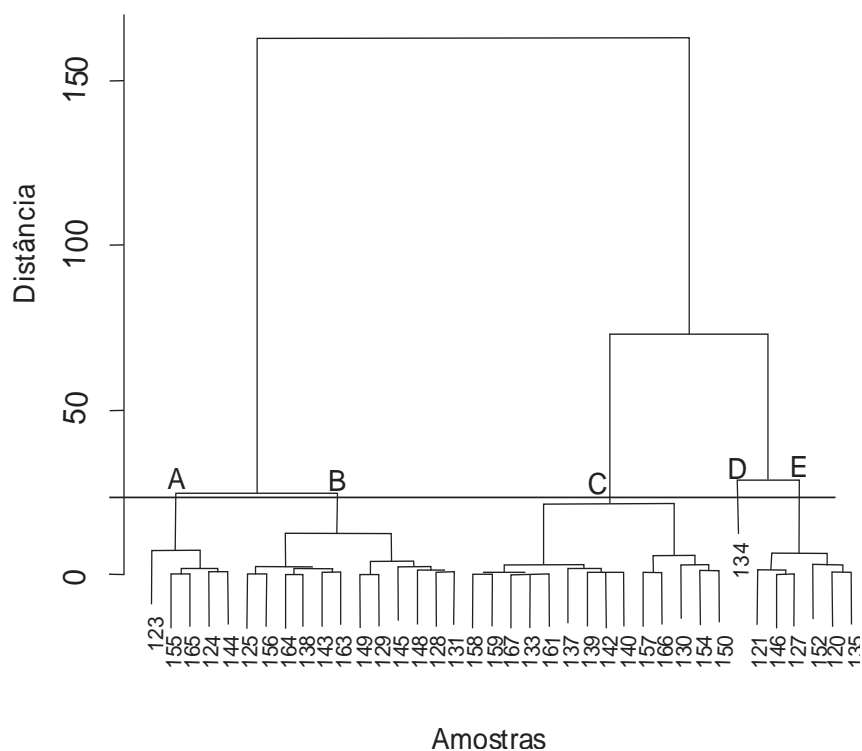


Figura 6 Dendrograma do Método de Ward com base na distância euclidiana quadrática agrupando as amostras considerando os teores sedimentológicos (Carbonato, Siliciclástico e Granulometria)

O valor do índice de correlação cofenética obtido para essa associação de amostras correspondeu a 0,7. No dendrograma foi realizado um corte na distância 24, onde foram observados cinco grupos A, B, C, D, E assim definidos: O Grupo A foi formado por cinco amostras, constituídas por altas taxas de areia

(em média 96,1%) e siliciclasto (em média 65,5%) e baixos teores de cascalho (em média 3,2%), carbonato (em média 34,5%) e lama (em média 0,8%).

No Grupo B foram reunidas doze amostras com alto teor de areia (em média 81,5%), carbonato (em média 58,2%) e cascalho (em média 15%), e baixas taxas de siliciclasto (em média 41,8%) e lama (em média 3,6%).

Os grupos A e B foram similares em relação aos teores de areia e carbonato, mas diferiram, principalmente, em relação ao teor de cascalho. Dessa forma o grupo A foi caracterizado pela presença de areia siliciclasta e o grupo B foi caracterizado pela presença de areia carbonática cascalhosa.

No Grupo C foram reunidas 14 amostras com alto teor de areia (em média 60,1%), carbonato (em média 91,8 %) e cascalho (em média 33,8%) e baixas percentagens de siliciclasto (em média 8,2%) e lama (em média 6,1%).

No Grupo D ficou isolada a amostra 134, que apresentou índice bastante elevado de lama (em média 67,9%) e de carbonato (em média 81%), e baixas percentagens de areia (em média 28,2%), siliciclasto (em média 19%) e de cascalho (em média 3,8%). No Grupo E foram agrupadas seis amostras que apresentaram altos teores de areia (em média 69,5%), carbonato (em média 80,7 %) e lama (em média 24,8%), e baixas taxas de siliciclasto (em média 19,3%) e cascalho (em média 5,6 %).

Os grupos C e E foram similares em relação aos teores de carbonato e areia, mas diferiram, principalmente, em relação ao teor de lama e teor de cascalho. Portanto, o grupo C foi caracterizado pela presença de areia carbonática cascalhosa e o grupo E foi caracterizado pela presença de areia carbonática lamosa.

O grupo D foi similar aos grupos C e E em relação ao teor de carbonato, mas diferiu desses grupos, principalmente, pelo alto teor de lama e baixo teor de areia. Dessa forma o grupo D foi caracterizado pela presença de lama carbonática arenosa (Tab.10).

