

**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA DA UNIVERSIDADE
FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE FÍSICA E MATEMÁTICA**

**DISSERTAÇÃO DE
MESTRADO**

**DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO *SELENE* (CARANGIDAE) DA
COSTA NORDESTE DO BRASIL, UTILIZANDO TÉCNICAS DE MORFOMETRIA
GEOMÉTRICA**

AUTOR: JOSÉ DE MELO LIMA FILHO

**RECIFE – PERNAMBUCO – BRASIL
FEVEREIRO – 2004**

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE FÍSICA E MATEMÁTICA**

DOIS IRMÃOS - RECIFE - PE - BRASIL

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE FÍSICA E MATEMÁTICA**

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA

**DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO *SELENE* (CARANGIDAE) DA
COSTA NORDESTE DO BRASIL, UTILIZANDO TÉCNICAS DE MORFOMETRIA
GEOMÉTRICA**

JOSÉ DE MELO LIMA FILHO

RECIFE – PERNAMBUCO – BRASIL

FEVEREIRO– 2004

**DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO *SELENE*
(CARANGIDAE) DA COSTA NORDESTE DO BRASIL,
UTILIZANDO TÉCNICAS DE MORFOMETRIA
GEOMÉTRICA**

JOSÉ DE MELO LIMA FILHO

**DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO *SELENE*
(CARANGIDAE) DA COSTA NORDESTE DO BRASIL,
UTILIZANDO TÉCNICAS DE MORFOMETRIA
GEOMÉTRICA**

Dissertação submetida ao Programa
de Pós-Graduação em Biometria da
Universidade Federal Rural de
Pernambuco, para obtenção do
título de Mestre em Biometria.

Área de Concentração: Métodos Estatísticos Aplicados às
Ciências Biológicas

ORIENTADORA: PROF^a. Dra. ROSÂNGELA PAULA LESSA

CO-ORIENTADOR: PROF. Dr. BORKO STOSIC

RECIFE – PERNAMBUCO – BRASIL

FEVEREIRO – 2004

Catálogo na Fonte
Setor de Processos Técnicos da Biblioteca Central – UFRPE

L732d Lima Filho, José de Melo
Discriminação de espécies do gênero *Selene* (Carangidae)
da costa nordeste do Brasil / José de Melo Lima Filho. –
2004.
59 f. : il.

Orientador: Rosângela Paula Lessa.
Dissertação (Mestrado em Biometria) – Universidade Fe-
deral Rural de Pernambuco. Departamento de Física e Mate-
mática.
Inclui bibliografia e anexo.

CDD 574.018 2

1. Morfometria
2. *Selene*
3. Marcos anatômicos
4. Centróide
5. Análise multivariada
6. Biometria
 - I. Lessa, Rosângela Paula
 - II. Título

**DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO *SELENE*
(CARANGIDAE) DA COSTA NORDESTE DO BRASIL, UTILIZANDO
TÉCNICAS DE MORFOMETRIA GEOMÉTRICA**

José de Melo Lima Filho

Dissertação defendida e aprovada pela banca examinadora.

ORIENTADORA:

Prof.^a. Dra. Rosângela Paula Lessa – UFRPE

EXAMINADORES:

Prof. Dr. Eduardo César de Miranda Loureiro - UPE

Prof. Dr. Borko Stosic - UFRPE

Prof. Dr. Romildo Albuquerque Nogueira - UFRPE

DEDICATÓRIA

Dedico este trabalho ao meu pai, José de Melo Lima (*in memorian*), à minha mãe, Silvana de Souza Melo (*in memorian*), à minha esposa, Maria Auricéia Vasconcelos Fraga, aos meus filhos Clarisse, Marina, João e também aos meus irmãos, Vanildo (*in memorian*), Valdemar, Vanice e Vandui.

AGRADECIMENTOS

À Professora Dra. Rosângela Lessa, pela orientação e apoio imprescindíveis à conclusão deste trabalho.

Ao Professor Dr. Borko, pela valiosa e eficiente atuação como co-orientador em todos os momentos.

A todos os professores de Biometria e, em particular, aos doutores Eufrázio de Souza Santos, Gauss Cordeiro, Rinaldo Luiz Caraciolo, Paulo de Paula Mendes e Maria Adélia, com os quais tive interação mais direta ao cursar disciplinas fundamentais para minha formação. Aos professores que fizeram parte das bancas dos meus Seminários: Doutores Borko, Romildo Nogueira e Socorro Brasileiro, pelas sugestões valiosas e pelo incentivo constante.

A todos os colegas da Biometria, especialmente a Lázaro, Niedja, Patrícia, Cristina, Rômulo, Jairo, José Carlos, Cleto, Dirac, Edwird, Moacir, Wellington, Telmo e Gilvandro pelo convívio, amizade e ajuda mútua.

A todos os que participam do grupo de pesquisa da Dra. Rosângela, especialmente a Paulo Duarte, José Lúcio, Chico, Mércia, Marcelo, Jailson e Kaio pelo convívio, amizade e grandiosa ajuda.

Aos amigos José Wilson, José Luiz, Joca Oliveira, Maria Vandelize, Adriano, Amâncio, André, Fred, Lídia Gonçalves, Bayma, Caitano Cintra e Olavo Otávio Nunes por serem amigos de todas as horas.

Aos amigos Hélio Nemésio do Nascimento e Aloísio Teles, que me estimulavam a fazer Mestrado. Quanta saudade!

A todos os colegas do Centro Federal de Educação Tecnológica de Pernambuco – CEFET PE, e em especial aos Profesores José Wilson Vieira, Ruberval e Tetsuo Usui pelos ensinamentos, incentivos e ajuda ao longo de todo o curso.

À todos que contribuíram de alguma forma para realização dessa dissertação.

*QUE OS NOSSOS ESFORÇOS
DESAFIEM AS IMPOSSIBILIDADES.
LEMBRAI-VÓS DE QUE AS
GRANDES PROEZAS DA HISTÓRIA
FORAM CONQUISTAS DO QUE
PARECIA IMPOSSÍVEL.*

Charlie Chaplin

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE FÍSICA E MATEMÁTICA
BIOMETRIA**

DISSERTAÇÃO DE MESTRADO

**DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO *SELENE*
(CARANGIDAE) DA COSTA NORDESTE DO BRASIL,
UTILIZANDO TÉCNICAS DE MORFOMETRIA
GEOMÉTRICA**

ALUNO: JOSÉ DE MELO LIMA FILHO

ORIENTADORA: PROF^a Dra. ROSÂNGELA LESSA

CO-ORIENTADOR: PROF. Dr. BORKO STOSIC

Recife, Fevereiro de 2004

**DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO
SELENE (CARANGIDAE) DA COSTA NORDESTE DO
BRASIL, UTILIZANDO TÉCNICAS DE
MORFOMETRIA GEOMÉTRICA**

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

RESUMO

ABSTRACT

1	INTRODUÇÃO	8
2	REVISÃO DE LITERATURA	10
2.1	ANÁLISE MULTIVARIADA DE DADOS MORFOMÉTRICOS	12
2.2	MORFOMETRIA GEOMÉTRICA	12
3	MATERIAIS E MÉTODOS	18
3.1	MARCOS ANATÔMICOS	20
3.2	TRATAMENTO DE IMAGEM	21
3.2.1.	Obtenção das Distâncias Entre os Marcos Anatômicos	25
3.3	MÉTODO 1- ANÁLISE MULTIVARIADA DOS DADOS	26
3.3.1.	Análise Estatística	26
3.4	MÉTODO 2 – ANÁLISE DE DEFORMAÇÕES	28
3.4.1.	Análise de Deformações Relativas (Função Thin Plate Splines).....	28
4	RESULTADOS E DISCUSSÃO	31
4.1	MÉTODO 1	31
4.1.1.	Análise de Variância Multivariada (MANOVA).....	31
4.1.2.	Análise De Componentes Principais	31
4.1.3.	Análise De Variáveis Canônicas	34
4.1.4.	Análise Discriminante	36
4.2	MÉTODO 2	36
5	CONCLUSÕES	43
	REFERÊNCIAS	44
	APÊNDICE 1	48
	APÊNDICE 2	55
	ANEXO	60

LISTA DE FIGURAS

Figura 1: Área da cadeia da amostra de <i>S. spixii</i> e <i>S. vomer</i> no Rio Grande do Norte (Caiçara do Norte).....	9
Figura 2: Distribuição de frequências de comprimentos zoológicos de (a) <i>S. spixii</i> e (b) <i>S. vomer</i>	18
Figura 3: Imagens de peixes <i>S. spixii</i> e <i>S. vomer</i> (à direita) em 3 fases distintas. ...	19
Figura 4: Rede de treliças para <i>Selene vomer</i> obtida a partir dos 9 marcos anatômicos indicados na figura.	20
Figura 5: Imagem 4 no formato tif, ajustada para uma largura de 250 pixels.....	21
Figura 6: Janela principal do Scion Image, exibindo a imagem em tons de cinza do peixe 4.....	22
Figura 7: Diálogo <i>Import File</i> do Scion Image com os dados necessários para acessar a matriz 3D.	23
Figura 8: Matriz 3D descompactada no Scion Image.....	24
Figura 9: Marcos anatômicos (ampliados) colocados no peixe 4, usando o Scion Image	25
Figura 10: Peixe nº 279, que corresponde à linha 199 da matriz 2D.	26
Figura 11: Gráficos dos primeiros componentes Principais	33
Figura 12: Projeção dos escores canônicos.....	35
Figura 13: Configuração média (ou consenso) dos 389 peixes da amostra estudada	37
Figura 14: Configuração dos 389 peixes com vetores indicando direção e sentido de cada deformação em cada marco anatômico em relação à CM	37
Figura 15: Componentes uniformes x e y.....	39
Figura 16: Ordenação dos espécimes por deformações relativas. No ponto origem situa-se a CM, e os pontos evidenciados em vermelho estão associados a <i>S. spixii</i> , enquanto os evidenciados em azul estão associados <i>S. vomer</i>	41
Figura 17: Grades de deformação. (a) configurações de <i>S. spixii</i> (peixes 64, 4 e 121 de cima para baixo), (b) grade da configuração média, (c) configurações de <i>S. vomer</i> (peixes 365, 375 e 312, de cima para baixo). Os espécimes escolhidos estão posicionados em ordem crescente das distâncias à CM, como está melhor evidenciado na Figura 16.	42

LISTA DE TABELAS

Tabela 1: Variáveis morfológicas no estudo morfométrico de peixes <i>Selene spixii</i> e <i>Selene vomer</i>	21
Tabela 2: Matriz (parcial) correspondente à imagem 4 da Figura 6	22
Tabela 3: Quadro de Análise da Variância Multivariada (MANOVA) - Delineamento Inteiramente Casualizado.....	27
Tabela 4: Matriz de Planejamento.....	27
Tabela 5: Análise de variância multivariada para as espécies <i>S. spixii</i> e <i>S. vomer</i> ...	31
Tabela 6: Resumo da Análise de Componentes Principais.....	32
Tabela 7: Coeficientes de ponderação (correlação das variáveis estudadas com os 3 primeiros componentes principais) e o coeficiente de alometria de cada variável	32
Tabela 8: Resumo da Análise de Variáveis Canônicas.....	34
Tabela 9: Coeficientes padronizados das 16 variáveis, das amostras combinadas de <i>S. spixii</i> e <i>S. vomer</i> no eixo canônico	35
Tabela 10: Número de Observações e Percentual Classificado	36
Tabela 11: Coordenadas da configuração média.....	36
Tabela 12: Resumo da análise das deformações relativas	38
Tabela 13: Matriz de correlação entre as variáveis tamanho do centróide e componentes uniformes	39
Tabela 14: Coeficientes padronizados nas amostras combinadas de <i>S. spixii</i> e <i>S. vomer</i> no único eixo canônico	39
Tabela 15: Coeficientes de ponderação na análise de deformações relativas para os três primeiros eixos principais	40

DISCRIMINAÇÃO DE ESPÉCIES DO GÊNERO *SELENE* (CARANGIDAE) DA COSTA NORDESTE DO BRASIL, UTILIZANDO TÉCNICAS DE MORFOMETRIA GEOMÉTRICA

Autor: José de Melo Lima Filho

Orientadora: Prof^a Dra. Rosângela Lessa

Co-orientador: Prof^o Dr. Borko Stosic

RESUMO

Os Carangidae são importantes componentes da ictiofauna em todas as fases da vida e as espécies *Selene spixii* (Lua-do-Caribe) Spix e Agassiz (1831) e *Selene vomer* (galo-de-penacho) Linnaeus (1758) estão distribuídas pelo Atlântico Oeste de Nova Escócia (Canadá) ao longo das costas do Golfo do México até a América do Sul. Esses peixes são encontrados em águas rasas costeiras, sobre fundos duros ou arenosos. Vivem em cardumes, mas também podem ser observados em pequenos grupos ou em pares. Alimentam-se de pequenos caranguejos, camarões, peixes e poliquetas. Sua carne apresenta um sabor excelente, sendo comercializado fresco. Estes peixes podem ser cultivados em aquários, são valorizados na pesca esportiva, e ambas espécies apresentam grande semelhança na fase jovem o que dificulta a identificação. O objetivo deste trabalho foi analisar os padrões de diferenciação morfológicos entre as espécies do gênero *Selene*, utilizando como método 1, técnicas de análise multivariada das 16 distâncias entre marcos anatômicos, definidas na rede de treliças; e como método 2, a análise de deformações relativas (função *thin plate spline*) sobre as coordenadas de 9 marcos anatômicos. Os dados morfométricos foram obtidos em imagens digitalizadas por *scanner*, as coordenadas dos marcos anatômicos pelo programa tpsDig, as distâncias obtidas por programa desenvolvido em linguagem C e a análise de deformações relativas através do programa tpsRelw. Pretendeu-se discriminar similaridades entre as duas espécies por meio de Componentes Principais, Análise de Variáveis Canônicas e Análise Discriminante com base em características morfométricas. Foram utilizados os seguintes marcos anatômicos: 1. ponta do focinho; 2. base da nadadeira ventral (pélvica); 3. início da primeira nadadeira dorsal; 4. início da nadadeira anal; 5. início da segunda nadadeira dorsal; 6. final da nadadeira anal; 7. final da segunda nadadeira dorsal; 8. início da nadadeira caudal inferior; 9. início da nadadeira caudal superior. No método 1, as medidas D 1-3 e D 3-4 foram as mais importantes, enquanto que no método 2, dentre as variações de forma, a que mais contribuiu foi a componente uniforme x.

Palavras Chaves: morfometria geométrica, *selene*, marcos anatômicos, centróide, análise multivariada.

