

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE AQUICULTURA**

Eduardo Luiz Tavares Gonçalves

**MODELOS PREDITIVOS COMO FERRAMENTAS PARA
ESTIMAR A QUANTIDADE DE PROTOZOÁRIOS PARASITOS
EM PEIXES**

Dissertação submetida ao Programa de Pós-Graduação em Aqüicultura da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Aqüicultura.

Orientador:

Prof. Dr. **Maurício Laterça Martins**

Co-orientador:

Prof. Dr. **Alex Pires de Oliveira Nuñez**

Florianópolis
2011

Catálogo na fonte pela Biblioteca Universitária
da
Universidade Federal de Santa Catarina

G635m Gonçalves, Eduardo Luiz Tavares

Modelos preditivos como ferramentas para estimar a quantidade de protozoários parasitos em peixes [dissertação] / Eduardo Luiz Tavares Gonçalves ; orientador, Maurício Laterça Martins. - Florianópolis, SC, 2011.

139 p. : il. , grafs. , tabs.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências Agrárias. Programa de Pós-Graduação em Aquicultura.

Inclui referências

1. Aquicultura. 2. Controle preditivo. 3. Protozoário. 4. Tilápia (Peixe). I. Martins, Maurício Laterça. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Aquicultura. III. Título

CDU 639.3

Eduardo Luiz Tavares Gonçalves

**MODELOS PREDITIVOS COMO FERRAMENTAS PARA
ESTIMAR A QUANTIDADE DE PROTOZOÁRIOS PARASITOS
EM PEIXES**

Esta Dissertação foi julgada adequada para obtenção do Título de Mestre em Aquicultura e aprovada, em sua forma final, pelo Programa de Pós-Graduação em Aquicultura da Universidade Federal de Santa Catarina.

Florianópolis, 15 de Junho de 2011.

Prof. Dr. Evoy Zaniboni Filho
Coordenador do Curso

Banca Examinadora:

Prof., Dr. Maurício Laterça Martins
Orientador
Universidade Federal de Santa Catarina

Prof.^a, Dr.^a Sonia Palomino Bean
Membro
Universidade Federal de Santa Catarina

Prof., Dr. Cláudio José Von Zuben
Membro
Universidade Estadual Paulista

Dedico este trabalho aos meus pais,
minha avó, meus irmãos, meu amor, e
todos aqueles com quem eu realmente
posso contar.

AGRADECIMENTOS

Agradeço aos meus pais, José Pedro e Celina Maria, que sempre orientaram e apoiaram o meu crescimento pessoal e profissional;

À minha namorada, Joana Aparecida da Silveira do Amarante, que ficou ao meu lado, dando-me força e incentivando-me a continuar;

Ao Jivago Medeiros Ribeiro, pela programação do software e pela amizade de longa data;

Aos meus orientadores Prof. Dr. Maurício Laterça Martins e Prof. Dr. Alex Pires de Oliveira Nuñez, pelo apoio dado apesar da distância;

Aos professores que ajudaram a me aventurar por áreas até então desconhecidas, Prof. Dr. José Francisco Danilo de Guadalupe Correa Fletes e Prof. Dra. Sonia Palomino Bean;

Ao Roberto Hoppe, da Fundação Municipal 25 de Julho, e ao Juan Ramon Esquivel Garcia, da Piscicultura Panamá, pela grande ajuda e paciência no fornecimento dos peixes;

Aos meus amigos do laboratório AQUOS: Aline Brum, Ágata Passeto, Daniela Bambi, Gabriela Hashimoto, Gabriela Tomas Jerônimo, Geovana Dotta, Gustavo Melchert, Jerko Ledic, Karen Tancredo, Katina Roumbedakis, Natália Marchiori e Samantha Fontanella;

Aos meus amigos pessoais, que sempre estão presentes não importando a distância: Alfredo dos Santos Weber, Breno Batista Liberato Souza Souza, João Luís Chainça Neto, Heitor Fabrício Klaus, Maíra Matos, Nathasha Karla Pimenta Said e Ynaiã Silva Benthroldo;

Por fim agradeço a todos que acreditaram em mim, mostrando que é possível seguir uma idéia e transformá-la em algo concreto.