Tabela 10 Valores percentuais de carbonato, siliciclásto, cascalho, areia e lama dos sedimentos em cada amostra. Carb = carbonato; Silic = Siliciclasto

AMOSTRAS	CARB	SILIC	CASCALHO	AREIA	LAMA
123	6,3	93,7	0,5	99	0,5
124	30,7	69,3	0,6	98,7	0,7
125	37,3	62,7	20	78,6	1,4
156	40,3	59,7	23,6	76,4	0
164	53	47	14,9	84,8	0,3
155	50,3	49,7	7	92,4	0,6
143	54,7	45,3	18,7	73,4	7,9
154	83	17	35,3	56	8,7
138	56,3	43,7	17,5	82	0,5
144	38	62	4,4	95	0,6
133	93	7	32,8	66,5	0,7
161	90,7	9,3	30,5	68,4	1
163	43,7	56,3	16,6	75,3	8,1
148	63	37	6,1	93,3	0,7
149	74,7	25,3	21,8	76,9	1,3
165	47	53	3,3	95,3	1,4
130	70,7	29,3	38	42,1	19,9
120	72,7	27,3	3,4	72,5	24,1
129	69	31	18,8	76,9	4,4
157	93	7	52,1	45,9	2
134	81	19	3,8	28,2	67,9
142	98,3	1,7	27,8	62,7	9,5
139	88,7	11,3	22,6	66,5	10,9
145	82	18	3,7	92,9	3,4
158	97,7	2,3	34,6	64,7	0,7
159	96,3	3,7	33	66	1
140	100	0	22,1	67,7	10,2
167	89,3	10,7	27,1	70,3	2,6
137	92	8	19	78,2	2,8
152	67,3	32,7	18,3	60,7	21
150	92,3	7,7	43,9	47,9	8,1
128	63	37	12	82,9	5
166	100	0	55	38,4	6,6
146	94,7	5,3	4,9	72,5	22,6
131	61,3	38,7	6	84,1	9,9
127	93,3	6,7	2,9	76,3	20,8
121	87,7	12,3	4,5	66,7	28,8
135	68,7	31,3	0	68,4	31,6

O método k-Médias foi aplicado utilizando-se como sementes iniciais os cinco grupos formados pelo método de Ward. Observou-se que das 38 amostras, 35 (92%) foram classificadas nos mesmos agrupamentos pelos dois métodos.

O coeficiente R^2 obtido foi 80% indicando que os grupos encontrados são homogêneos internamente. Entretanto, analisando as medidas descritivas verificou-se que as amostras 125, 156 e 163 pertencentes ao grupo B pelo método de Ward foram deslocadas para ao Grupo A, pelo método k-Médias. Observou-se que estas amostras apresentaram índices elevados de siliciclasto, principal característica das amostras do Grupo A. Portanto, o método k-Médias representou melhor o agrupamento A (Tab.11 e 12).

Tabela 11 Medidas descritivas dos agrupamentos de amostras formados pelo método de Ward.

Variáveis (%)	Grupo A		Grupo B		Grupo C		Grupo D		Grupo E	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
Carbonato	34,5	17,52	58,2	13,56	91,8	7,72	81,0		80,7	34,27
Siliciclasto	65,5	17,52	41,8	13,55	8,2	7,72	19,0		19,3	11,89
Cascalho	3,2	2,74	15,0	6,60	33,8	10,68	3,8		5,6	33,85
Areia	96,1	2,77	81,5	6,53	60,1	11,98	28,2		69,5	23,82
Lama	0,8	0,37	3,6	3,48	6,1	5,57	67,9		24,8	4,43

Tabela 12 Medidas descritivas dos agrupamentos de amostras formados pelo método k-Médias

Variáveis (%)	Grupo A		Grupo B		Grupo C		Grupo D		Grupo E	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
Carbonato	36,7	13,71	64,1	9,61	91,8	7,72	81,0		80,7	34,27
Siliciclasto	63,3	13,71	35,8	9,56	8,2	7,72	19,0		19,3	11,89
Cascalho	9,5	9,18	13,3	6,62	33,8	10,68	3,8		5,6	33,85
Areia	88,8	10,25	83,0	6,86	60,1	11,98	28,2		69,5	23,82
Lama	1,7	2,64	3,7	3,44	6,1	5,57	67,9		24,8	4,43

4. Discussão dos resultados

A maior abundância de *Peneroplis carinatus* em plataforma interna, motivo pelo qual essa espécie se mostrou em um grupo a parte no dendograma (Grupo A - Figura 2), também foi observada por Figueiredo *et al.* (2011), em estudo de análise do zoneamento batimétrico das assembléias de foraminíferos da plataforma e do talude continentais do município de Conde (BA).

A associação observada entre as espécies *Peneroplis proteus* e *Archaias angulatus* nos grupos B e C na plataforma interna, assemelha-se a registrada por Bruno (2008), em batimetria entre 25-35m na plataforma continental de Ilhéus, na Bahia. Li *et al.* (1998) também exibiram *Archaias angulatus* como principal espécie em batimetria entre 12-40m, além das espécies *Astigerina carinata*, *Quinqueloculina agglutinans* e *Valvulina oviedoiana*, na baía aberta de Frank Sound – Grand Cayman próximo de Cuba. Segundo Murray (2006), a espécie *Archaias angulatus* é confinada no oceano Atlântico, principalmente Caribeano em profundidades até 20m. Araújo & Araújo (2010), observaram que a abundância de *Archaias angulatus* e *Peneroplis proteus* decresce no sentido da plataforma interna para a externa no litoral norte do Estado da Bahia.