DISCRIMINATION OF *SELENE* SPECIES (FAMILY CARANGIDAE) FROM NORTHEASTERN BRAZILIAN COAST, USING TECHNIQUES OF GEOMETRIC MORPHOMETRY

Author: José de Melo Lima Filho

Adviser: Prof^a Dra. Rosângela Lessa

Co-adviser: Prof^o Dr. Borko Stosic

ABSTRACT

Species of the Carangidae family are important components of fish fauna during all phases of the life span. Both *Selene spixii* (full moon fish) Spix and Agassiz (1831) and *Selene Vomere* (Atlantic look down) Linnaeus (1758) range in the Western Atlantic Ocean from New Scotia (Canada) to South America including the Gulf of Mexico. They are found in coastal shallow waters, either on hard or sandy bottoms. They usually establish schools, but are also observed in small groups or pairs. They feed on small crustaceans, shrimps, fish and worms. Their flesh tastes good; these species are trade fresh. They can also be reared, they are highly prized in sport fishing. Both species are very similar in early phases of their life cycle, which renders the identification at the species level quite difficult. The aim of the present study was to carry out analyses of the pattern of morphological differentiation between the above mentioned species of *Selene* genre, using as method 1, techniques of multivariate analysis of the 16 distances between landmarks, defined in the truss net; and as method 2, the relative warps analysis (thin plate spline function) on the coordinates of 9 landmarks. Morphometric data were obtained from scanner digitalized images and the coordinates of the landmarks by using tpsdigw program, the distances were obtained from a program developed in language C and the relative warps analysis from tpsRelw program. Similarities were discriminated between the studied species using both Principal Components Analysis, Canonic as well as Discriminant Analyses based on morphometric features. The landmarks used were: 1. tip of the nose, 2. base of pelvic fin, 3. beginning of the first dorsal fin, 4. beginning of anal fin, 5. beginning of the second dorsal fin, 6. end of anal fin, 7. end of the second dorsal fin, 8. beginning of lower caudal fin, 9. beginning of upper caudal fin. According to method 1, measures D 1-3 and D 3-4 were the most important whereas for method 2, among shape variations, the one which contributes the most was the component x.

Key words: geometric morphometric, *selene*, landmarks, centroid, multivariate analysis.

1 INTRODUÇÃO

As espécies estudadas estão distribuídas pelo Atlântico Oeste de Nova Escócia (Canadá) ao longo das costas do Golfo do México até a América do Sul. *Selene vomer* se distribui até o Uruguai. São encontradas em águas rasas costeiras, sobre fundos duros ou arenosos. Juvenis, podem ser encontrados em áreas estuarinas e livres em praias arenosas. Vivem em cardumes, mas também podem ser observadas em pequenos grupos ou em pares. Alimentam-se de pequenos caranguejos, camarões, peixes e poliquetas. Sua carne apresenta um sabor excelente, sendo comercializado fresco. Estes peixes podem ser cultivados e são valorizados na pesca esportiva (FISHBASE, 2003).

No nordeste do Brasil, as espécies *Selene spixii* e *Selene vomer* ocorrem nas mesmas áreas e apresentam grandes semelhanças nos estágios iniciais, o que constitui-se uma dificuldade na identificação ao nível de espécie. A identificação imprecisa de espécies pode levar a erro sobre aspectos populacionais como idade, crescimento e sexo, falseando a avaliação do estado de exploração dos estoques. Desta forma, a análise morfométrica contribui para a minimização dos erros, e para o conhecimento taxonômico das espécies.

Os peixes da família Carangidae são importantes componentes da ictiofauna em todas as fases da vida e por isso é desejável poder avaliar os impactos da pesca sobre suas populações. Essa necessidade implica na exigência de identificar as espécies dessa família corretamente utilizando-se os melhores métodos disponíveis.

A palavra espécie vem de uma raiz indo-européia extremamente antiga, SPEK, que significa “aquilo que se vê” – de onde também deriva, por exemplo, “retrospectiva”, olhar para trás (BARTLEBY, 2003). Isto significa que a forma é a mais antiga das fontes de informação que o homem usa sobre os organismos. Embora, somente ao longo do século XX, aconteceram várias tentativas para se definir metodologias adequadas para a análise da forma, sempre com barreiras difíceis de superar.

Somente a partir do final da década de 80, com a disponibilidade de ferramentas computacionais mais poderosas, novos conceitos e a solução de problemas matemáticos, ingressou-se em uma nova era nas análises morfométricas.

A Morfometria Geométrica corresponde a uma área de conhecimento extremamente nova, cujo alcance e extensão começam a ser compreendidas apenas agora. Hoje já corresponde a uma ferramenta importante em análises Taxonômicas e Sistemáticas e outras áreas, uma vez que aumenta a capacidade de reconhecer diferenças entre grupos em comparação com a “velha” morfologia. Os avanços recentes permitem melhor aproximação para medidas de variação de forma entre organismos e a relação dessas variações com o ambiente dos organismos.

O propósito do presente trabalho é apresentar um estudo feito em amostras de peixes da família Carangidae (perciformes) das espécies *selene spixii* (Spix e Agassiz, 1831) (lua-do-caribe) e *Selene vomer* (Linnaeus, 1758) (galo-de-penacho), capturados artesanalmente na praia Caiçara do Norte (RN), no intuito de discriminar similaridades entre indivíduos dentro de uma mesma espécie e entre espécies, com base em características morfométricas definidas no protocolo da “rede de treliças”.



Figura 1: Área da cadeia da amostra de *S. spixii* e *S. vomer* no Rio Grande do Norte (Caiçara do Norte)

2 REVISÃO DE LITERATURA

A Ciência Biologia Pesqueira, na qual se insere o presente estudo, teve origem na Europa, onde a exploração da solha (*Pleuronectes platessa*) e do bacalhau (*Gadus morhua*) com redes-de-arrasto, no século XIX, já era considerada uma atividade predatória. Havia uma crença generalizada de que as populações tinham capacidade infinita de produção, devido a enorme fecundidade de algumas espécies (FONTELES, 1989). Entretanto, no século XX, começaram a aparecer sinais de que essa produção era limitada, e daí a necessidade de se estudar a abundância e a estrutura dos estoques, para obter explicações sobre os declínios das biomassas coletadas anualmente.

No início, os estudos eram realizados com base nos aspectos biológicos como reprodução, alimentação, crescimento e migração, e a análise dos dados tinha caráter empírico, pois não existiam modelos matemáticos adequados à interpretação da dinâmica populacional e sua relação com a pesca. Em 1918, surgiu o primeiro trabalho utilizando análise matemática de dados, pelo pesquisador russo F. I. Baranov cuja teoria foi amplamente utilizada em trabalhos subseqüentes (FONTELES, 1989).

Sabe-se que a Biologia Pesqueira estuda as características biológicas da população a partir do conhecimento das características individuais que apresentam grande variabilidade devido às diferenças no patrimônio genético e às modificações no meio ambiente. Na avaliação dos processos dinâmicos por que passa uma população, o crescimento é um dos principais parâmetros pelos quais se medem as variações que ocorrem em seu tamanho e estrutura, como forma de compensação ao decréscimo da abundância em termos numéricos. Isto se deve ao fato de que os indivíduos de mesma idade, desde seu nascimento, são sujeitos a vários níveis de predação, que se reduzem à medida que aumentam de tamanho.

Assumindo que as amostras de diferentes populações exploradas devem apresentar diferenças nos caracteres morfométricos; os indivíduos de uma mesma população são biologicamente semelhantes entre si e constituem uma unidade reprodutiva

distinta das demais; a adaptação significa o ajustamento da população ao seu ambiente por força de seleção natural. Espera-se daí, que as diferenças morfométricas de populações geneticamente isoladas possam ser analisadas quantitativamente.

Por muito tempo, o termo *morfometria* foi utilizado para qualquer estudo que analisasse quantitativamente a variação da morfologia encontrada nos organismos. A palavra foi cunhada por Blackith (1965), para designar métodos que serviam para medir a distância de forma *entre* espécies e, a partir daí, construir fenogramas (diagramas em forma de árvore onde o comprimento de um ramo é proporcional ao grau de semelhança fenotípica das espécies).

Durante os anos 80, ocorreu uma revolução metodológica nesta área fazendo com que houvesse a necessidade de um embasamento teórico e filosófico, culminando com o surgimento de uma área de pesquisa na fronteira entre a Biologia, a Estatística e a Geometria (MONTEIRO e REIS, 1999). Fred L. Bookstein, um dos maiores responsáveis pela “revolução morfométrica”, define morfometria como “o estudo estatístico da covariância entre mudanças de formas e fatores causais” (BOOKSTEIN, 1991) o que implica na busca de causas das diferenças de *forma* entre organismos, sejam elas ecológicas ou filogenéticas. Incorporando-se conceitos geométricos na análise morfométrica, a definição de *forma* mais utilizada passou a ser: qualquer propriedade de uma configuração de pontos que não se altera por efeitos de tamanho, posição e orientação (BOOKSTEIN, 1989).

Em 1982, Strauss & Bookstein (MONTEIRO e REIS, 1999) sugeriram um protocolo geométrico preciso para a quantificação da forma de organismos, notadamente peixes, denominado “rede de treliças”, que possibilita a obtenção de medidas a partir da associação dos marcos anatômicos definidos com base na morfologia externa, pressupondo-se a homologia dos mesmos entre as formas comparadas.

Este capítulo está dividido em duas partes. A primeira tem por tema a Análise Multivariada de dados morfométricos, e a segunda, a Morfometria Geométrica.

2.1 ANÁLISE MULTIVARIADA DE DADOS MORFOMÉTRICOS

Análises multivariadas são técnicas usadas para estudar e descrever a covariação entre variáveis, entre indivíduos ou entre variáveis e indivíduos, objetivando: redução de dimensões ou de dados; a busca de um padrão estrutural para a massa de dados e testar a adaptação dos dados a modelos ou para discriminar populações. A força da Análise Multivariada reside na capacidade de examinar quantitativamente muitas variáveis simultaneamente [GAUCH,1989], com a finalidade de otimizar a interpretação de grandes conjuntos de dados. Ela analisa medições múltiplas efetuadas em uma ou diversas amostras de indivíduos, com as variáveis múltiplas consideradas em conjunto, tal como um sistema de medidas (FERREIRA e SOUZA,1997).

A maioria dos estudos morfométricos recentes faz uso de técnicas de análise estatística multivariada, que permitem considerar simultaneamente a variação de diversos caracteres quantitativos. Técnicas como Análise de Componentes Principais, Análise das Variáveis Canônicas e de Funções Discriminantes são extremamente úteis na ordenação dos dados morfométricos, pois permitem que parâmetros biológicos subjacentes às relações morfológicas entre indivíduos ou grupos possam ser mais facilmente detectados e interpretados (MONTEIRO e REIS, 1999).

2.2 MORFOMETRIA GEOMÉTRICA

Nos últimos vinte anos foram desenvolvidos métodos para descrever e comparar estatisticamente os diversos tipos de curvas em formas biológicas. Estes métodos foram os primeiros a capturar geometricamente a variação na forma dos organismos. Até recentemente, a discriminação de tais variações de forma entre organismos era baseada em parte sobre avaliações subjetivas. Ocorreram avanços que permitem identificar variações de forma entre organismos, e relacionar essas variações com o ambiente. A identificação de variações de forma se dá através de aproximações quantitativas que têm sido chamadas de “Morfometria Geométrica”. Essas técnicas

geométricas têm várias vantagens: são relativamente econômicas quando comparadas com a genética através do DNA, por exemplo; implicam em menor uso de técnicas laboratoriais e têm sido reconhecidas por serem extremamente sensíveis para indicar variações de forma quando são devidas ao meio que estão inseridas (CLARK, 1992).

As famílias de técnicas dividem-se em análises para contornos abertos e fechados (ROHLF, 1990), entretanto, um ponto comum a todas estas técnicas é que todas são funções que se ajustam à forma do contorno estudado, e produzem parâmetros que são utilizados como variáveis para o estudo da forma.

As características de marcos anatômicos em organismos que podem ser abstraídos, estudados e comparados não são dimensões lineares, mas as características baseadas nas suas coordenadas através do uso da morfometria geométrica para descrever diferenças em tamanho e forma passam a ser lineares. Esse método detecta diferenças de forma entre organismos por superposição de um sobre outro, expressando variação e diferenças estatisticamente e visualmente (BOOKSTEIN, 1991).

Na morfometria geométrica, o tamanho é definido em termos do centróide da configuração de marcos anatômicos. O centróide é o ponto médio dos marcos anatômicos, em outras palavras, é o centro de gravidade do conjunto de marcos anatômicos daquele objeto. O tamanho do centróide é a raiz quadrada da soma dos quadrados das distâncias entre cada marco anatômico e o centróide, ou equivalentemente, é a raiz quadrada da soma das variâncias dos marcos anatômicos em relação ao centróide. Pelo fato do tamanho do centróide ser a única variável de tamanho que não se correlaciona com a forma, quando o modelo nulo da alometria (ausência de correlação da forma com o tamanho) é verdadeiro, justifica-se a utilização do tamanho do centróide como variável de tamanho padrão em todas as análises geométricas, em detrimento de outras variáveis (BOOKSTEIN, 1991).

A proporcionalização é feita igualando-se o tamanho do centróide de cada configuração a 1, o que é obtido pela divisão do valor de cada componente do

centróide por sua norma. Para se eliminar o efeito indesejável da orientação, aplica-se à rotação, para que os marcos anatômicos correspondentes nas duas configurações fiquem o mais próximo possível. A rotação é uma transformação de similaridade que é aplicada para se comparar formas de figuras, removendo-se os efeitos da orientação nas coordenadas de pontos homólogos. Esses procedimentos estão descritos em detalhes em MONTEIRO e REIS (1999).

Quando as configurações de marcos anatômicos estão centradas em um lugar comum, proporcionalizadas de modo que todas tenham o mesmo tamanho e rotacionadas segundo um critério de otimização, as coordenadas resultantes destas transformações localizam um ponto no espaço de forma ("*shape space*"), chamado Espaço de Forma de Kendall (KENDALL, 1984; 1989).

A concentração dos dados no espaço de Kendall, permite que análises estatísticas comuns sejam utilizadas para estudar a variação da forma, mesmo dentro desse espaço não-linear. Uma comparação pode ser feita com um espaço curvo muito familiar: a superfície da Terra.

A métrica do espaço de Kendall é chamada *distância de Procrustes*, que é a mínima soma dos quadrados das distâncias entre pontos homólogos num par de configurações após a aplicação do processo de superposição. A métrica de Procrustes determina que o espaço de forma seja não-euclidiano. A distância de Procrustes equivale a uma distância em corda no espaço de forma. Distâncias geodésicas (ou Riemannianas) podem ser obtidas a partir de simples transformação da distância em corda (MONTEIRO e REIS, 1999). Em escala local na Terra, as distâncias em corda e geodésica têm valores muito próximos, dando a impressão de linearidade. Uma opção ao problema do espaço de forma é trabalhar com espaços que aproximam o espaço de forma tangente e que apresentam características lineares.

O espaço tangente é construído a partir de projeções dos indivíduos que se encontram no espaço de forma. Nenhuma informação é perdida na projeção do espaço de forma para o espaço tangente, pois eles apresentam sempre a mesma

dimensão $k.(p - k)$, para p marcos em k dimensões. Também por causa da concentração dos dados, as distâncias entre indivíduos no espaço tangente avizinham-se das distâncias geodésicas no espaço de forma (MONTEIRO e REIS, 1999).

A maioria dos conjuntos de dados estudados em biologia está concentrada no espaço de forma, de modo que eles se comportam como se fossem lineares (MONTEIRO e REIS, 1999).

As componentes uniformes ou afins podem ser descritas como as diferenças de forma após a configuração ter sido transladada, rotacionada e proporcionalizada para igualar-se a configuração média (CM), tal que todo conjunto de linhas paralelas sobre a configuração inicial permaneça paralelo após a transformação. Em outras palavras, elas expressam quanto um espécime é esticado ou contraído numa direção com respeito a outro, ou cisalhado até que assuma tamanho e local comum ao da CM. A variação de forma não uniforme consiste dos outros movimentos que são não lineares, e que estão localizados em regiões diferentes do corpo (CORTI *et al*, 1996). A primeira componente uniforme explica os estiramentos ao longo do eixo x da configuração, enquanto a segunda componente uniforme indica dilatações ou compressões ao longo do eixo y .

A função “*thin plate splines*” é utilizada em animação e computação gráfica para transformar suavemente uma figura em outra. Esta função modela o comportamento de uma placa de metal bastante fina, de proporções lineares infinitas, quando ela tem que ser deformada para se ajustar a uma configuração de marcos que esteja sobre ela, numa configuração abaixo da mesma (BOOKSTEIN, 1989b).