Durchaus studirt, mit heißem Bemühn.
Da steh' ich nun, ich armer Thor!
Und bin so klug als wie zuvor.

(Goethe, 1808)

RESUMO

Protozoários parasitos são causadores de mortalidade em peixes cultivados, porém existe um número muito pequeno de trabalhos com análises quantitativas sobre o assunto no Brasil. Uma gama de métodos para quantificar parasitos é conhecida, entretanto eles geralmente requerem equipamentos caros e mão de obra especializada. Neste estudo foi desenvolvido um modelo preditivo e, a partir dele, um software, com o objetivo de auxiliar na quantificação de protozoários parasitos nas brânquias de peixe, tornando-a mais precisa e confiável, levando em consideração a realidade brasileira. Tilápias do Nilo (n=129) foram coletadas e tiveram um total de oito variáveis aferidas e testadas no processo de desenvolvimento do modelo preditivo. Os protocolos de coleta e aferição das variáveis, juntamente com a modelagem por meio de MARSplines, provaram ser eficientes e resultaram em um modelo preditivo bastante preciso e robusto. Esses resultados subsidiaram o desenvolvimento de um software, batizado de *ProtoCount*, que é capaz de estimar a quantidade total de protozoários nas brânquias, a partir dos valores de peso e comprimento do peixe e da média de protozoários observados em amostras de 1,5 ml, retiradas de 200 ml de solução formalina contendo raspado das brânquias do peixe amostrado e contadas em câmara de McMaster. *ProtoCount* demonstrou ser prático e confiável, assim, espera-se que sua utilização possa contribuir com a expansão, padronização, precisão e confiabilidade das análises quantitativas feitas por laboratórios nacionais de parasitologia de peixes.

Palavras-chave: Aquicultura. Modelagem. Protozoários. Tilápia do Nilo. MARSplines.

ABSTRACT

Protozoan parasites are a major cause of cultured fish mortality, but there are surprisingly few data with quantitative measurements in Brazil. There are few reliable methods for quantifying parasite protozoa, but either they require expensive equipments or highly specialized professionals. This study has developed a predictive model and a software, with the objective of providing more precise and reliable quantifications of protozoan parasites on fish gills more, considering Brazilian reality. Nile tilapia ($n = 129$) were collected and a total of eight variables were measured and tested for the predictive model development. Collecting and measurement procedures, along with MARSplines modeling, were efficient and resulted in a very precise and robust predicting model. Resulting data helped in developing a software, called *ProtoCount*, that manages to estimate the total protozoa burned on fish gills, based on fish's weight, length and mean protozoa count on a 1,5 ml sample. Each 1,5 ml corresponds to 10 homogenized samples, taken from 200 ml formalin solution, containing scraped gills from sample fish, and counted using McMaster's chamber. *ProtoCount* proved itself to be simple and reliable, therefore capable of increasing the number, standardization, precision and reliability of quantitative analysis performed in national fish parasitology laboratories.

Keywords: Aquaculture. Modeling. Protozoa. Nile tilapia. MARSplines.

LISTA DE FIGURAS

Figura A: Diagrama de Venn com relação hospedeiro-parasita-ambiente.....	26
Figura 1: Fluxograma do processo de coleta dos animais.....	52
Figura 2: Fluxograma simplificado do processo de aferição das variáveis para modelagem.....	52
Figura 3: Fluxograma detalhado do processo de aferição das variáveis para modelagem	56
Figura 4: Fluxograma simplificado do processo de desenvolvimento do modelo.	57
Figura 5: Fluxograma detalhado do processo de desenvolvimento do modelo..	60
Figura 6: a) Histograma do peso do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	62
Figura 7: a) Histograma do comprimento do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	63
Figura 8: a) Histograma da altura do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	64
Figura 9: a) Histograma da área superficial relativa do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	66
Figura 10: a) Histograma do peso da brânquia hidratada com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	67
Figura 11: a) Histograma do peso seco ósseo com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	68
Figura 12: a) Histograma do peso seco de tecido branquial na amostra (PSTA) com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	70
Figura 13: a) Histograma do peso seco de tecido branquial (PST) com estatística descritiva b) médias com intervalo de confiança (95%) das classes.....	71
Figura 14: Resultado da ANOVA para medidas repetidas comparando as médias e intervalos de confiança (95%) das repetições das alíquotas de 1,5ml.....	72
Figura 15: Captura de tela com resultado de ANOVA via <i>bootstrap</i> de 100.000 repetições comparando as repetições das alíquotas de 1,5ml.....	73
Figura 16: Curvas de regressão linear, onde X = peso do peixe e Y = peso da brânquia hidratada (PBH), peso seco ósseo (PSO), peso seco do tecido branquial na amostra (PSTA) e peso seco total do tecido branquial (PST).	76
Figura 17: Curvas de regressão potencial, onde X = peso do peixe e Y = peso da brânquia hidratada (PBH), peso seco ósseo (PSO), peso seco do tecido branquial na amostra (PSTA) e peso seco total do tecido branquial (PST).	77
Figura 18: Plotagem dos resíduos da estimativa de peso da brânquia hidratada (PBH).....	78
Figura 19: Plotagem dos resíduos da estimativa de peso seco ósseo (PSO).	79
Figura 20: Plotagem dos resíduos da estimativa de peso seco do tecido branquial na amostra (PSTA).	80