Os gêneros *Archaias* e *Peneroplis* são denominados de “macroforaminíferos”, sendo característicos de ambientes de zona fótica com sedimentação carbonática. Segundo Hoheggerem *et al.* (1999), foraminíferos com maiores dimensões, a exemplo dos gêneros *Archaias* e *Peneroplis*, representam componentes de significativa importância nas comunidades de mares tropicais e subtropicais, sendo predominantemente encontrados desde a zona intermarés até o limite da zona eufótica.

Na plataforma média as espécies *Amphistegina lessonii*, *A. gibbosa*, *Archaias angulatus*, *Homotrema rubrum* e *Heterostegina suborbicularis* dentre outras, foram também encontradas por Andrade (1997) na margem costeira até

45m de profundidade no litoral norte do Estado da Bahia entre Itapuã e Praia do Forte.

Bruno (2008) encontrou as espécies *Amphistegina lessonii*, *Amphistegina gibbosa*, *Archaias angulatus*, *Discorbis mira*, *Peneroplis pertusus*, *Peneroplis proteus*, *Peneroplis carinatus* e *Cibicides aknerianus* em locais cuja batimetria estava entre 40-50m na plataforma de Ilhéus na Bahia. Neste trabalho foi registrada associação semelhante na plataforma externa. Figueiredo *et al.* (2011) também constatou expressiva abundância de *Amphistegina lessonii* e *A. gibbosa*, *Archaias angulatus* em plataforma média a externa do município de Conde (BA), reforçando a significativa presença destas espécies em plataformas de águas quentes e carbonáticas.

No talude os subgrupos foram formados considerando a abundância das espécies. A maior representatividade da espécie *Amphistegina lessonii* neste intervalo batimétrico reflete o transporte por fluxo de massa e/ou correntes de turbidez de sedimento transportado da plataforma continental. Leipnitz & Leipnitz (1996) relata as formas vivas dessa espécie com batimetria máxima de 129m, porque são epifaunais, epífitas, herbívoras e possuem simbiontes. *Amphistegina lessonii* é encontrada nas praias de Salvador, sendo comum em plataformas e áreas recifais do litoral norte da Bahia (Andrade, 1997; Machado *et al.*, 2006; Araújo & Araújo, 2010). Araújo & Machado 2008b assinala que a maior abundância dessa espécie encontra-se no talude em frente a recifes da praia de Guarajuba no litoral norte da Bahia. Em Porto Rico ela é relatada como uma das espécies mais abundantes entre as assembleias recifais (Seiglei, 1968). *Amphistegina lessonii* ocorre também, no Indo-Pacífico e Mediterrâneo do sul, alcançando 0-90m de profundidade (Murray, 2006) e no Hawaii é abundante em áreas com menos de 10m de profundidade (Hallock, 1984). O gênero *Cassidulina* tem sido registrado em uma ampla variedade de ambientes, de nerítico a batial (Miao & Thunell, 1993) e frequentemente responde à entrada de

nutrientes no meio ambiente (Nees,1998). *Cassidulina laevigata* é epifaunal, típica de plataformas banhadas por águas quentes do Atlântico (Mackensen & Hald, 1988), enquanto *Cassidulina subglobosa* apresenta-se geralmente em diversas batimetrias (Miao & Thunell, 1993). O gênero *Cibicides* ocorre geralmente em substrato duro, águas frias a quentes, em ambiente de plataforma até batial a 2000m de profundidade (Murray, 1991; 2006; Culver, 1988). Segundo Araújo & Araújo (2010), a abundância de *Amphistegina lessonii* e *A. gibbosa*, nesta área estudada, aumenta no sentido da plataforma interna para externa.

Os resultados observados a partir do agrupamento das 38 amostras segundo as variáveis sedimentológicas (carbonato, siliciclasto, areia, cascalho e lama) evidenciam as relações entre o tipo de sedimento com a batimetria e consequentemente com as espécies abundantes em cada subambiente (plataformas interna, média e externa, e talude). Essa relação está associada à efetiva contribuição dos de diferentes organismos na produção de bioclastos, nos distintos intervalos de profundidade. De acordo com Dias (2000), na plataforma continental brasileira observa-se uma grande variação nos tipos de bioclastos tanto regionalmente quanto em função da profundidade.

Considerando os dados obtidos neste trabalho, o grupo C (Figura 6) foi caracterizado pela presença de areia carbonática cascalhosa e suas amostras estão localizadas, predominantemente, nas plataformas média e externa e talude. Bruno (2008) registra sedimento constituído de areia e cascalho carbonático em batimetria entre 40-50m na plataforma de Ilhéus, na Bahia.