O processo de transformação da configuração tangente, de modo que todos os marcos de uma estejam superpostos aos marcos homólogos da configuração sob a placa, envolve uma deformação da placa de metal, com conseqüente gasto de energia para tal. A energia necessária para se deformar um grupo de marcos anatômicos depende da distância que estes se encontram uns dos outros. Quanto mais próximos os marcos, mais energia é necessária para a deformação da placa.

A matriz L^{-1} (chamada matriz de energia de deformação) é simétrica, e é decomposta para encontrar seus autovalores e autovetores, ou seja,

$$L_p^{-1} = E\Lambda E^T$$

onde Λ é a matriz diagonal ($p \times p$) de autovalores e E é uma matriz ($p \times p$) com os autovetores (linhas = marcos, colunas = autovetores normalizados) da matriz L_p^{-1} . Esses autovetores são as deformações principais (“principal warps”) que expressam os modos de variação possível nos espécimes (BOOKSTEIN, 1989) e cada um descreve uma possível mudança de forma aplicável à configuração de referência em diferentes escalas espaciais. Os maiores autovalores estão associados às deformações principais que descrevem mudanças de forma em pequena escala (muito localizadas), os menores autovalores estão associados às deformações principais que descrevem transformações em maior escala (MONTEIRO e REIS, 1999).

A matriz de energia de deformações não tem posto completo, três autovetores terão sempre autovalores nulos, pois correspondem à redução de dimensões imposta pelas restrições lineares que retiram os efeitos de tamanho, posição, orientação. Daí as deformações principais aparecerem em número de $(p-3)$ para as figuras bidimensionais.

A partir da matriz de energia de deformação pode-se construir um subespaço localizado, tangente ao espaço de forma, na vizinhança da configuração de referência. Daí, a escolha mais lógica da referência ser uma configuração média da amostra calculada pela superposição ortogonal por mínimos quadrados (MONTEIRO e REIS, 1999).

Para se transformar a configuração de referência em uma configuração alvo, a modificação da forma observada é dada pelas projeções das coordenadas da configuração alvo nas deformações principais. Estes escores indicam quanto de cada deformação principal é necessário para realizar a deformação desejada. Estes escores são chamados de pesos (“weights”) (ROHLF, 1993; 1996), porque ponderam as mudanças de forma em cada deformação principal. Enquanto que os vetores de escores são chamados *deformações parciais* (“partial warps”) (BOOKSTEIN, 1991), porque cada uma descreve uma porção da variação total.

A projeção das coordenadas é feita pela fórmula

$$W_i = Y_i E \Lambda^{-\alpha/2}$$

onde Y_i é uma matriz com as diferenças entre coordenadas x e y de cada espécime para a referência, E é a matriz de deformações principais e Λ é a matriz de autovalores que pondera os escores de acordo com o valor de α . Considerando-se $\alpha=0$, os escores são calculados com peso igual para todas as deformações principais, sendo a opção mais indicada para estudos exploratórios (MONTEIRO e REIS, 1999).

A análise das deformações relativas, método sugerido por Bookstein (1989), é uma análise de componentes principais da matriz W formada pelos escores das deformações parciais. A análise é realizada calculando-se uma matriz de variância-covariância desses escores e calculando-se os autovalores e autovetores (componentes principais). Estes componentes são as deformações relativas. Esse nome foi dado porque os autovalores relativos representam uma razão entre a variância dos escores e a energia de deformação (BOOKSTEIN, 1991).

Após a introdução do expoente α no cálculo dos escores das deformações parciais (ROHLF, 1993), passou a ser possível a combinação das deformações parciais com os componentes uniformes para o cálculo das deformações relativas. Isto, no entanto só é possível quando $\alpha=0$.

As deformações relativas foram utilizadas para estudos exploratórios buscando detectar evidências de heterogeneidade na forma da amostra. A análise das deformações relativas foi utilizada em vários estudos de Sistemática, como os exemplificados pelos trabalhos de Loy *et al.* (1993), em crânios de toupeiras, Rohlf (1993), com asas de mosquitos, Sarà (1996), com mandíbulas de musaranhos, Dam (1996), em dentes de roedores, Carpenter (1996), na forma do corpo de peixes imperadores, Monteiro (1997) e Monteiro *et al.* (1999) para estudar, respectivamente, a variação ontogenética da forma no crânio de largatos e de roedores (MONTEIRO e REIS, 1999), Cavalcanti M.J. *et al.* (1999) para análise morfométrica em espécies de peixes Serranidae (Perciformes: Teleostei).

3 MATERIAIS E MÉTODOS

Uma amostra de peixes *Selene spixii* (número de espécimes $n = 281$; comprimento zoológico CZ: 5,48 a 13,41 cm) e *S. vomer* ($n = 108$; CZ: 5,34 a 13,17 cm) (Figura 2) composta por espécimes procedentes de Caiçara do Norte (RN), pescados com rede de arrasto em junho e julho de 2001, foi objeto deste estudo. A obtenção do tamanho dos indivíduos de amostras de peixes explorados pela pesca faz parte da rotina de todos os laboratórios de investigação pesqueira. A medição do tamanho é a técnica mais utilizada em biologia pesqueira pois as demais características biológicas (peso, idade, maturação sexual) estão relacionadas com o tamanho. A escolha da dimensão a medir recai sempre sobre a mais fácil de medir e sujeita a menos erros (com menor variabilidade). A medida do comprimento zoológico (comprimento da extremidade do focinho (lábio superior) à furca) foi utilizada em virtude das espécies possuírem barbatana caudal furcal.

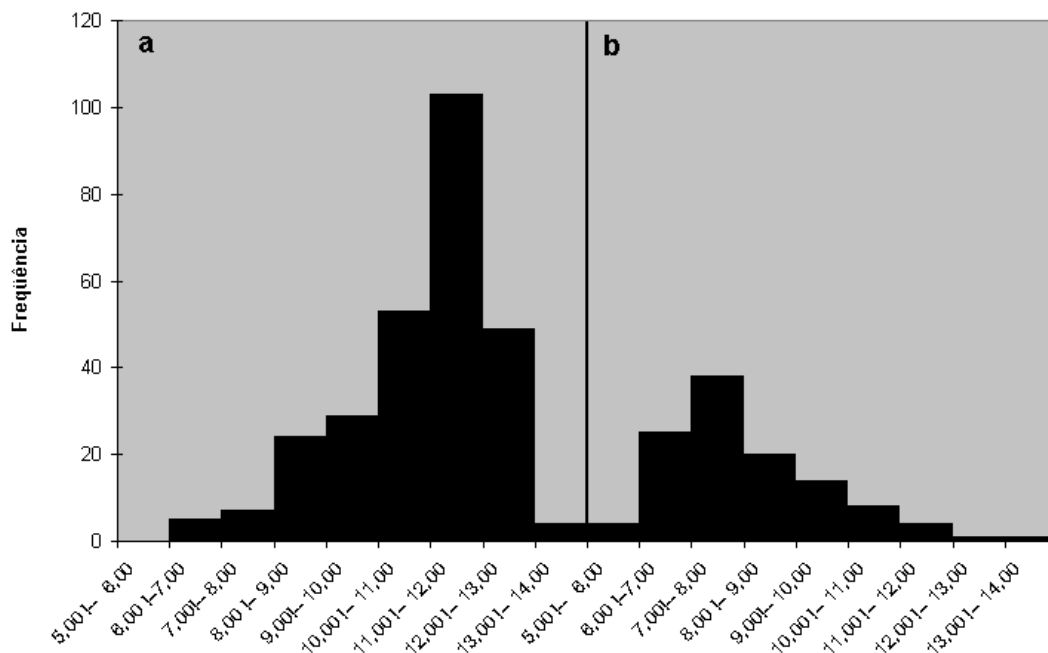


Figura 2: Distribuição de freqüências de comprimentos zoológicos de (a) *S. spixii* e (b) *S. vomer*.

A Figura 3 mostra 3 peixes de cada espécie em fases diferentes.



S. spixii – peixe 270



S. vomer– peixe 383



S. spixii – peixe 057



S. vomer– peixe 366



S. spixii – peixe 001



S. vomer– peixe 385

Figura 3: Imagens de peixes *S. spixii* e *S. vomer* (à direita) em 3 fases distintas.

3.1 MARCOS ANATÔMICOS

As medidas foram baseadas no protocolo da rede de treliças, relativas a 9 marcos anatômicos cujas posições estão exibidas na Figura 4, usando-se o lado ventral esquerdo de cada peixe.

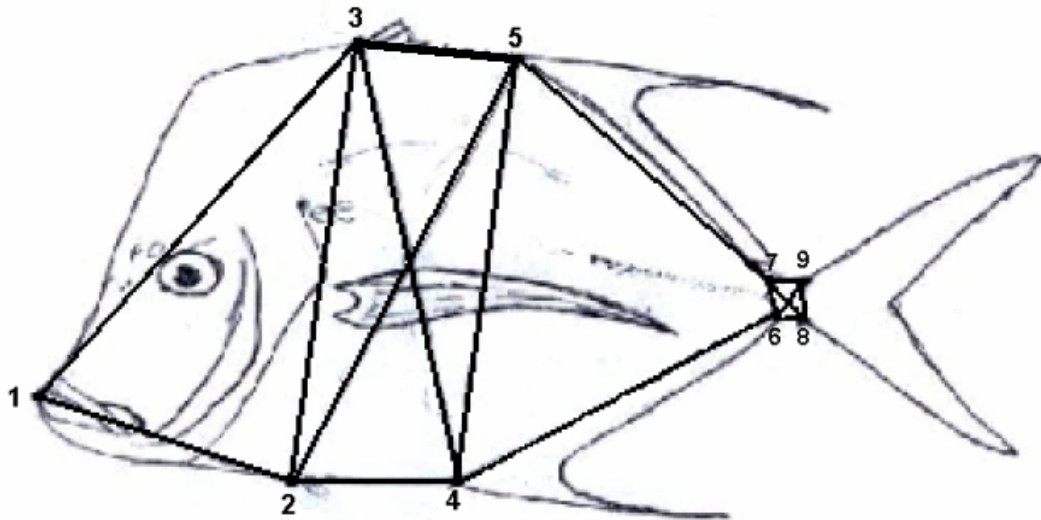


Figura 4: Rede de treliças para *Selene vomer* obtida a partir dos 9 marcos anatômicos indicados na figura.

Os dados morfométricos mostrados na Tabela 1 foram obtidos para as duas espécies a partir dos seguintes marcos anatômicos, homólogos em todos os espécimes da amostra:

- 1 - Ponta do focinho;
- 2 - Base da nadadeira ventral (pélvica);
- 3 - Início da primeira nadadeira dorsal
- 4 - Início da nadadeira anal;
- 5 - Início da segunda nadadeira dorsal;
- 6 - Final da nadadeira anal;
- 7 - Final da segunda nadadeira dorsal;
- 8 - Início da nadadeira caudal inferior;
- 9 - Início da nadadeira caudal superior.

Tabela 1: Variáveis morfológicas no estudo morfométrico de peixes *Selene spixii* e *Selene vomer*.

VARIÁVEL	NOTAÇÃO
Distância do focinho à base da Nadadeira ventral	D 1-2
Distância do focinho ao início da 1ª nadadeira dorsal	D 1-3
Distância da base da nadadeira ventral ao início da 1ª nadadeira dorsal	D 2-3
Distância da base da nadadeira ventral ao início da nadadeira anal	D 2-4
Distância da base da nadadeira ventral ao início da 2ª nadadeira dorsal	D 2-5
Distância do início da 1ª nadadeira dorsal ao início da nadadeira anal	D 3-4
Distância do início da 1ª nadadeira dorsal ao início da 2ª nadadeira dorsal	D 3-5
Distância do início da nadadeira anal ao início da 2ª nadadeira dorsal	D 4-5
Distância do início da nadadeira anal ao final da nadadeira anal	D 4-6
Distância do início da 2ª nadadeira dorsal ao final da 2ª nadadeira dorsal	D 5-7
Distância do final da nadadeira anal ao final da 2ª nadadeira dorsal	D 6-7
Distância do final da nadadeira anal ao início da nadadeira caudal inferior	D 6-8
Distância do final da nadadeira anal ao início da nadadeira caudal superior	D 6-9
Distância do final da 2ª nadadeira dorsal ao início da caudal inferior	D 7-8
Distância do final da 2ª nadadeira dorsal ao início da caudal superior	D 7-9
Distância do início da nadadeira caudal inferior ao início da superior	D 8-9

3.2 TRATAMENTO DE IMAGEM

Os peixes foram retirados do *freezer* e colocados numa solução de álcool a 10% durante 30 minutos e para ficarem enxutos foram expostos por 30 minutos ao ar livre. Em seguida, cada peixe teve sua imagem capturada por um *scanner* TCE S430. Os dados primários para este estudo foram constituídos por 389 imagens originalmente obtidas no formato JPEG (Joint Photographic Experts Group), que foram convertidas para o formato TIF (Tag Image File), para uso do *scion*, ajustando a largura para 250 *pixels*. A Figura 5 mostra a imagem número 4 no formato tif.



Figura 5: Imagem 4 no formato tif, ajustada para uma largura de 250 pixels.

Em seguida, utilizando o software *Scion Image* (SCION, 2002), as 389 imagens foram transformadas em 389 matrizes numéricas bidimensionais (2D). Para transformar cada imagem na matriz 2D correspondente, deve-se entrar no programa *Scion Image*, através do seu executável, clicar no menu *File* e depois em *Import...*, para ter acesso aos discos do computador. A seguir, localizar a pasta onde estão as imagens e dar um clique duplo no arquivo de interesse. A imagem é exibida em tons de cinza (que variam de zero (branco) a 255 (preto)) na janela principal do programa, como ilustrado na Figura 6.

Para obter a matriz contendo os valores dos tons de cinza da imagem visualizada, deve-se clicar em *FILE* → *Export...*, digitar um nome para a matriz, marcar a opção *Text* e clicar no botão *Salvar*. Como padrão, o arquivo de texto é salvo na mesma pasta em que estiver a imagem. A Tabela 2 mostra os valores de tons de cinza da região do olho do peixe 4 (Figura 6).



Figura 6: Janela principal do Scion Image, exibindo a imagem em tons de cinza do peixe 4.

Tabela 2: Matriz (parcial) correspondente à imagem 4 da Figura 6

0	0	10	3	88	78	22	32	62	100	116	112	101
1	2	7	43	34	52	62	45	18	85	72	87	29
0	3	13	19	5	0	10	17	70	112	69	123	86
0	0	0	0	71	121	136	144	114	24	33	88	84
21	0	58	150	160	172	177	168	171	168	170	115	65
69	142	136	175	175	187	197	195	207	215	180	188	164
139	133	180	169	161	200	171	183	189	165	200	188	182
130	189	171	182	196	200	197	194	196	196	171	177	178
166	169	179	179	184	170	175	124	93	153	183	157	171
175	170	176	177	172	156	61	86	89	23	134	165	165
150	123	123	167	159	182	192	169	171	191	169	144	143
130	147	143	155	202	203	189	195	168	186	179	173	153

Para uniformizar as dimensões de todas as matrizes 2D, foi desenvolvido um programa em C (SCHILDT, 1996). Este programa retornou 389 matrizes com 128 linhas por 240 colunas.