Figura 21: Plotagem dos resíduos da estimativa de peso seco total do tecido branquial (PST).	81
Figura 22: Captura de tela com resultado da análise de correlação por <i>bootstrap</i> entre peso do peixe e peso da brânquia hidratada (PBH).....	82
Figura 23: Captura de tela com resultado da análise de correlação por <i>bootstrap</i> entre peso do peixe e peso seco ósseo (PSO).	83
Figura 24: Captura de tela com resultado da análise de correlação por <i>bootstrap</i> entre peso do peixe e peso seco do tecido branquial na amostra (PSTA).	84
Figura 25: Captura de tela com resultado da análise de correlação por <i>bootstrap</i> entre peso do peixe e peso seco do tecido branquial (PST).	85
Figura 26: Plotagem da função para PBH, onde X = peso do peixe, Y = comprimento do peixe e Z = peso da brânquia hidratada.	89
Figura 27: Plotagem da função para PSTA, onde X = peso do peixe, Y = comprimento do peixe e Z = Peso seco de tecido branquial em amostra de 1,5 ml.....	90
Figura 28: Plotagem da função para PST, onde X = peso do peixe, Y = comprimento do peixe e Z = Peso seco total do tecido branquial.....	91
Figura 29: a) Curva de regressão entre os valores observados vs. estimados de peso da brânquia hidratada (PBH) b) plotagem dos resíduos das estimativas de peso da brânquia hidratada (PBH).	93
Figura 30: a) Curva de regressão entre os valores observados vs. estimados de peso seco do tecido branquial na amostra (PSTA) b) plotagem dos resíduos das estimativas de peso seco do tecido branquial na amostra (PSTA).	95
Figura 31: a) Curva de regressão entre os valores observados vs. estimados de peso seco do tecido branquial (PST) b) plotagem dos resíduos das estimativas de peso seco do tecido branquial (PST).	97
Figura 32: Resultado do teste T de Student para amostras pareadas para PBH. obs. = observados; est. = estimados; DP = desvio padrão.	98
Figura 33: Resultado do teste T de Student para amostras pareadas para PSTA. obs. = observados; est. = estimados; EP = erro padrão.	99
Figura 34: Resultado do teste T de Student para amostras pareadas para PST. obs. = observados; est. = estimados; EP = erro padrão.	100
Figura i: Captura de tela com a interface do software <i>ProtoCount</i>	118
Figura ii: Captura de tela com detalhe do menu “Arquivo” e do comando de “Importar Planilha”.	119
Figura iii: Captura de tela com as informações contidas na janela “Sobre” do software.	119
Figura iv: Captura de tela com detalhe da mensagem confirmando o processamento dos dados importados com a planilha .xls.	120
Figura v: Fluxograma de funcionamento do software <i>ProtoCount</i>	121
Figura vi: Gráfico de dispersão com valores de QTP estimados pelo software <i>ProtoCount</i> e pelo método original.	122
Figura vii: Gráfico de dispersão com valores de QTP estimados pelo software <i>ProtoCount</i> e pelo método original.	123