Andrade (1997), em estudo envolvendo foraminíferos na zona de transição carbonato-siliciclasto entre Itapuã e Praia do Forte, no litoral norte da Bahia, encontrou nas amostras silicicláticas de areia fina e lamosa as formas: *Ammonia beccarii*, *Nonionella atlântica*, *Hanzawaia concentrica*, *Elphidium poeyanum*, *Bolivina compta*, *Discorbis mira*, *D. floridanus* e *Cibicides*

aknerianus. Nas regiões carbonáticas, onde ocorrem recifes de coral (Praia do Forte) foram encontradas as espécies: *Amphistegina lessonii*, *A. gibbosa*, *Archaias angulatus*, *Homotrema rubrum*, *Heterostegina suborbicularis*, dentre outras. Macedo & Machado (1995), assinalaram as espécies *Amphistegina lessonii* e *Archaias angulatus* como formas dominantes associadas ao substrato arenoso. Li *et al.* (1998) registrou a espécie *Archaias angulatus* em sedimento siliciclastos e *Amphistegina gibbosa* em sedimento carbonático na baía de Frank Sound – Cayman, próximo de Cuba. Renema (2006), na plataforma leste da Indonésia, no arquipélago de Supermonde, que consiste de sedimento carbonático-recifal encontrou *Amphistegina papillosa*, *A. radiata*, *Operculina ammonoides* e *Parasorites orbitolitoïdes*. As espécies *Peneroplis pertusus*, *Peneroplis proteus* e *Archaias angulatus* são citadas como formas representantes do ambiente carbonático por Bruno (2008) e Machado *et al.* (2006) na Bahia e por Batista *et al.* (2007) no ambiente recifal dos Parrachos de Maracajaú, Rio Grande do Norte.

5. Conclusões

A utilização conjunta dos dois métodos no agrupamento das espécies (modo R) e das amostras (modo Q) permitiu determinar através do método de Ward o número de grupos adequados a serem utilizados na análise dos dados. O método k-Médias otimizou estes resultados ao classificar as espécies e as amostras nos grupos formados, demonstrando a homogeneidade dentro dos grupos e heterogeneidades entre grupos.

Os resultados obtidos nesse trabalho evidenciaram que a plataforma interna do litoral norte da Bahia apresenta sedimento constituído de areia siliciclástica e as espécies de foraminíferos predominantes são: *Peneroplis carinatus*, *Archaias angulatus* e *Peneroplis proteus*. Além dessas espécies

formam a assembléia dessa plataforma as seguintes espécies: *Textularia agglutinans*; *Massilina pernambucensis*; *Quinqueloculina moynensis*; *Q. microcostata*; *Textularia gramen*; *Peneroplis bradyi*; *Textularia candeiana*.

A plataforma média é caracterizada por sedimento composto de areia siliciclástica e carbonática apresentando principalmente as espécies: *Amphistegina lessonii*, *A. gibbosa* e *Peneroplis carinatus*. Também compõem a assembléia dessa plataforma as espécies: *Quinqueloculina bicarinata*; *Hanzawaia concentrica*; *Elphidium discoidale*; *Hanzawaia bertheloti*; *Massilina pernambucensis*; *Q. microcostata*; *Textularia agglutinans*; *Q. bicostata*; *Articulina atlantica*; *Peneroplis proteus*; *Q. disparilis curta*; *Textularia candeiana*; *T. gramen*; *Peneroplis bradyi*. A plataforma média é a área de transição da sedimentação de areia siliciclástica da plataforma interna e areia carbonática da plataforma externa e do talude.

Na plataforma externa o sedimento é formado por areia carbonática cascalhosa e as espécies de foraminíferos predominantes são: *Amphistegina lessonii*, *A. gibbosa*, *Peneroplis carinatus*, *Globigerinoides ruber* e *Discorbis mira*. Também fazem parte da assembléia dessa plataforma as espécies: *Peneroplis bradyi*; *Elphidium discoidale*; *Bigenerina nodosaria*; *Hanzawaia concentrica*; *Quinqueloculina bicarinata*; *Textularia gramen*; *Q. parkeri*.

No talude observa-se sedimento formado por areia carbonática lamosa tendo as seguintes espécies predominando: *Amphistegina lessonii*, *Cassidulina laevigata*, *Cibicides pseudoungerianus* e *Globigerinoide ruber*. A assembleia de foraminíferos presentes no talude é composta também por: *Cassidulina curvata*; *Siphonina reticulata*; *C. subglobosa*; *Angulogerina angulosa angulosa*; *Sigmavirgulina tortuosa*; *C. laevigata*; e *Cibicides pseudoungerianus*.