Com o Programa Fantomas (VIEIRA *et al.*, 2003), obteve-se a matriz 3D (389, 128, 240) e voltou-se para o *Scion Image* para visualizações e ajustes. Para acessar a matriz 3D no *Scion Image*, deve-se clicar em *File* → *Import...* e preencher o diálogo *Import File* como indicado na Figura 7. Clicando nos botões *OK* e *Abrir*, o arquivo com a matriz 3D é exibido na janela principal.

Usando as teclas > ou <, escolhe-se a imagem desejada, que pode ser editada e exportada como imagem ou como matriz.

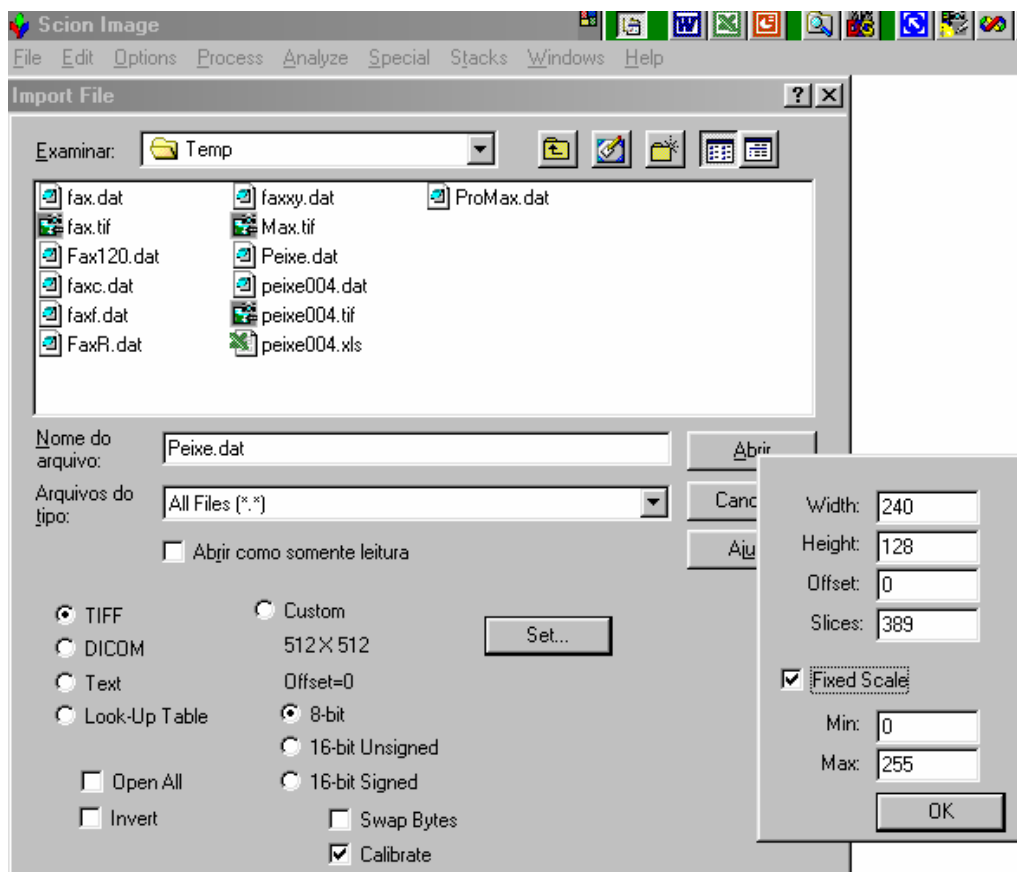


Figura 7: Diálogo *Import File* do *Scion Image* com os dados necessários para acessar a matriz 3D.

Para descompactar a matriz 3D, deve-se clicar no menu *Stacks* → *Stack to Windows*. A Figura 8 mostra a pilha de imagens resultantes deste procedimento.

A etapa seguinte consistiu fixar o tom de cinza de maior número identificador (ID) para 250. Para fazer isto no conjunto de matrizes, usou-se um programa em C.

Na seqüência, as matrizes foram compactadas no programa Fantomas e a matriz 3D resultante foi visualizada no *Scion Image*, onde foram marcados os pontos notáveis (marcos anatômicos). Cada ponto correspondeu a um *pixel* preto (ID = 255), marcado manualmente. A Figura 9 mostra os marcos colocados no peixe 4.

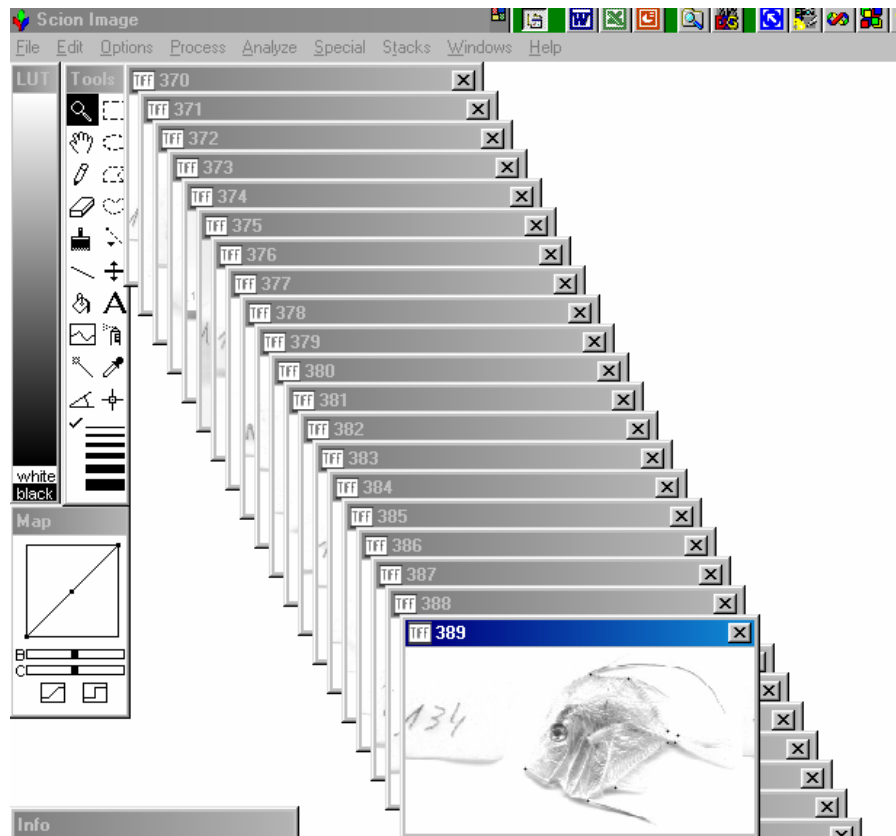


Figura 8: Matriz 3D descompactada no Scion Image



Figura 9: Marcos anatômicos (ampliados) colocados no peixe 4, usando o Scion Image

O novo conjunto de imagens/matrizes 2D, com os 9 marcos anatômicos, foi usado para a obtenção dos valores das variáveis desejadas.

3.2.1. Obtenção das Distâncias Entre os Marcos Anatômicos

As 16 distâncias que constituem as variáveis analisadas foram obtidas através de um programa em C, cujos passos são descritos a seguir.

PROGRAMA *Distâncias.c*

PASSO 1: Incluir bibliotecas padrões c.

PASSO 2: Definir constantes como as dimensões das matrizes de entrada e da matriz de saída.

PASSO 3: Declarar variáveis inteiras, reais e ponteiros para os arquivos de entrada e saída.

PASSO 4: Iniciar um laço para lê cada uma das 389 matrizes 2D com os tons de cinza e verifica se o valor do tom de cinza atual é 255. Se for, guarda os valores que indexam as linhas e colunas para o tom de cinza atual (i, j). Quando termina o processamento de uma matriz, o programa armazena as coordenadas dos 9 marcos anatômicos como a k-ésima linha da matriz de saída. Quando termina o laço, a matriz de saída contém 389 linhas, cada uma com as coordenadas dos 9 marcos anatômicos, em *pixel*.

PASSO 5: Iniciar um laço para calcular as 16 distâncias pretendidas, armazenando-as em uma nova matriz. Usando a matriz gerada no PASSO 4 e a fórmula

$d = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$, as distâncias são calculadas em *pixel*, convertidas para centímetro e guardadas em uma matriz 390×16 . No final do laço, esta matriz contém 390 linhas preenchidas. A 1ª linha, com a notação (coluna 2 da Tabela 1) que identifica cada distância e as demais com as 16 distâncias pré-selecionadas.

PASSO 6: Imprimir, em arquivo ASCII, a matriz 390×16 resultante do laço do PASSO 5.

PASSO 7: Fechar o programa.

Para converter *pixel* para centímetro (PASSO 5) foi usada a imagem 199 (Figura 10) que tem uma régua graduada em milímetros. Observou-se que 240 *pixels* equivalem a 16,8 cm, o que fornece como fator de conversão $0,07 = 16,8/240$, ou seja, cada medida em *pixel* multiplicada por 0,07 é convertida em centímetro.



Figura 10: Peixe nº 279, que corresponde à linha 199 da matriz 2D.

3.3 MÉTODO 1- ANÁLISE MULTIVARIADA DOS DADOS

3.3.1. Análise Estatística

A análise de variância multivariada (MANOVA), a análise de componentes principais, a análise das variáveis canônicas e a análise discriminante foram utilizadas na descrição multivariada dos dados. Os valores das 16 distâncias (Tabela 1) obtidas da rede de treliças foram transformados em logaritmos decimais para normalizar a distribuição e homogeneizar as variâncias ($d_{ij} = \log D_{i-j}$) e utilizados no cálculo da matriz de covariância entre as variáveis (SOKAL & ROHLF, 1981).

A MANOVA foi aplicada sobre a massa de dados oriundos de experimento inteiramente casualizado, segundo o Quadro da Tabela 3, com o objetivo de se verificar a existência de diferenças entre as duas espécies estudadas, em relação às variáveis, simultaneamente.

Tabela 3: Quadro de Análise da Variância Multivariada (MANOVA) - Delineamento Inteiramente Casualizado

FONTE DE VARIAÇÃO	GRAU DE LIBERDADE	MATRIZ DA SOMA DE QUADRADOS E PRODUTOS
Tratamentos	$I - 1$	$H = \sum_{i=1}^I n_i (\bar{x}_i - \bar{x})^T (\bar{x}_i - \bar{x})$
Resíduo	$\sum_{j=1}^J (n_j - i)$	$E = \sum_{i=1}^I \sum_{j=1}^J (\bar{x}_{ij} - \bar{x})^T (\bar{x}_{ij} - \bar{x})$
Total	$\sum_{j=1}^J (n_j - 1)$	$A = H + E$

A = matriz de somas de quadrados e produtos referente a totais;

I = número de tratamentos;

J = número de repetições e

n_j = número de observações na j-ésima repetição.

Para este estudo considerou-se os dados organizados segundo a matriz de planejamento mostrada na Tabela 4.

Tabela 4: Matriz de Planejamento

Tratamentos	Repetições	Variáveis d12 d13 ... d89	Matrizes
Espécie <i>S. spixii</i>			
1	Espécime 1		$H = \sum (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$
1	Espécime 2		
1	Espécime 3		
.....		
1	Espécime 281		
Espécie <i>S. vomer</i>			
2	Espécime 1		$E = \sum \sum (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T$
2	Espécime 2		
2	Espécime 3		
.....		
2	Espécime 108		
			$A = H + E$

A análise de componentes principais e a análise das variáveis canônicas foram utilizadas na descrição multivariada dos dados morfométricos. As distâncias foram utilizadas no cálculo da matriz de covariância entre as variáveis. Desta matriz foram extraídos os autovalores e seus autovetores associados que, combinados às variáveis originais, produziram as combinações lineares que definem os componentes principais. Os padrões de discriminação entre as espécies foram examinados por meio da análise de variáveis canônicas e análise discriminante. Maiores detalhes são encontrados no Apêndice 1.

Coeficientes alométricos multivariados para cada variável foram obtidos segundo JOLICOEUR (apud CAVALCANTI & LOPES, 1993, p. 56), calculando-se a média do primeiro componente principal e dividindo-se cada valor pelo resultado, podendo ser interpretados como coeficientes de crescimento relativo. Os valores dos coeficientes alométricos maiores do que 1 indicam alometria positiva em relação ao tamanho, enquanto que os valores menores que 1 indicam alometria negativa. As variáveis morfométricas alometricamente positivas crescem relativamente mais rápido que a média, contribuindo mais intensivamente para o fator do tamanho geral representado pelo 1º componente principal.

As análises estatísticas foram realizadas em um microcomputador AMD-DURON - 1400, com o pacote estatístico SAS, versão 6.12 - *Institute SAS* (1999).

3.4 MÉTODO 2 – ANÁLISE DE DEFORMAÇÕES.

3.4.1. Análise de Deformações Relativas (Função *Thin Plate Splines*)

Para a análise de variações de forma entre indivíduos reportada aqui, foram usadas técnicas morfométricas desenvolvidas por Bookstein (1989, 1991) e Rohlf (1990). A análise de deformações relativas ("*thin plate spline*") não é amplamente conhecida, assim, as análises técnicas estão resumidas aqui e em mais detalhes no Apêndice 2.

As coordenadas dos marcos anatômicos de todos espécimes foram alinhadas usando-se o método de superposição ortogonal mínimos quadrados Procrustes generalizado, que superpõe uma configuração sobre outra, minimizando a soma de

quadrados das distâncias entre os marcos anatômicos homólogos. A configuração média de marcos resultante desse processo serviu como a referência ou como a configuração tangente (definindo o ponto de tangência entre o espaço de forma não linear e o espaço tangente).

Para cada espécime foram calculados o tamanho do centróide e as componentes uniformes x e y . Neste estudo, o eixo- x corresponde ao eixo ântero-posterior, e o eixo- y corresponde ao eixo dorso-ventral do corpo do peixe.

As componentes uniformes foram estimadas pelo método linearizado Procrustes de Bookstein (1996b). Análises de correlação foram realizadas entre tamanho do centróide e as componentes uniformes. O tamanho do centróide foi testado para identificar diferenças entre espécies por análise de variância (ANOVA). As componentes uniformes foram testadas para identificar diferenças entre espécies por análise de variância multivariada (MANOVA).

As coordenadas de todos os espécimes alinhados foram usadas para análise de deformação relativa (*thin plate splines*), para analisar e mostrar a direção de diferenças de forma entre espécies. A matriz de energia de deformação resultante deste processo é então decomposta em eixos ortogonais, que são as deformações principais.

A projeção dos espécimes superpostos através das deformações principais produz os escores (pesos) deformações parciais, que descrevem seus desvios da CM e que são usadas como variáveis em análise estatística multivariada. A configuração média de marcos anatômicos foi usada como configuração de referência na análise de deformações relativas, e a referência foi alinhada para seus eixos principais. As deformações relativas foram calculadas com a opção $\alpha=0$ e no espaço de forma completo (incluindo as componentes uniformes e não uniformes) como recomendado por Bookstein (1996a) para estudos exploratórios. Foram usadas deformações de grades, pela função *thin plate splines* para retratar graficamente os padrões de variação de forma entre marcos anatômicos.

Padrões de variação entre espécies no espaço de forma total foram examinados por análise de variáveis canônicas da matriz W de escores de deformações parciais com as componentes uniformes incluídas.

Para digitalizar os marcos anatômicos em cada imagem foi utilizado o programa TPSDIG 32, versão 1.40, 2003, de F. James Rohlf. A análise de deformações relativas e cálculos dos escores de deformações parciais foram feitos usando o software TPSRelw 32, 2003, de F. James Rohlf. As análises estatísticas foram realizadas com o pacote estatístico SAS, versão 6.12 - *Institute SAS* (1999). Todos os cálculos foram realizados em um microcomputador AMD-DURON -1400.