LISTA DE TABELAS

Tabela 1: Resumo dos resultados da estatística da regressão. Florianópolis, junho de 2011.	74
Tabela 2: Estatística descritiva para as variáveis utilizadas na regressão múltipla. Florianópolis, junho de 2011.*	74
Tabela 3: Distribuição dos resultados da regressão múltipla para a variável peso da brânquia hidratada (PBH). Florianópolis, junho de 2011.*	75
Tabela 4: Análise de variância dos componentes da regressão múltipla para a variável peso da brânquia hidratada (PBH). Florianópolis, junho de 2011. *	75
Tabela 5: Matriz de correlações obtidas na regressão múltipla. Florianópolis, junho de 2011.*	75
Tabela 6: Comparação entre os valores de média e desvio padrão observados e estimados, para cada uma das variáveis dependentes. Florianópolis, junho de 2011.*	77
Tabela 7: Resultados do Teste T pareado comparando os valores estimados de peso da brânquia hidratada (PBH) com os observados. Florianópolis, junho de 2011.	78
Tabela 8: Resultados do Teste T pareado comparando os valores estimados de peso seco ósseo (PSO) com os observados. Florianópolis, junho de 2011.	79
Tabela 9: Resultados do Teste T pareado comparando os valores estimados de peso seco do tecido branquial na amostra (PSTA) com os observados. Florianópolis, junho de 2011.	80
Tabela 10: Resultados do Teste T pareado comparando os valores estimados de peso seco total do tecido branquial (PST) com os observados. Florianópolis, junho de 2011.	81
Tabela 11: Resultado da análise prévia para importância dos preditores. Florianópolis, junho de 2011.	86
Tabela 12: Especificações do modelo desenvolvido através da MARSplines. Florianópolis, junho de 2011.	87
Tabela 13: Distribuição do resultado da estatística descritiva da MARSplines. Florianópolis, junho de 2011.	88
Tabela 14: Distribuição dos coeficientes do modelo desenvolvido através da MARSplines. Florianópolis, junho de 2011.*	88
Tabela 15: Importância dos preditores no modelo desenvolvido através da MARSplines. Florianópolis, junho de 2011.....	88
Tabela 16: Resumo da qualidade de ajuste para a equação de peso da brânquia hidratada (PBH). Florianópolis, junho de 2011.	92
Tabela 17: Resumo da qualidade de ajuste para a equação de peso seco do tecido branquial na amostra (PSTA). Florianópolis, junho de 2011.	94
Tabela 18: Resumo da qualidade de ajuste para a equação de peso seco do tecido branquial (PST). Florianópolis, junho de 2011.	96

Tabela i: Resultados do Teste T pareado comparando os valores de quantidade total de protozoários (QTP) estimados pelo software <i>ProtoCount</i> e pelo método original. Florianópolis, junho de 2011.....	122
Tabela ii: Resultados do Teste T pareado comparando os valores de quantidade total de protozoários (QTP) estimados pelo software <i>ProtoCount</i> e pelo método original. Florianópolis, junho de 2011.....	123

LISTA DE ABREVIATURAS E SIGLAS

ASR = Área superficial relativa

ANOVA = Análise de variância

C = comprimento do peixe

QPA = Quantidade de protozoários contados na câmara de McMaster em amostra de 1,5 ml

QTP = Quantidade total de protozoários nas brânquias do peixe

MARSplines = *Multivariate adaptive regression splines*

P = peso do peixe

PBH = Peso da brânquia hidratada

PSO = Peso seco ósseo

PST = Peso seco total do tecido branquial

PSTA = Peso seco de tecido em uma amostra de 1,5 ml de solução formalina 10% com raspado de brânquias