Referências

Ab'Saber, A.N., Tundisi, J.G., Forneris, L., Marino, M.C. & Rocha, O., Tundisi, T., Schaeffer-Novelli, Y., Vouno, Y.S. & Watanabe, S. 1997. *Glossário de ecologia*. 2ª ed., Editora Aciesp, São Paulo, SP, Brasil, 352p.

Andrade, E. J. 1997. Distribuição dos Foraminíferos Recentes na Transição Carbonatos / Siliciclásticos na Região de Praia do Forte, Litoral Norte do Estado da Bahia, 111p, 11estampas. Dissertação de Mestrado, Instituto de Geociências da Universidade Federal da Bahia.

Araújo, H.A.B. 2009. Assembléias de foraminíferos indicadores de mudanças ambientais no complexo recifal de Abrolhos, Bahia, 209 p.,7 estampas.Tese de Doutorado em Geologia, Universidade Federal da Bahia, BA, Brasil.

Araújo, T.M.F. & Araújo H.A.B. 2010. Assembléias de foraminíferos encontradas nos sedimentos superficiais da plataforma continental e talude superior do norte da Bahia. *Revista de Geologia*, 23 (1): 115-134.

Araújo, T.M.F. & Machado, A.J.M. 2008a. Eventos bioestratigráficos, paleoclimáticos e paleobatimétricos do talude continental superior da Bahia, evidenciados através do estudo da fauna dos Foraminíferos. *Gravel*, 6(1): 27-45.

Araújo, T.M.F. & Machado, A.J.M. 2008b. Foraminíferos da subsuperfície do talude continental superior do norte da Bahia, Brasil. *Revista de Geologia*, 21(1): 49-77.

Araújo, T.M.F. & Machado, A.J.M. 2008c. Análise sedimentar e micropaleontológica (Foraminíferos) de Seções Quaternárias do Talude Continental Superior do Norte da Bahia, Brasil. *Revista Pesquisas em Geociências*, 35(2): 97-113.

Araújo, T.M.F. 2004. Estudo da microfauna de foraminíferos do sedimento da superfície e da subsuperfície da plataforma e do talude continentais da região norte do Estado da Bahia (Salvador à Barra do Itariri), 528p. Tese de Doutorado em Geologia, Universidade Federal da Bahia, Bahia, Brasil.

Asmus, H.E. 1984. *Geologia da margem continental brasileira*. In: Schobbenhaus, C., Campos, D. A., Derzei, G.R., Asmus, H.E. (eds), Geologia do Brasil-Texto explicativo do mapa geológico do Brasil e da área adjacente incluindo depósitos minerais, escala 1:500.000. Departamento Nacional de Produção Mineral, 501p.

Batista, D.S., Vilela, C.G. & Koutsoukos, E.A.M. 2007. Influência dos Fatores Ambientais na Preservação da Microfauna de Foraminíferos Bentônicos no Ambiente Recifal dos Parrachos de Maracajaú, RN, Brasil. *Anuário do Instituto de Geociências*, 30: 92-103.

Bergue C.T. & Coimbra J.C. 2008. Abordagens faunísticas e geoquímicas em microfósseis calcários e suas aplicações à paleoceanografia e paleoclimatologia. *Boletim do Museu Paraense Emílio Goeldi, Ciências Naturais*, 3(2): 115-126.

Bruno, R.L.M. 2008. Avaliação das assembleias de foraminíferos na Plataforma de transição carbonato-siliciclasto, região de Ilhéus, Bahia, 121p., 6 est. Dissertação de Mestrado, Instituto de Geociências da Universidade Federal da Bahia, BA, Brasil.

Bussab, W.O., Miazaki, É.S. & Andrade, D.F. 1990. *Análise de agrupamentos*, São Paulo, IME-USP.

Clarke, K.R. & Warwick, R.M. 1994. *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*. Plymouth: Plymouth Marine Laboratory, 144 p.

Clarke, K.R. & Warwick, R.M. 2001. *Change in marine communities: an approach to statistical analysis and interpretation*. 2nd ed., Plymouth, PRIMER-E, UK, 172p.

Culver, S.J. 1988. New Foraminiferal Depth Zonation of the Northwestern Gulf of Mexico. SEPM, *Palaios*, 3: 39-85.

Dajoz, R. 1983. *Ecologia Geral*. Petrópolis, Vozes, 4^a ed., 472p.

Dias, G.T.M. 2000. Granulados bioclásticos – algas calcárias. *Revista Brasileira de Geofísica*, 18(3): 307-318, 2000.

Drooger, C.W. & Kaasschietter, J.P. H. 1958. Foraminifera of Orenoco-Trinidad-Paria shelf expedition. Verhand. Kon. Nedel. Akad. Wetensch. Afd. Natnurf., 4 (22): 108.

Eichler, P. P. B., Eichler, B. B., Gupta, B. K. S. & Rodrigues, A.R. 2012. Foraminifera as indicator of marine pollutant contamination on inner continental shelf of southern Brazil. *Marine Pollution Bulletin*, 64: 22-30.