4 RESULTADOS E DISCUSSÃO

4.1 MÉTODO 1

4.1.1. Análise de Variância Multivariada (MANOVA)

A MANOVA foi utilizada para testar a igualdade dos vetores médios das medidas das variáveis. O teste de significância utilizado foi o teste de Wilks (Λ), que é um teste realizado entre o resíduo e o total. Quando existe diferença sistemática obtém-se $\Lambda < 1$, quanto mais significativa a diferença, menor o seu valor, desta forma a hipótese de nulidade é rejeitada.

Pelo Teste de Wilks ($\Lambda=0,136$) observou-se a existência de leve diferença entre vetores médios das 16 distâncias referentes às espécies *S. spixii* e *S. vomer* de acordo com os resultados apresentados na Tabela 5.

Tabela 5: Análise de variância multivariada para as espécies *S. spixii* e *S. vomer*

Λ Wilks	F	gl num	gl den	Pr>F
0,13596	147,76	16	372	<0,0001

Onde: gl num: graus de liberdade do numerador

gl den: graus de liberdade do denominador

Pr : probabilidade real

4.1.2. Análise De Componentes Principais

O resumo da Análise de Componentes Principais está inserido na Tabela 6. Os três primeiros componentes (Y_1 , Y_2 e Y_3) explicaram 82,88% da variação total, sendo que o primeiro componente explicou 65,66%, o segundo 11,92% e o terceiro 5,29% da variação total.

Os demais componentes absorveram menos de 18% do total da variância, isso mostra que os três primeiros componentes principais podem ser utilizados para as análises, uma vez que explicaram satisfatoriamente a variabilidade do conjunto de dados.

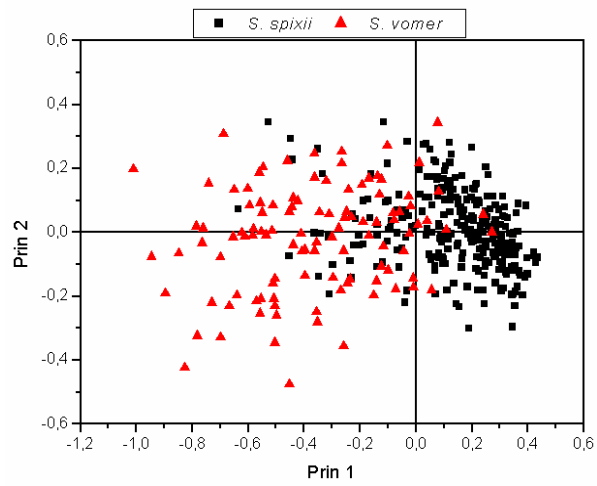
Tabela 6: Resumo da Análise de Componentes Principais

Componentes Principais	Autovalores	Diferenças	Proporção de Variância Total	Proporção Acumulada da Variação Total
Y ₁	0,09862906	0,08072313	0,6566	0,6566
Y ₂	0,01790593	0,00995511	0,1192	0,7759
Y ₃	0,00795082	0,00118767	0,0529	0,8288
Y ₄	0,00676315	0,00144154	0,0450	0,8738
Y ₅	0,00532161	0,00159558	0,0354	0,9092
Y ₆	0,00372603	0,00062961	0,0248	0,9341
Y ₇	0,00309642	0,00028132	0,0206	0,9547
Y ₈	0,00281510	0,00123232	0,0187	0,9734
Y ₉	0,00158279	0,00056710	0,0105	0,9839
Y ₁₀	0,00101569	0,00031725	0,0068	0,9907
Y ₁₁	0,00069844	0,00030706	0,0046	0,9954
Y ₁₂	0,00039138	0,00019258	0,0026	0,9980
Y ₁₃	0,00019880	0,00012759	0,0013	0,9993
Y ₁₄	0,00007120	0,00003816	0,0005	0,9998
Y ₁₅	0,00003305	0,00003019	0,0002	1,0000
Y ₁₆	0,00000286	–	0,0000	1,0000

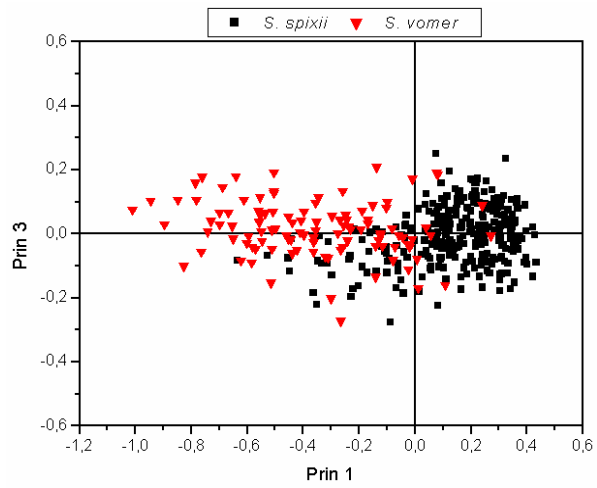
Os coeficientes do primeiro autovetor apresentaram valores positivos (Tabela 7), indicando que todas as variáveis são correlacionadas positivamente com este componente, sugerindo que este autovetor é um componente caracterizador do tamanho. O segundo e o terceiro autovetores, apresentaram coeficientes positivos e negativos, expressando informação sobre a forma nestas espécies.

Tabela 7: Coeficientes de ponderação (correlação das variáveis estudadas com os 3 primeiros componentes principais) e o coeficiente de alometria de cada variável

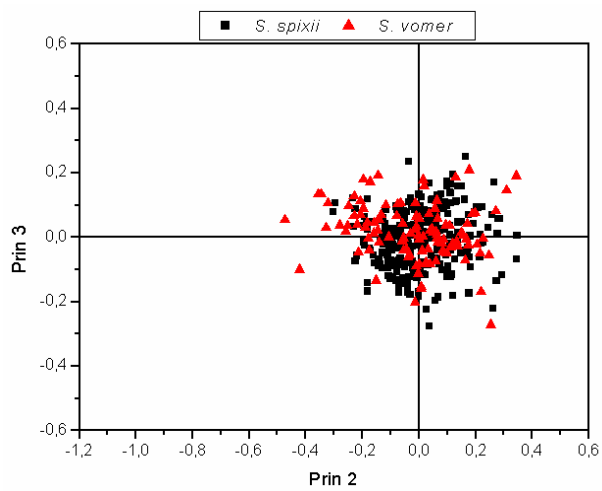
Variável	Y ₁ (Corr)	Y ₂ (Corr)	Y ₃ (Corr)	k
d12	0,2436 (0,8880)	0,1637 (0,2543)	0,1151 (0,1192)	0,99
d13	0,1559 (0,7542)	0,1810 (0,3731)	-0,0055 (-0,0075)	0,64
d23	0,1818 (0,8568)	0,2078 (0,4172)	-0,0769 (-0,1029)	0,74
d24	0,2535 (0,7211)	0,1184 (0,1435)	0,2270 (0,1834)	1,04
d25	0,1832 (0,8742)	0,2006 (0,4078)	-0,0641 (-0,0868)	0,75
d34	0,2250 (0,8926)	0,2326 (0,3933)	-0,0706 (-0,0796)	0,92
d35	0,1696 (0,6528)	0,1230 (0,2017)	0,2058 (0,2250)	0,69
d45	0,2247 (0,8804)	0,2386 (0,3983)	-0,0986 (-0,1097)	0,92
d46	0,2951 (0,8886)	0,2768 (0,3552)	-0,0140 (-0,0119)	1,20
d57	0,3146 (0,9162)	0,2572 (0,3192)	0,0512 (0,0424)	1,28
d67	0,2391 (0,7442)	-0,0095 (-0,0126)	-0,1235 (-0,1092)	0,98
d68	0,2865 (0,7179)	-0,4017 (-0,4288)	0,7171 (0,5102)	1,17
d69	0,2819 (0,8242)	-0,3709 (-0,4621)	-0,1151 (-0,0956)	1,15
d78	0,2752 (0,8516)	-0,1545 (-0,2037)	0,0746 (0,0655)	1,12
d79	0,3233 (0,7732)	-0,4713 (-0,4803)	-0,3650 (-0,2479)	1,32
d89	0,2657 (0,8036)	-0,1705 (-0,2197)	-0,4285 (-0,3680)	1,08



(a)



(b)



(c)

Figura 11: Gráficos dos primeiros componentes Principais

A partir da projeção (Figura 11(a)) dos indivíduos em relação aos primeiro e segundo componentes principais (Y_1 e Y_2) nota-se que há definição visual de dois grupos, havendo má classificação para alguns indivíduos o mesmo ocorre em relação aos primeiro e terceiro (Y_1 e Y_3), (Figura 11(b)), enquanto que para os componentes principais Y_2 e Y_3 (Figura 11(c)) nota-se que não houve definição visual de grupos.

Na projeção dos escores de *S. spixii* e *S. vomer*, nota-se uma dependência das projeções dos escores do 2º componente em relação ao 1º (Figura 7(a)), a qual é devido provavelmente à alometria.

Dessa forma, um aumento no tamanho geral tende a acarretar uma alteração na forma destas espécies durante seu crescimento. Observa-se que os espécimes maiores de *S. vomer* superpõem-se aos menores de *S. spixii* no espaço dos dois primeiros componentes principais.

4.1.3. Análise de Variáveis Canônicas

A Análise de Variáveis Canônicas forneceu uma única variável canônica que explica 100% da variação total disponível (Tabela 8). A partir da projeção dos indivíduos em relação a este eixo canônico (Figura 12) nota-se que há definição visual dos dois grupos *S. spixii* e *S. vomer*.

Tabela 8: Resumo da Análise de Variáveis Canônicas

Variáveis Canônicas	Autovalores	Diferenças	Proporção de Variância Total	Proporção Acumulada da Variação Total
Vcan ₁	6,3553	-	1,00	1,00

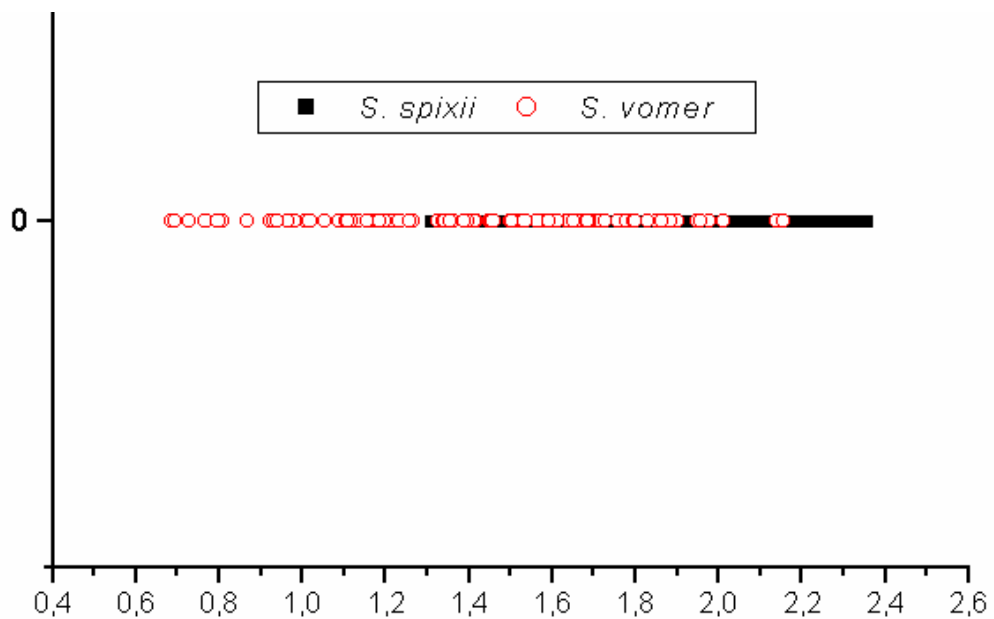


Figura 12: Projeção dos escores canônicos

As variáveis d13 e d34 foram as mais importantes para a discriminação das espécies, pois contribuíram com os maiores valores absolutos dos coeficientes canônicos padronizados (Tabela 9), que correspondem às distâncias D1-3 e D3-4 (da ponta do focinho ao início da primeira nadadeira dorsal e desse último ponto ao início da nadadeira anal), tomadas diagonalmente na direção dorso-ventral.

Tabela 9: Coeficientes padronizados das 16 variáveis, das amostras combinadas de *S. spixii* e *S. vomer* no eixo canônico

Variável	Can1	Variável	Can1
d12	0,1548	d46	1,1227
d13	-2,0928	d57	-0,6353
d23	-0,9734	d67	-0,0344
d24	0,4317	d68	0,1124
d25	0,1269	d69	0,2287
d34	1,9249	d78	-0,1363
d35	-0,5193	d79	0,1614
d45	0,6993	d89	0,1883

4.1.4. Análise Discriminante

Com a análise discriminante verificou-se a separação bem definida dos dois grupos (Tabela 10) com a seguinte função discriminante de Fischer:

$$\begin{aligned} \text{FDF} = & 0.253478933*d_{12} - 2.763827920*d_{13} - 3.097783229*d_{23} + 0.717573984*d_{24} + \\ & 2.710304506*d_{25} + 4.030502413*d_{34} - 2.060012805*d_{35} - 0.738376211*d_{45} + \\ & 1.560720280*d_{46} - 1.106410790*d_{57} - 1.253112164*d_{67} - 0.194166586*d_{68} + \\ & 2.237788482*d_{69} - 0.059187738*d_{78} + 0.255544928*d_{79} + 1.320731789*d_{89} \end{aligned}$$

A distância de Mahalanobis (D^2) entre os dois grupos foi de 31,526 ($F=147,76$; $P<0,0001$). Quanto maior for a distância de Mahalanobis entre dois grupos melhor, pois isto indica que os grupos são bastante distintos e portanto a classificação de um novo indivíduo num dos grupos se fará com a máxima probabilidade de acerto.

Tabela 10: Número de Observações e Percentual Classificado

Designação a priori	Função Discriminante		Total	Classificação Errada
	Selene spixii	Selene vomer		
Selene spixii	281	0	281	0%
Selene vomer	0	108	108	0%
Total	281	108	389	-

4.2 MÉTODO 2

Os marcos anatômicos da configuração média (CM) (a média das coordenadas dos marcos anatômicos dos 389 peixes) estão exibidos na Figura 13 e as coordenadas na Tabela 11.

Tabela 11: Coordenadas da configuração média

Marco anatômico	Abcissa	Ordenada
1	-0,50642	-0,10264
2	-0,22637	-0,28465
3	-0,20239	0,31820
4	-0,08238	-0,26476
5	-0,03140	0,30881
6	0,21645	-0,01544
7	0,23077	0,03884
8	0,28786	-0,03117
9	0,31389	0,03280

Os vetores que aparecem na configuração geral dos 389 peixes da Figura 14 mostram a direção e o sentido da deformação ocorrida em cada ponto correspondente ao respectivo ponto da CM.



Figura 13: Configuração média (ou consenso) dos 389 peixes da amostra estudada

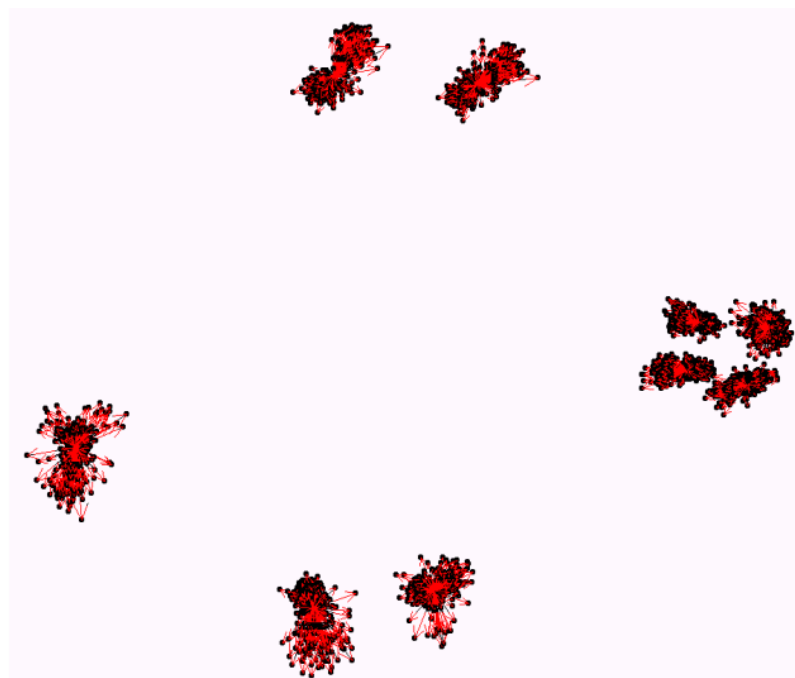


Figura 14: Configuração dos 389 peixes com vetores indicando direção e sentido de cada deformação em cada marco anatômico em relação à CM

Na análise das deformações relativas foram obtidos os autovalores e a proporção de variância total explicada pelos eixos associados (Tabela 12).

Tabela 12: Resumo da análise das deformações relativas

Eixos de deformações relativas	Autovalores	Proporção de Variância Total	Proporção Acumulada da Variação Total
X1	1,16503	0,6353	0,6353
X2	0,52151	0,1273	0,7626
X3	0,37917	0,0673	0,8299
X4	0,31828	0,0474	0,8773
X5	0,24917	0,0291	0,9063
X6	0,22553	0,0238	0,9302
X7	0,19881	0,0185	0,9487
X8	0,19330	0,0175	0,9661
X9	0,15147	0,0107	0,9769
X10	0,12440	0,0072	0,9841
X11	0,10733	0,0054	0,9895
X12	0,09378	0,0041	0,9936
X13	0,08445	0,0033	0,9970
X14	0,08050	0,0030	1,0000

Na análise de variância do tamanho do centróide ($F=216,23$; $p<0,0001$), observou-se que houve diferença significativa entre eles nas duas espécies.

A análise de variância multivariada das variáveis tamanho do centróide e componentes uniformes x e y, nas duas espécies estudadas, obteve-se ($Wilks \Lambda=0,291$; $F=312,37$; $p<0,0001$), o que demonstra existir leve diferença entre os vetores de média dos dois tratamentos. A distância de Mahalanobis (D^2) entre os dois tratamentos foi 12,074. A análise de variáveis canônicas apresentou um único autovalor de 2,4341 ($F=312,37$; $p<0,0001$) explicando toda a variação.

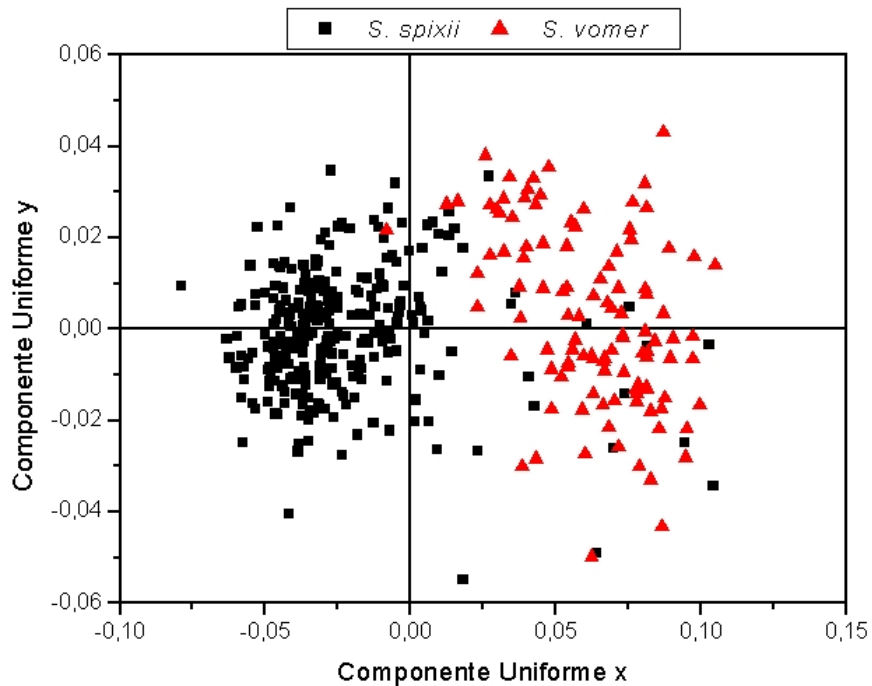
Tabela 13: Matriz de correlação entre as variáveis tamanho do centróide e componentes uniformes

	Tamanho do centróide	Componente uniforme x	Componente uniforme y
Tamanho do centróide	1,000		
Componente uniforme x	-0,7130	1,0000	
Componente uniforme y	0,1453	-0,1531	1,0000

Tabela 14: Coeficientes padronizados nas amostras combinadas de *S. spixii* e *S. vomer* no único eixo canônico

Variável	Can 1
Tamanho do centróide	0,3753
Componente uniforme x	1,2395
Componente uniforme y	0,2021

A variável mais importante para a discriminação das espécies, contribuindo com o maior valor dos coeficientes canônicos padronizados (Tabela 14) é Componente uniforme x. No gráfico da Figura 15 verificou-se que as variações de forma registradas revelaram a existência de duas espécies, e que a Componente uniforme x foi a que mais contribuiu para isto.

**Figura 15: Componentes uniformes x e y**

Houve uma correlação negativa de -0,7130 ; $p < 0,0001$, da componente uniforme x com o Tamanho do centróide, ao passo que houve uma insignificante correlação positiva de 0,1453 ; $p = 0,0041$, da segunda componente uniforme com o tamanho do centróide.

Tabela 15: Coeficientes de ponderação na análise de deformações relativas para os três primeiros eixos principais

Variáveis de mudança de forma	Eixo 1	Eixo 2	Eixo 3
x1	-0,033376	-0,040577	-0,138683
y1	-0,006951	0,014867	0,034557
x2	0,012102	0,097242	0,353126
y2	-0,093825	-0,039813	-0,068540
x3	0,178659	0,164562	0,043496
y3	0,150885	0,115139	0,221057
x4	0,092025	-0,029015	-0,034320
y4	-0,161600	-0,096300	0,316456
x5	0,098238	-0,047581	-0,257897
y5	-0,167257	0,059028	-0,108420
x6	-0,439381	0,142785	-0,486959
y6	-0,302862	0,871399	0,192056
Comp. Unif. x	0,760318	0,364213	-0,206206
Comp. Unif. Y	0,011383	-0,136066	0,550723

Na análise de variância multivariada dos escores de deformações parciais com a inclusão das componentes uniformes x e y, matriz W, nas duas espécies estudadas, obteve-se (Wilks $\Lambda = 0,1287$; $F = 180,91$; $p < 0,0001$) e a distância de Mahalanobis (D^2) entre os dois tratamentos foi de 33,59. A análise de variáveis canônicas apresentou um único autovalor de 6,7721 ($F = 180,91$; $p < 0,0001$) explicando toda a variação.

Na Tabela 15 estão os coeficientes de ponderação das 14 variáveis de mudança de forma da análise de deformações relativas nos 3 primeiros eixos principais, e nela observou-se que a componente uniforme x (Comp.unif. x) foi a que mais contribuiu na discriminação.

A Figura 16 mostra a ordenação de todos os indivíduos da amostra, através das deformações relativas, onde se verifica que a predominância dos pontos à esquerda do eixo vertical é a daqueles associados a números que variam de 1 a 281 (*S. spixii*), e à direita a predominância é dos números que variam entre 282 e 389 (*S. vomer*).

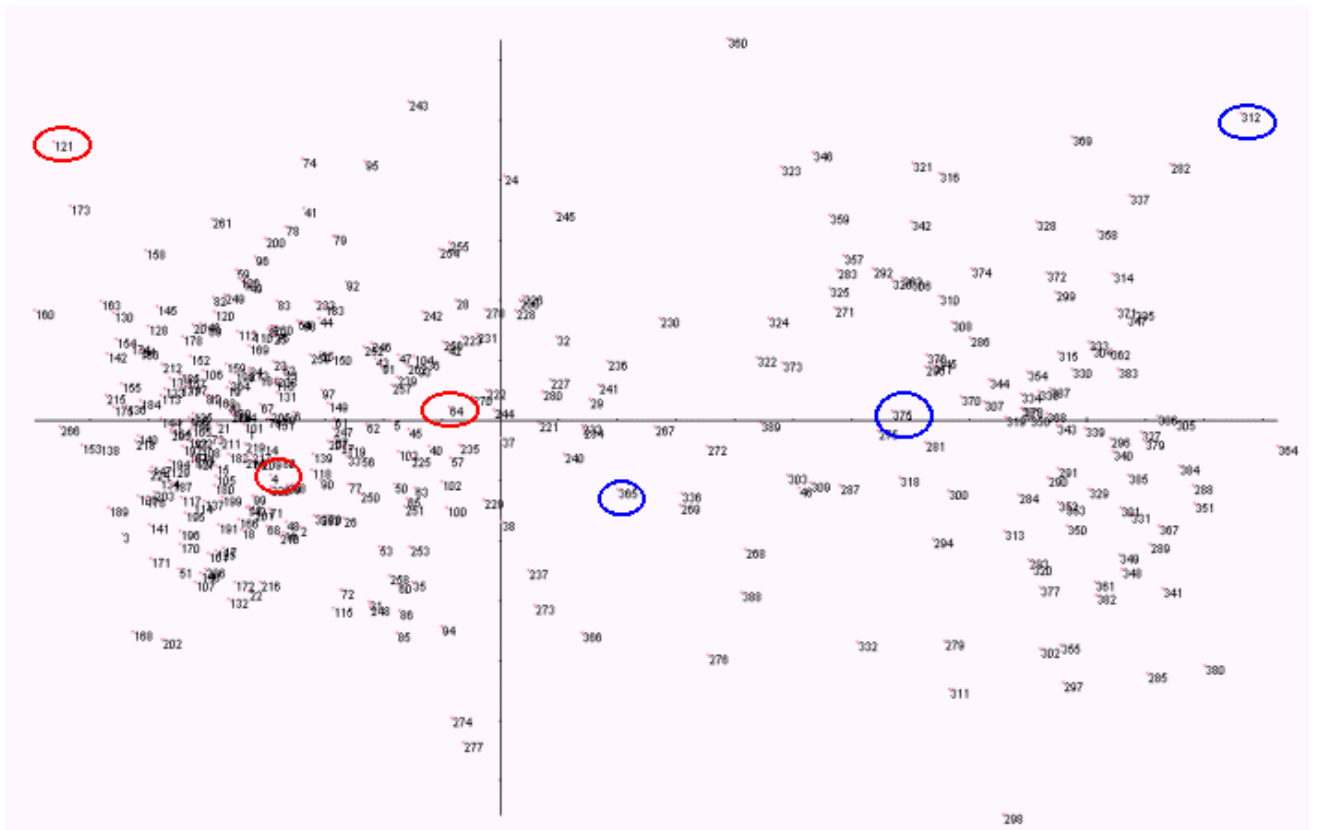


Figura 16: Ordenação dos espécimes por deformações relativas. No ponto origem situa-se a CM, e os pontos evidenciados em vermelho estão associados a *S. spixii*, enquanto os evidenciados em azul estão associados *S. vomer*

Na Figura 17 aparecem 9 grades, que representam mudanças globais (uniformes e não uniformes) de forma em relação à forma da CM (grades do centro). As grades da esquerda, de cima para baixo, mostram deformações ocorridas nos espécimes correspondentes aos pontos evidenciados em vermelho na ordem crescente da distância à CM. As grades da direita se relacionam aos pontos evidenciados em azul com o mesmo critério.

Observou-se que as grades associadas às configurações de *S. spixii* sofrem deformações horizontais para direita, que tornam-se mais fortes à medida que se afastam da CM. Enquanto que as grades associadas às configurações de *S. vomer* sofrem deformações horizontais para esquerda, que tornam-se mais fortes à medida que se afastam da CM. As deformações verticais são de pouca magnitude para as duas espécies.

5 CONCLUSÕES

É possível concluir, a partir dos resultados obtidos através do Método 1, que *S. spixii* e *S. vomer* podem ser discriminadas por medições orientadas diagonalmente em relação ao eixo ântero-posterior do corpo. As medidas D 1-3 e D 3-4 (distância da ponta do focinho ao início da primeira nadadeira dorsal e desse último ponto ao início da nadadeira anal) foram as mais importantes para a discriminação dessas espécies, pois contribuíram com os maiores valores absolutos dos coeficientes canônicos padronizados (Tabela 9).

A partir do Método 2, foi possível identificar a melhor variável para discriminar as espécies. Dentre as variáveis tamanho do centróide, componente uniforme x e componente uniforme y, a mais importante para a discriminação das espécies foi a componente uniforme x, pois contribuiu com o maior valor dos coeficientes canônicos padronizados (Tabela 14). Também verificou-se (Figura 15) que as variações de forma registradas pelos valores das componentes uniformes, discriminaram as espécies, e a componente uniforme x foi a que mais contribuiu.

A análise de deformações relativas através da matriz de escores das deformações parciais com as componentes uniformes, mostrou que houve diferença significativa entre as espécies. Nesta análise a variável que mais contribuiu na discriminação das espécies foi também a componente uniforme x que possui o maior coeficiente de ponderação do Eixo 1 (Tabela 15). A ordenação dos espécimes por deformações relativas (Figura 16) mostra a separação dessas espécies. Com a função “*thin plate spline*” foram geradas grades de deformação que possibilitam avaliações minuciosas de mudanças de forma, que favorecem estudos da Sistemática dessas espécies.

Portanto, os resultados das análises do Método 1 e do Método 2 convergem para a mesma conclusão, qual seja, é possível fazer a discriminação entre as espécies de peixes enfocadas neste trabalho, desde que utilizada a metodologia apropriada e a variável que melhor contribua para a discriminação.

REFERÊNCIAS

- BARTLEBY, www.bartleby.com/ – **Dictionary _ Appendix I _ Indo-European Roots.** página acessada em novembro-2003.
- BOOKSTEIN, F.L., **Morphometric Tools for Landmark Data.** Geometry and Biology, Cambridge University Press: New York, 1991.
- BOOKSTEIN, F.L., **Size and shape: a comment on semantics.** Syst. Zool. 38; 173 – 180.1989a.
- BOOKSTEIN, F.L., **Principal warps: thin plate splines and the decomposition of deformations.** IEEE Trans. Pat. Anal. Mach. Intel. 11: 567 – 585.1989b.
- BOOKSTEIN, F.L., **Combining the tools of geometric morphometrics,** p. 131 – 151. Advances in Morphometrics. Life Sciences. Vol 284. Plenum Publishing Corp., New York. 1996a.
- BOOKSTEIN, F.L., **Standard formula for the uniform shape component in landmark data,** p. 153 – 168. Advances in Morphometrics. Life Sciences. Vol 284. Plenum Publishing Corp., New York. 1996b.
- CARPENTER, K. E., **Morphometric pattern and feeding mode in emperor fishes (Lethrinidae, Perciformes), Advances in Morphometrics.** Life Sciences, Vol.284. Plenum Publishing Corp., New York. p. 479 – 488. 1996.
- CAVALCANTI, M. J. & LOPES, P. R. D., **Análise Morfométrica Multivariada de cinco espécies de serranidae (teleostei, perciformes).** P.53-64. Acta Biologica Leopoldensia. Vol.15. Nº 1. 1993
- CLARK, G. M., **Fluctuating asymmetry: a technique for measuring developmental stress of genetic and environmental origin.** Acta Zoologica Fennica 191, 31 – 36.1992.
- CORTI M. & CROSETTI., **Geographic variation in the grey mullet: a geometric morphometric analysis using partial warp scores.** Journal of Fish Biology 255-269 – 48. 1996.