França, A.M.C. 1979. Geomorfologia da Margem Continental Leste Brasileira e da Bacia Oceânica Adjacente. In: Chaves, H.A.E. (ed), *Geomorfologia da Margem Continental Brasileira e das Áreas oceânicas Adjacentes*. Série Projeto REMAC, v.7, p.89-127.

Ferreira, D. F. 2008. *Estatística multivariada*, Lavras, Editora UFLA, 662p.

Figueiredo, SMC; Machado, A.J.; Araújo, T.M.F. & Araújo, H.A.B. 2011. Zoneamento batimétrico das assembleias de foraminíferos da plataforma e talude continentais do município de Conde, Bahia. *Revista Gravel*, 9(1): 1-10.

Hair, J. F., Tatham, R. L., Anderson, R. E. & Black, W. C. 2005. *Análise multivariada de dados*. Porto Alegre: Bookman, 593p.

Hallock, P. 1984. Distribution of selected species of living algal symbiont-bearing foraminifera on two Pacific coral reefs. *Journal Foraminiferal Research*, 23(3): 162-179.

Hohenegger, J.; Yordanova, E.; Nakano, Y. & Tatzreiter, F. 1999. Habitats of larger foraminifera on the upper reef slope of Sasoko Island, Okinawa, Japan. *Marine Micropaleontology*, 36: 109-168.

Landim, P.M.B. 2011. *Análise estatística de dados geológicos multivariados*, São Paulo. Oficina de Textos, 239p.

Lattin, J., Carroll, J.D. & Green, P.E. *Análise de dados multivariadas*, São Paulo, Cengage Learning, 2011. 455p.

Leipnitz, I.I. & Leipnitz, B. 1996. Nova proposta de subdivisão biogeográfica com base nos foraminíferos para as porções norte e nordeste da Plataforma Continental Brasileira. In: XXXIX CONGRESSO BRASILEIRO DE GEOLOGIA, Salvador – Bahia, 1996. *Resumos Expandidos*, **39**: 270-273.

Li, C., Jones, B. & Kalbfleisch, W.B.C. 1998. Carbonate sediment transport path ways based on foraminífera: case study from Frank Sound, Grand Cayman, British West Indies. *Sedimentology* 45: 109-120.

Loeblich, A.R.Jr. & Tappan, H. 1988. *Foraminiferal genera and their classification*. Van Nostrand Reinhold Ed., New York, USA, 970p, 847 est.

Macedo, D.L.S. & Machado, A.J. 1995. A fauna de foraminíferos dos sedimentos superficiais de uma área selecionada na plataforma continental do estado da Bahia (Arembepe ao Morro de São Paulo). *Acta Geologica Leopoldensia*, **42**: 119-137.

Machado, A.J., Andrade, E. D. & Araújo, H.A.B. 2006. Fauna de Foraminíferos do Litoral Norte do Estado da Bahia. *Revista de Geologia*, 19(2): 147-154.

Mackensen, A. & Hald, M.1988. *Cassidulina teretis* Tappan and *C. laevigata* d'Orbigny: their modern and late Quaternary distribution in northern seas. *Journal of Foraminiferal Research*, 18: 16-24.

Miao, Q. & Thunell, R.C. 1993. Recent deep-sea benthic foraminiferal distributions in the south China and Sulu Seas. *Marine Micropaleontology*, 22: 1-32.

Mingoti, S. A. 2005. Análise de dados através de métodos de estatísticas multivariadas: uma abordagem aplicada. Belo Horizonte, Editora UFMG, 297p.

Murray, J.W. 1991. *Ecology and Palaeocology of benthic foraminífera*. Longman Scientific & Technical, 397p.

Murray, J.W. 2006. *Ecology and Applications of Benthic Foraminifera*. Editora: Cambridge University Press. 426p.

Nees, S. 1998. Ground truth in palaeoceanography – good news from the sea floor. Proceedings of the 6th International Conference on palaeoceanography. Lisbon. 5p.

Passos, R.F., Souza, S.H.M. & Mahiques. M. M. 2001. Mudanças ambientais ocorridas entre Abrolhos (BA) e Cabo Frio (RJ) ao longo do holoceno, e sua resposta nas associações de foraminíferos. *Pesquisa em Geociências*, 28(2): 223-231.

Pawlowski, J. & Holzmann, M. 2008. Diversity and geographic distribution of benthic foraminifera: a molecular perspective. *Biodivers Conserv.*, 17: 317–328.

PNUMA – Programa das Nações Unidas para o Meio Ambiente. 2004. *Perspectivas do Meio Ambiente Mundial 2002 – GEO-3*. Nairobi, Kenya, Disponível em: <www.unep.org/GEO/geo3/index.htm>. Acesso em: 16 abr. 2012.

R Development Core Team. 2012. *R: a language and environment for statistical computing*. Vienna, Austria. Disponível em: < <http://www.R-project.org> >. Acesso em: junho de 2012.

Rencher, C. A. 2002. *Methods of multivariate analysis*. New York: J. Wiley, 708p.

Renema, W. 2006. Large benthic foraminífera from the deep photic zone of a mixed siliciclastic-carbonate shelf off East Kalimantan, Indonesia. *Marine Micropaleontology*, 68:73-82.

Schlee, J. 1973. Atlantic Continental Shelf and Slope of the United States sediment texture of the northeastern part. *U.S. Geological Survey Professional Paper*, 529-L, 64p.

Seiglie, G. A. 1968. Relationship between the distribution of *Amphistegina* and the submerged Pleistocene reefs off western Puerto Rico. *Tulane Studies in Geology*, 6: 138-147.

Shepard, F.S. 1954. Nomenclature based on sand-silt-clay ratios. *Journal of Sedimentary Petrology*, 24(3): 151-158.

Testa, V. 1996. Quaternary sediments of the shallow shelf, Rio Grande do Norte, NE Brazil. London, 416p. PhD Thesis, Royal Holloway University of London.

Vieira, F.S. 2011. Caracterização microfaciológica e ambiental com base em foraminíferos bentônicos na plataforma continental da bacia de Campos, Rio de Janeiro. 137p. Tese de Doutorado em Geologia, Universidade Federal da Bahia, BA, Brasil. 137p.

Wentworth, C.K. 1962. A scale of grade and class terms for clastic sediments. *Journal Geology*, 30: 377-392.

ANEXOS

ANEXO A – Estampas

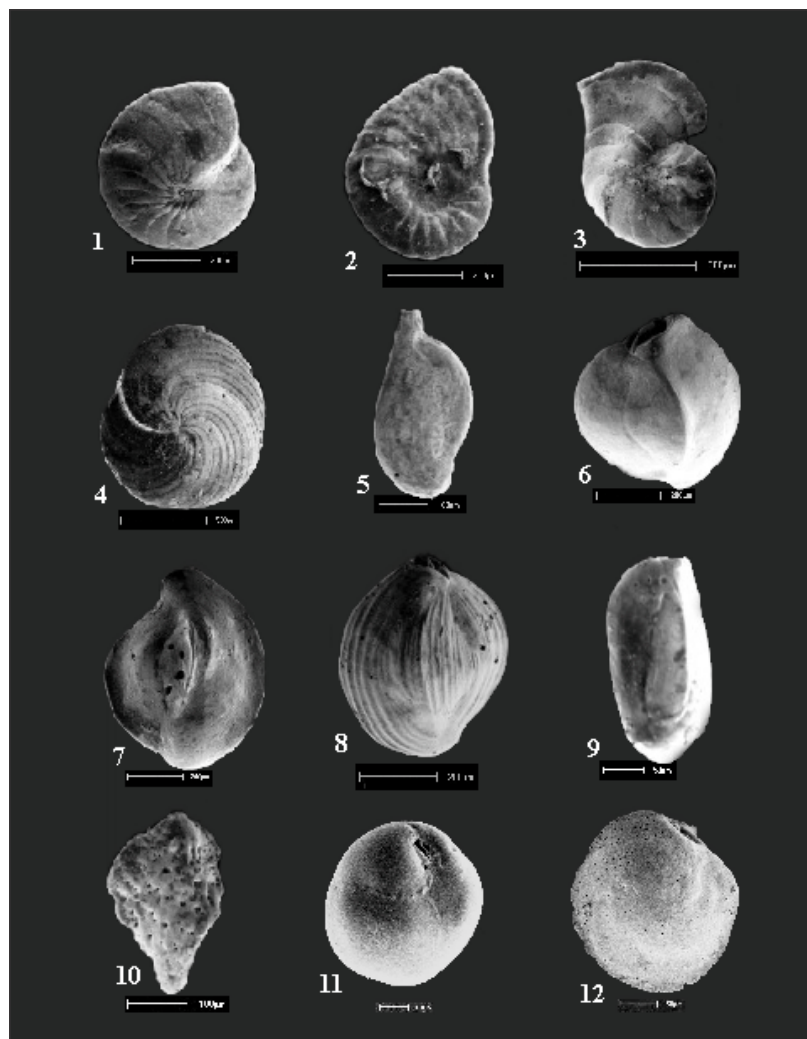


Figura 1 *Peneroplis carinatus*. 2. *P. bradyi*. 3. *P. proteus*. 4. *Archaias angulatus*. 5. *Massilina pernambucensis*. 6. *Quinqueloculina bicarinata*. 7. *Q. bicostata*. 8. *Q. microcostata*. 9. *Q. moynensis*. 10. *Sigmavirgulina tortuosa*. 11. *Cassidulina subglobosa*. 12. *C. curvata*