CRUZ, C. D.; REGAZZI, A. J., **Modelos Biométricos Aplicados ao Melhoramento Genético**. Viçosa: U.F.V., 390p.1994

DIAS, L. A. DOS S.; KAGEYAMA, P. Y.; CASTRO, G. C. T., **Divergência Fenética Multivariada na Preservação de Germoplasma de Cacau (*Theoproma cacao* L.)** Agrotrópia. Centro de Pesquisa de Cacau, Ilheus. Bahia. 29 – 40p, 1997.

DOUGLAS M. E., W. J. MATTHEWS., **Does morphology predict ecology? Hypothesis testing within a freshwater stream fish assemblage**. *Oikos* 65; 213 – 224. 1992.

DRYDEN, I. L. & K. V. MARDIA., **Multivariate shape analysis**. *Sankhya* 55; 460 - 480.1993.

DRYDEN, I. L. & K. V. MARDIA., **Statistical shape analysis**. Wiley, New York.1998.

FERREIRA, R. L. C.; SOUZA, A. L., **Técnicas de Análise Multivariada Aplicadas ao Manejo Florestal no Brasil** – Universidade Federal Viçosa **Sociedade de Investigações Florestais** – MG, Viçosa: Boletim Técnico, 14. 1 – 16p.1997.

FISHBASE, www.fishbase.org, página acessada em novembro-2003.

FONTELES F. A. A., **Recursos Pesqueiros: Biologia e Dinâmica Populacional**. Fortaleza, Imprensa Oficial do Ceará, 296p, 1989.

GAUCH, H. G. J., **Multivariate analysis in community ecology**. Cambridge University Press, 198p. 1989.

GOWER, J. C., **Statistical methods of comparing different multivariate analysis of the same data, Mathematics in the Archaeological and Historical Sciences**. Edimburgh University Press, Edimburgh, p. 138 –149. 1970

JOHNSON R. A. and WICHERN D. W., **Applied multivariate statistical analysis**, 2ª ed., Englewood Cliffs: Prentice – Hall, 607p. 1992.

KENDALL, D. G., **Shape Manifolds, procrustean Metrics, and Complex Projective Spaces**. *Bull. Lond. Math. Soc.* 16:81 – 121. 1984.

KENDALL, D. G., **A Survey of the Statistical Theory of Shape**. Statistical Science 4(2): 87 – 120. 1989.

KLECKA, W., **Discriminant analysis**. Beverly Hills: Sage Publications, 71p. 1980.
4(2): 87 – 120. 1989.

MANLY B. R. J., **Multivariate Statistical Methods: a primer**, 2^a ed., London: Chapman and Haal, 215p, 1995.

MARDIA K. V., KENT J. T. and BIBBY J. M., **Multivariate Analysis**, London, Academic Press, 521p., 1997.

MONTEIRO L. R. e REIS S. F. dos, **Princípios de morfometria geométrica**, Ribeirão Preto, Ed. Holos, 198p., 1999.

MORPHOMETRICS at SUNY Stony Brook, <http://life.bio.sunysb.edu/morph/>, Última atualização, Julho-2002.

MORRISON, D. F., **Multivariate Statistical Methods**. New York, McGraw-Hill, 338p. 1976.

ROHLF, F. J., **Fitting curves to outlines**, Proceedings of the Michigan morphometrics workshop. Special publication number 2, University of Michigan Museum of Zoology, Ann Arbor. p. 167 – 178. 1990.

ROHLF, F. J., **Multivariate analysis of shape using partial warp scores**, Current issues in statistical shape analysis. Leed University Press, Leeds, 1995.

ROHLF, F. J., **Morphometric spaces, shape components and the effect of linear transformations**, **Advances in Morphometrics**, Life Sciences, vol. 284 Plenum Publishing Corp., New York. p. 117 – 130.1996.

ROHLF, F. J., **tpsDig 32**. Version 1.40. Programa para digitalizar marcos anatômicos para análises da morfometria geométrica. de N.Y.: State University at Stony Brook (program). 2003.

ROHLF, F. J., **tpsRelw: analysis of relative warps**. Version 1.35. Programa para realizar análise de deformações relativas. Department of Ecology and Evolution, N.Y.: State University at Stony Brook (program). 2003.

ROHLF, F. J. & D. E. SLICE., **Extensions of the Procrustes method for the optimal superimposition of landmarks**. Syst. Zool. 39; 40 – 59. 1990.

SCION CORPORATION, <http://www.scioncorp.com>, acessada em novembro 2003.

SCHILDT H., **C Completo e Total**, 3ª ed., MAKRON Books, São Paulo, SP, 1996.

SOKAL, R. R. & ROHLF, F. J., **Biometry: The Principles and Practice of Statistics in Biological Research**. W. H. Freeman, San Francisco, 859p. 1981.

TRIBE, C. J. A., **Mastofauna do estado do Rio de Janeiro, com especial referência à ordem Polyprotodontia (marsupiais)**. Rio de Janeiro, Dissertação (Mestrado em Zoologia). Universidade do Rio de Janeiro. 244p. 1987.

VALENTIN, J. L.; **Ecologia Numérica – Uma introdução à análise multivariada de dados ecológicos**. Editora Interciência. Rio de Janeiro. 2000.

VIEIRA, J. W., STOSIC B., LIMA F. R. A. KRAMER R., KHOURY H., **Um Software Para Editar Fantomas de Voxels e Calcular Coeficientes de Conversão para a Proteção Radiológica**. X Jornadas Portuguesas de Protecção Contra Radiações, Lisboa - Portugal, Dez. 2003.

APÊNDICE 1

ANÁLISE MULTIVARIADA

A Análise de Variância Multivariada (MANOVA)

É o estudo de diferenças de grupos localizados num espaço de medidas multidimensional. O objetivo desta técnica é verificar a existência de diferenças entre tratamentos, com relação a diversas variáveis, simultaneamente.

A MANOVA será utilizada para testar a igualdade dos vetores médios das medidas das variáveis. O teste é realizado entre o resíduo e o total. Um teste de significância muito utilizado é o teste de Wilks (Λ), assim definido:

$$\Lambda = \frac{\det E}{\det(H + E)}$$

em que:

det = determinante da matriz;

E = matriz de somas de quadrados e de produtos referente a resíduo;

H = matriz de soma de quadrados e de produtos referente a tratamentos.

Quando existe diferença sistemática esperamos obter $\Lambda < 1$ quanto mais significativa a diferença, menor o seu valor, desta forma a hipótese de nulidade é rejeitada.

A Análise dos Componentes Principais

A Análise de Componentes Principais (ACP) foi apresentada de forma original em 1901 por Karl Pearson e em 1933 e 1936 foi aplicada por Hotelling em diversas áreas científicas, porém seu uso só foi acentuado quando houve disponibilidade de recursos na área da informática.

Um conjunto de variáveis de um espaço multidimensional é transformado em um outro conjunto equivalente, plotado no espaço bi ou tridimensional, com a preservação da variância dos dados originais, e obtendo-se assim sua representação gráfica que é a forma mais simples de se obter alguma informação

sobre a existência de padrões em termos de associações entre indivíduos ou variáveis. Essa técnica é utilizada para a redução do número de variáveis a um número menor de índices (chamado de componentes principais) que são combinações lineares das variáveis originais (MANLY, 1995), independentes entre si e estimados com o propósito de reter, em ordem de estimação, o máximo de informações em termos de variação total, contida nos dados iniciais (CRUZ e REGAZZI, 1994). Segundo Johnson e Wichern (1992) os primeiros componentes devem absorver de 80 a 90% da variação total, para serem usados como substitutos dos dados originais.

Os componentes principais são assim definidos:

$$Y_i = a_{i1} \cdot X_1 + a_{i2} \cdot X_2 + \dots + a_{ip} \cdot X_p = \sum_{j=1}^p a_{ij} X_j$$

sendo: Y_i = i-ésimo componente principal ($i = 1, 2, \dots, p$);

X_j = j-ésima variável ($j = 1, 2, \dots, p$);

a_{ij} = coeficiente de ponderação da j-ésima variável no i-ésimo componente.

Cada componente associado a cada variável representa a contribuição dessa variável para o componente principal e o seu valor é função do conjunto de variáveis incluídas na análise. Os componentes são estimados de modo que o primeiro retenha a maior variância, o segundo retenha a segunda maior variância e assim sucessivamente (MANLY, 1995). Logo:

$$\text{Var}(Y_1) > \text{Var}(Y_2) > \dots > \text{Var}(Y_p) \text{ e } \sum_{i=1}^p \text{var}(Y_i) = \sum_{j=1}^p \text{var}(X_j).$$

Esta propriedade é importante por permitir a identificação e o descarte das variáveis menos informativas para a divergência, ao mesmo tempo que permite formação de agrupamentos (DIAS *et al*, 1997).

Os Componentes Principais são obtidos pela solução do sistema $\det(S - \lambda_i I) = 0$; onde: S é a matriz de covariância; λ_i são as raízes características ou autovalores de S e I é a matriz identidade $p \times p$.

A importância relativa de um componente é avaliada pela percentagem da variação total que ele explica (JOHNSON e WICHERN, 1992);

$$\text{Importância de } Y_i = \frac{\lambda_i}{\sum_{k=1}^p \lambda_k} = \frac{\lambda_i}{\text{Tr}(S)}, \quad K = 1, 2, \dots, p$$

sendo $\text{Tr}(S)$ o traço da matriz S .

Associado ao autovalor λ_i existe o autovetor v_i obtido do sistema $(S - \lambda_i I)$, $v_i = 0$, e fazendo-se $V_i^* = \frac{V_i}{\|V_i\|}$ encontra-se o vetor normalizado. Os autovalores de S correspondem às variâncias de cada componente e os autovetores normalizados, aos coeficientes de ponderação dos caracteres.

A correlação entre as variáveis e os componentes principais é definida por Johnson e Wichern (1992) como

$$r_{ij} = \frac{\text{cov}(X_j, Y_i)}{\sqrt{\text{var}(X_j)} \cdot \sqrt{\text{var}(Y_i)}}.$$

A Análise das Variáveis Canônicas

A Análise das Variáveis Canônicas (AVC) fornece uma descrição das diferenças entre grupos especificados a priori em um conjunto de dados multivariados. Os grupos podem representar indivíduos de populações de uma espécie ou espécies diferentes. O interesse é verificar a relação da magnitude de diferenças entre os grupos, relativa àquela dentro dos grupos (MONTEIRO e REIS, 1999).

Os princípios fundamentais de redução de dimensionalidade e explicação da variação em termos de variáveis latentes (obtidas por meio de autovalores e autovetores) aplicam-se também para a análise de variáveis canônicas. No caso da ACP delineada para analisar matrizes de dados não-estruturados, ou seja, uma única população multivariada, os autovalores e autovetores são extraídos de uma única matriz de variâncias e covariâncias. No caso da AVC, é necessário construir

uma matriz que quantifique a variação dentro dos grupos (considerando a soma dos quadrados dentro), W , e uma outra matriz que quantifique a variação entre os grupos (considerando a soma dos quadrados entre grupos), B , e são definidas como

$$W = \sum \sum (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T \text{ e } B = \sum (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$$

onde

x_{ij} denota a observação de uma das p variáveis no j -ésimo indivíduo no i -ésimo grupo,

$\bar{x}_i = \frac{\sum x_{ij}}{n_i}$ representa a média de uma dada variável no i -ésimo grupo e

$\bar{x} = \frac{\sum n_i \bar{x}_i}{n}$ indexa a média geral para uma dada variável para todos os grupos.

Os autovalores extraídos da matriz produto $W^{-1}B$ são interpretados como a quantidade de variação associada a cada autovetor ou eixo de maior variação.

A Análise Discriminante

A Análise Discriminante (AD) foi inicialmente desenvolvida por Fisher em 1936, procurando solucionar problemas em Antropologia e Biologia (KLECKA, 1980). A função discriminante linear de Fisher (FDLF) é uma combinação linear das características observadas que apresenta maior poder de discriminação entre os grupos e com a propriedade de minimizar as probabilidades de má classificação, quando as populações são normalmente distribuídas com parâmetros média μ_i e variância Σ_i conhecidos.

A AD permite ao pesquisador estudar as diferenças entre dois ou mais grupos, considerando diversas variáveis, simultaneamente. Este tipo de análise é destinado a interpretar grupos de objetos, definidos a priori. Portanto o objetivo da técnica não é estabelecer grupos, pois os mesmos já devem ser previamente conhecidos, mas interpretá-los a partir das variáveis (VALENTIN, 1995).

Segundo Klecka (1980), e Manly (1995), a definição matemática de uma função discriminante é a seguinte:

$$Z_{jk} = a_1 \cdot X_{1jk} + a_2 \cdot X_{2jk} + \dots + a_p \cdot X_{pjk}$$

sendo:

Z_{jk} = valor (escore) na função discriminante para o caso k no grupo j;

X_i = valor correspondente à i-ésima variável;

a_i = coeficiente de ponderação da i-ésima variável.

A Análise Discriminante consiste em obter funções que permitam classificar um indivíduo X, com base em medidas de um número p de características do mesmo, em várias populações π_i ($i = 1, 2, \dots, g$), visando minimizar a probabilidade de má classificação em uma população π_i , quando este pertence à população π_j para $i \neq j$, ($j=1, 2, \dots, g$) (MARDIA *et al.*, 1997).

Para classificação de um indivíduo X, com base em medidas de um número p de características do mesmo, em uma das duas populações π_1, π_2 distintas, buscando minimizar a probabilidade de classificar erroneamente um indivíduo, na população π_1 , quando ele realmente pertence à população π_2 , considera-se o seguinte (Johnson e Wichern, 1992):

I) $D(\tilde{x}) = \tilde{x}' \tilde{x} = [\mu_{\sim 1} - \mu_{\sim 2}]' \Sigma^{-1} \tilde{x}$ que é a função discriminante linear de Fisher (FDLF)

Sendo $D(\tilde{x}_{\sim 0}) = \tilde{x}_{\sim 0}' \tilde{x}_{\sim 0} = [\mu_{\sim 1} - \mu_{\sim 2}]' \Sigma^{-1} \tilde{x}_{\sim 0}$ o valor de FDLF para uma nova observação $\tilde{x}_{\sim 0}$;

II) $m = \frac{1}{2} \left(D\left(\mu_{\sim 1}\right) + D\left(\mu_{\sim 2}\right) \right)$ é o ponto médio entre as duas médias populacionais

univariadas $D\left(\mu_{\sim 1}\right); D\left(\mu_{\sim 2}\right)$

III) Regra de classificação:

Alocar $x_{\sim 0}$ em π_1 se $D(x_{\sim 0}) = [\mu_{\sim 1} - \mu_{\sim 2}]' \Sigma^{-1} x_{\sim 0} \geq m$;

Alocar $x_{\sim 0}$ em π_2 se $D(x_{\sim 0}) = [\mu_{\sim 1} - \mu_{\sim 2}]' \Sigma^{-1} x_{\sim 0} < m$;

IV) Estimativas dos parâmetros média $\mu_{\sim i}$ e variância Σ .