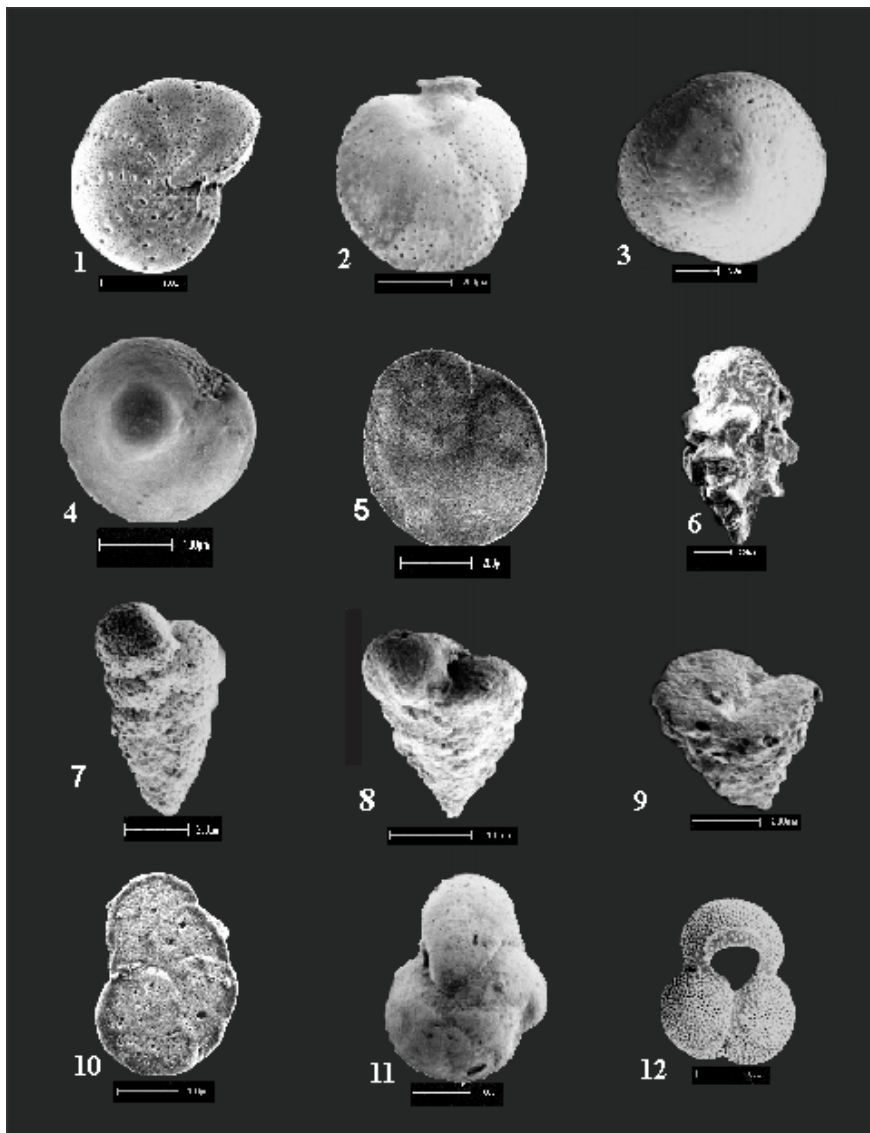


Figura 2 *Elphidium discoideale*. 2. *Siphonina reticulata*. 3. *Dicorbis mira*. 4. *Amphistegina gibbosa*. 5. *A. lessonii*. 6. *Bigenerina nodosaria*. 7. *Textularia agglutinans*. 8. *T. candeiana*. 9. *T. gramen*. 10. *Cibicides pseudoungerianus*, (lado dorsal). 11. *C. pseudoungerianus*, (lado ventral). 12. *Globigerinoides ruber*.

CONSIDERAÇÕES FINAIS

O objetivo principal desse trabalho foi utilizar métodos estatísticos para caracterizar a microfauna de assembleias de foraminíferos da plataforma e talude continentais do norte do Estado da Bahia.

O uso de semivariogramas mostrou eficiente na descrição da estrutura da dependência espacial das espécies de foraminíferos enquanto, por meio da krigagem, foi possível obter estimativas em locais não amostrados e, conseqüentemente, gerar mapas que permitiram entender o padrão da distribuição espacial das espécies de foraminíferos.

A análise multivariada de agrupamento permitiu não somente a caracterização sedimentológica da área de estudo mas também a detecção de grupos de espécies de foraminíferos.

Evidentemente, muitos estudos foram deixados de lado em virtude do tempo que foi destinado a esse projeto. Assim, como perspectivas de trabalhos futuros, pretende-se utilizar a geoestatística na análise de dados obtidos de 40 amostras da microfauna de foraminíferos dos testemunhos do talude continental do Litoral Norte da Bahia. Pretende-se também, explorar a construção de intervalos de confiança para os índices usuais de riqueza e diversidade de espécies (ex. Shannon, Simpson) que são, em geral, utilizados de forma pontual em estudos ecológicos. Esse estudo iria abordar não somente a obtenção de distribuições amostrais para esses indicadores mas também o uso de métodos de computação intensiva.