Sejam duas populações normais multivariadas π_1, π_2 . Supondo-se que n_1 é o número de observações da variável aleatória multivariada $X' = [x_1 \ x_2 \ \dots \ x_p]$ de π_1 e n_2 medidas destas quantidades de π_2 . As respectivas matrizes de dados são:

$X_1 = \begin{pmatrix} x_{\sim 11} & x_{\sim 12} & \dots & x_{\sim 1p} \end{pmatrix}$ sendo $x_{\sim 11} \ x_{\sim 12} \ \dots \ x_{\sim 1p}$ observações $N_p(\mu_{\sim 1}; \Sigma)$;

$X_2 = \begin{pmatrix} x_{\sim 21} & x_{\sim 22} & \dots & x_{\sim 2p} \end{pmatrix}$ sendo $x_{\sim 21} \ x_{\sim 22} \ \dots \ x_{\sim 2p}$ observações $N_p(\mu_{\sim 2}; \Sigma)$;

V) Vetores de médias amostrais:

$\bar{x}_{\sim 1} = \frac{1}{n_1} \sum_{j=1}^{n_1} x_{\sim 1j}$; $\bar{x}_{\sim 2} = \frac{1}{n_2} \sum_{j=1}^{n_2} x_{\sim 2j}$ são do tipo (px1);

VI) Matrizes de covariâncias:

$S_1 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (x_{\sim 1j} - \bar{x}_{\sim 1})(x_{\sim 1j} - \bar{x}_{\sim 1})'$; $S_2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (x_{\sim 2j} - \bar{x}_{\sim 2})(x_{\sim 2j} - \bar{x}_{\sim 2})'$, matrizes

(pxp);

$S_C = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2}$ matriz de covariância combinada do tipo (pxp).

Existindo S_C^{-1} então $D(x_{\sim 0}) = \hat{e}' x_{\sim 0} = [\bar{x}_{\sim 1} - \bar{x}_{\sim 2}]' S_C^{-1} x_{\sim 0}$ e segue $\hat{m} = \frac{1}{2} \left[\left(D(\bar{x}_{\sim 1}) \right) + D(\bar{x}_{\sim 2}) \right]$

VII) Regra de decisão:

Alocar $x_{\sim 0}$ em π_1 se $D(x_{\sim 0}) = [\bar{x}_{\sim 1} - \bar{x}_{\sim 2}]' S_C^{-1} x_{\sim 0} \geq \hat{m}$

Alocar $x_{\sim 0}$ em π_2 se $D(x_{\sim 0}) < \hat{m}$

VIII) Avaliação da função discriminante:

$$D^2 = [\bar{x}_{\sim 1} - \bar{x}_{\sim 2}]' S_C^{-1} \cdot [\bar{x}_{\sim 1} - \bar{x}_{\sim 2}] = D(\bar{x}_{\sim 1}) - D(\bar{x}_{\sim 2})$$

$H_0: \mu_{\sim 1} = \mu_{\sim 2}$ versus $H: \mu_{\sim 1} \neq \mu_{\sim 2}$

$$F_0 = \left(\frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p} \right) \cdot \left(\frac{n_1 \cdot n_2}{n_1 + n_2} \right) D^2 \text{ onde } n_1 \text{ e } n_2 \text{ são os tamanhos das amostras em}$$

π_1, π_2 respectivamente, e p é o número de variáveis que têm distribuição F com $v_1 = p$ e $v_2 = n_1 + n_2 - p - 1$ graus de liberdade.

IX) Método de estimação das probabilidades de má classificação, devido a Smith (1947).

$\hat{p}(j/i)$ é a estimativa da probabilidade de classificarmos erroneamente uma observação de π_i em $\pi_j, i \neq j$,

$$\hat{p}(j/i) = \hat{p}(i) = \frac{m_i}{n_i}; \quad i = 1, 2.$$

X) Taxa de erro aparente: $T \left(\begin{matrix} \hat{R} \\ \hat{f} \end{matrix} \right) = \frac{\sum_{i=1}^g m_i}{\sum_{i=1}^g n_i}$, sendo m_i = número de observações de má

classificação em π_i e n_i = número de observações em π_i .

XI) Distância de Mahalanobis entre dois grupos:

$$D^2_{ij} = \left(\bar{x}_{\sim i} - \bar{x}_{\sim j} \right)' S_C^{-1} \left(\bar{x}_{\sim i} - \bar{x}_{\sim j} \right) \text{ sendo } i \neq j = 1, 2, \dots, g,$$

Quanto maior for os D^2_{ij} 's melhor, pois isto indica que os grupos são bastante distintos e portanto a classificação de um novo indivíduo num dos grupos se fará com a máxima probabilidade de acerto,

APÊNDICE 2

ANÁLISE DE DEFORMAÇÕES (FUNÇÃO THIN PLATE SPLINES)

A análise começa com as coordenadas x e y dos p marcos anatômicos da configuração média ou organismo de referência, X_c é a matriz ($2 \times p$) de coordenadas da configuração média (CM), com a primeira linha sendo as abscissas x e a segunda linha as ordenadas y dos marcos anatômicos. A matriz Q ($p \times 3$) tem a primeira coluna de elementos iguais a 1, a segunda coluna com as abscissas x e a terceira com as ordenadas y dos marcos anatômicos da configuração média.

A segunda matriz simétrica P ($p \times p$), é formada usando-se os valores da função $U(r)$ que é definida como: $U(r) = r_{ij}^2 \ln(r_{ij}^2)$. Nesta função r_{ij} é a distância entre o i -ésimo e o j -ésimo marcos anatômicos da configuração média.

Uma matriz L , $((p+3) \times (p+3))$ é formada das submatrizes P e Q combinadas como segue:

$L = \begin{pmatrix} P & Q \\ Q^T & O \end{pmatrix}$ onde O é a matriz nula (3×3), Em geral, a formação de L é do seguinte modo:

$$L = \begin{pmatrix} 0 & U(r_{12}) & U(r_{13}) & \dots & U(r_{1p}) & 1 & x_1 & y_1 \\ U(r_{21}) & 0 & U(r_{23}) & \dots & U(r_{2p}) & 1 & x_2 & y_2 \\ U(r_{31}) & U(r_{32}) & 0 & \dots & U(r_{3p}) & 1 & x_3 & y_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ U(r_{p1}) & U(r_{p2}) & U(r_{p3}) & \dots & 0 & 1 & x_p & y_p \\ 1 & 1 & 1 & \dots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & x_3 & \dots & x_p & 0 & 0 & 0 \\ y_1 & y_2 & y_3 & \dots & y_p & 0 & 0 & 0 \end{pmatrix}$$

A inversa de L é indicada por L^{-1} . Uma submatriz L_p^{-1} é então formada considerando-se as p primeiras linhas e colunas de L^{-1} . A matriz L^{-1} (chamada matriz de energia de deformação) é simétrica, e é decomposta para encontrar seus autovalores e autovetores, ou seja,

$$L_p^{-1} = E \Lambda E^T$$

Onde

Λ é a matriz diagonal ($p \times p$) de autovalores e

E é uma matriz ($p \times p$) de autovetores cujas colunas são os autovetores de L_p^{-1} . Esses autovetores são as deformações principais que expressam os modos de variação possível nos espécimes.

Três dos autovalores em Λ são nulos, assim Λ torna-se uma matriz $((p-3) \times (p-3))$ e E torna-se uma matriz $(p \times (p-3))$.

O próximo passo é subtrair as x coordenadas da CM das coordenadas correspondentes dos x marcos anatômicos dos espécimes. Isto transforma cada coordenada X em X_x de seus valores observados no desvio da coordenada no objeto de referência. Isto é escrito em notação matricial como:

$$V_x = X_x - 1_n \otimes ([1 \ 0]X_c)$$

A expressão $(1_n \otimes [1 \ 0])$ denota o produto direto do vetor coluna que tem n linhas, com o valor 1 em cada linha com a matriz linha com “1” na primeira coluna e “0” na segunda. Este produto forma uma matriz $(n \times 2)$ com todos elementos da primeira coluna iguais a 1 e da segunda coluna iguais a zero. Multiplicando-se esta matriz pela matriz X_c produz-se uma matriz $(n \times p)$ com cada coluna contendo as coordenadas X da CM para cada marco anatômico correspondente. Similarmente, os desvios das coordenadas y da CM são obtidas como:

$$V_y = X_y - 1_n \otimes ([0 \ 1]X_c)$$

As matrizes V_x e V_y são concatenadas horizontalmente para formar a matriz V , onde as primeiras p colunas são formadas pelos desvios das X coordenadas x daquelas da CM, e da $(p+1)$ até a coluna $2p$ contém os desvios da coordenada Y . Cada linha de V corresponde a um único espécime, portanto há n linhas em V .

O próximo passo é posicionar os desvios dos espécimes individualmente em relação à posição do espécime de referência no espaço de deformações principais anteriormente definido como a matriz E . Este, e vários outros passos foram

realizados para formar uma matriz W de dimensões $(n \times 2(p-3))$ tal como:

$$W = \left(\frac{1}{\sqrt{n}} V \right) \left(I_2 \otimes (E \Lambda^{-\alpha/2}) \right)$$

Divide-se pela raiz quadrada de n , porque a matriz resultante foi uma matriz de soma de quadrados e produtos.

BOOKSTEIN (1991) sugeriu a introdução do parâmetro α no expoente de Λ na obtenção da matriz W de pesos. A matriz $\Lambda^{-\alpha/2}$ é usada para proporcionalizar os autovetores (deformações principais) E por seus autovalores criados por alguma potência que depende do valor de α . Diferentes valores de α produzem diferentes escalas nos escores, dependendo da quantidade de energia contida em cada deformação.

O valor $\alpha=0$ é de particular interesse para estudos explanatórios porque ele confere a todas as deformações principais o mesmo peso, confere o mesmo grau de importância às deformações principais que são usadas como base de vetores para o espaço. Neste espaço foi considerada a distância de Procrustes entre espécimes, uma vez que é livre de transformações tais como translação, rotação, e componente uniforme (ROHLF e SLICE, 1990).

Os desvios X e Y dos espécimes, foram posicionados separadamente da CM no sistema de coordenadas das deformações principais. Isto é devido ao produto direto da matriz identidade I_2 e $E \Lambda^{-\alpha/2}$, onde E uma matriz $(p \times (p-3))$, e Λ uma matriz $((p-3) \times (p-3))$, logo o produto é uma matriz de dimensões $(p \times (p-3))$. Considerando o produto direto de I_2 e o produto $E \Lambda^{-\alpha/2}$ formou-se uma matriz de quatro sub-matrizes $((p-3) \times p)$ tal como segue:

$$\begin{bmatrix} E \Lambda^{-\alpha/2} & 0 \\ 0 & E \Lambda^{-\alpha/2} \end{bmatrix}$$

As primeiras p colunas de V , correspondendo as abscissas X são transformadas por $E \Lambda^{-\alpha/2}$ (sub-matriz do canto superior esquerdo), e a segunda p colunas de V ,

correspondentes as ordenadas Y dos marcos anatômicos, são por sua vez transformadas por $E\Lambda^{-\alpha/2}$ (sub-matriz do canto inferior direito).

O resultado final é a matriz peso W ($n \times 2(p-3)$) que é uma concatenação horizontal dos pesos para as coordenadas X e Y, respectivamente. Colocando estes pontos (linhas = espécimes) representados em W num plano cartesiano obtém-se um gráfico de espécimes no espaço vetorial definido pela CM. W é portanto o conjunto de dados que representa a variação nas formas dos espécimes que pode ser adicionalmente analisada. Diferentes escalas produzem variâncias não homogêneas (heteroscedasticidade) na matriz W. Como a ACP é sensível à heteroscedasticidade, as deformações relativas terão mais influência das deformações parciais em maior escala.

A análise de deformações relativas pode ser feita pela decomposição do valor singular

$$SDR^T$$

Onde

S é uma matriz de escores com linhas correspondendo aos objetos e colunas às deformações relativas,

D é uma matriz diagonal de valores singulares (autovalores relativos),

R contém os coeficientes das deformações relativas (colunas) referentes às deformações parciais (linhas).

Isto é devido ao fator $\frac{1}{\sqrt{n}}$ que originalmente estava incluído na formação de W e é o equivalente aos graus de liberdade na análise de componentes principais.

S tem dimensões ($n \times \min(n-1, 2(p-3))$). Suas linhas correspondem aos espécimes, e cada coluna corresponde às deformações relativas com valor singular maior que zero. S é uma matriz de escores de cada espécime nos eixos associados às deformações relativas.

D representa uma matriz diagonal de valores singulares e tem dimensões $[\min(n - 1, 2(p - 3)) \times \min(n - 1, 2(p - 3))]$.

R tem dimensões $\min(n - 1, 2(p - 3)) \times \min(n - 1, 2(p - 3))$. As linhas de R correspondem às deformações principais e as colunas de R correspondem às deformações relativas. As deformações relativas são funcionalmente equivalentes aos componentes principais e podem ser examinados para padrões dentro do conjunto de dados.

A visualização dos efeitos da localização de um espécime ao longo das deformações relativas é feita pelo cálculo dos coeficientes relativos para cada marco anatômico (ao invés das deformações parciais). As deformações relativas podem ser representadas também como deformações de uma grade cartesiana sobreposta à referência.

A decomposição das mudanças de forma em localizadas e globais é muito importante para qualquer estudo morfométrico, pois permite estimativas da quantidade de integração total dos marcos durante a mudança de forma, assim como a verificação das mudanças de forma mais localizadas.

ANEXO

Tabela - Matriz de correlação ($p < 0,0001$) entre as 16 variáveis morfométricas dij

Var	d12	d13	d23	d24	d25	d34	d35	d45
d12	1,0000							
d13	0,7597	1,0000						
d23	0,7237	0,9029	1,0000					
d24	0,4299	0,4428	0,3542	1,0000				
d25	0,7221	0,9130	0,9837	0,4109	1,0000			
d34	0,7713	0,9066	0,9800	0,3923	0,9663	1,0000		
d35	0,6805	0,5183	0,4664	0,4279	0,5152	0,5321	1,0000	
d45	0,7277	0,9031	0,9768	0,3254	0,9805	0,9846	0,4854	1,0000
d46	0,7999	0,8298	0,7869	0,3749	0,8250	0,8425	0,6326	0,8521
d57	0,8544	0,8418	0,8109	0,5009	0,8256	0,8670	0,6090	0,8471
d67	0,4345	0,4663	0,4772	0,2514	0,4767	0,4922	0,2805	0,4851
d68	0,4489	0,3138	0,2104	0,2915	0,2289	0,2403	0,3372	0,2152
d69	0,4068	0,3839	0,2757	0,2898	0,2975	0,2892	0,3333	0,2771
d78	0,5077	0,4498	0,4136	0,2731	0,4111	0,4445	0,3454	0,4187
d79	0,3211	0,3134	0,1974	0,2608	0,2143	0,2189	0,2963	0,2017
d89	0,3719	0,4070	0,3586	0,2152	0,3708	0,3560	0,2660	0,3546

Var	d46	d57	d67	d68	d69	d78	d79	d89
d46	1,0000							
d57	0,9022	1,0000						
d67	0,3501	0,4194	1,0000					
d68	0,2869	0,3859	0,3517	1,0000				
d69	0,2580	0,3767	0,5126	0,6812	1,0000			
d78	0,4464	0,4652	0,5438	0,6665	0,4951	1,0000		
d79	0,2444	0,3004	0,2864	0,4879	0,7998	0,5823	1,0000	
d89	0,3275	0,3684	0,4359	0,3387	0,6910	0,5472	0,5928	1,0000