VA = volume da amostra quantificada em câmara de McMaster

VT = volume total da solução de formalina contendo raspado de tecido branquial

SUMÁRIO

INTRODUÇÃO	25
PARASITISMO E SUAS RELAÇÕES.....	25
PARASITOLOGIA E AQUICULTURA	28
<i>PISCINOODINIUM PILLULARE</i>	29
TRICODINÍDEOS.....	30
<i>ICHTHYOPHTHIRIUS MULTIFILIIS</i>	30
MÉTODOS UTILIZADOS NA QUANTIFICAÇÃO DE PROTOZOÁRIOS PARASITOS.....	31
ABORDAGEM MOLECULAR.....	31
ABORDAGEM ESTEREOLÓGICA.....	32
ABORDAGEM DE ANÁLISE DE IMAGENS.....	32
ABORDAGEM DE CULTURA.....	33
QUANTIFICAÇÃO DE PROTOZOÁRIOS ICTIOPARASITOS NO BRASIL.....	33
ABORDAGEM PROPOSTA	34
ESPÉCIE ALVO	35
COMPREENDENDO A MODELAGEM MATEMÁTICA	35
BOOTSTRAP	36
MÉTODO DOS MÍNIMOS QUADRADOS (MMQ)	36
APRENDIZAGEM DE MÁQUINA (MACHINE LEARNING)	37
CONJUNTO DE TREINAMENTO E CONJUNTO DE TESTE	37
QUALIDADE DE AJUSTE (GOODNESS OF FIT)	37
SOBRE-AJUSTE (OVERFITTING).....	37
PODA (PRUNING)	38
VALIDAÇÃO CRUZADA (CROSS-VALIDATION)	38
ERRO DE VALIDAÇÃO CRUZADA GENERALIZADA (GCV ERROR)	38
LIMIAR (THRESHOLD)	38
FERRAMENTAS DE MODELAGEM COMPUTACIONAL	39
REDES NEURAIS ARTIFICIAIS (ARTIFICIAL NEURAL NETWORKS - ANN)	39
K-NEAREST NEIGHBORS (K-NN)	40
SUPPORT VECTOR MACHINE (SVM)	40
MULTIVARIATE ADAPTIVE REGRESSION SPLINES (MARSPLINES)	41
OBJETIVOS.....	42
OBJETIVO GERAL.....	42
OBJETIVOS ESPECÍFICOS.....	42

CAPÍTULO 1.....	43
DESENVOLVIMENTO DE UM MODELO PREDITIVO PARA ESTIMAR PARÂMETROS UTILIZADOS NA QUANTIFICAÇÃO DE PROTOZOÁRIOS PARASITOS DE PEIXES	43
ABSTRACT	44
RESUMO	45
INTRODUÇÃO	46
MATERIAIS E MÉTODOS.....	51
1.1. COLETA DOS ANIMAIS.....	51
1.2. AFERIÇÃO DAS VARIÁVEIS PARA A MODELAGEM	52
2.2.1. <i>Peso do peixe</i>	52
2.2.2. <i>Comprimento, altura e área superficial relativa do peixe (ASR)</i>	52
2.2.3. <i>Peso da brânquia hidratada (PBH)</i>	53
2.2.4. <i>Peso seco ósseo (PSO)</i>	53
2.2.5. <i>Peso seco de tecido branquial em amostra de 1,5 ml (PSTA)</i>	54
2.2.6. <i>Peso seco total do tecido branquial (PST)</i>	55
1.3. DESENVOLVIMENTO DO MODELO	57
2.3.1. <i>Análises prévias</i>	57
2.3.2. <i>Teste de homogeneidade de amostras</i>	57
2.3.3. <i>Correlação dos dados</i>	58
2.3.4. <i>Modelagem</i>	59
RESULTADOS	61
2.1. ESTATÍSTICA DESCRITIVA	61
3.1.1. <i>Peso do peixe</i>	61
3.1.2. <i>Comprimento</i>	62
3.1.3. <i>Altura</i>	63
3.1.4. <i>Área superficial relativa (ASR)</i>	65
3.1.5. <i>Peso da brânquia hidratada (PBH)</i>	66
3.1.6. <i>Peso seco ósseo (PSO)</i>	67
3.1.7. <i>Peso seco de tecido branquial em amostra de 1,5 ml (PSTA)</i>	69
3.1.8. <i>Peso seco total do tecido branquial (PST)</i>	70
2.2. DESENVOLVIMENTO DO MODELO	72
3.2.1. <i>Teste de homogeneidade de amostras</i>	72
3.2.2. <i>Regressões e análise das correlações</i>	73
3.2.3. <i>Modelagem</i>	86
DISCUSSÃO.....	101
3.1. UMA QUESTÃO DE LÓGICA	101
3.2. ACERCA DO PROCESSO DE COLETA.....	102
3.3. ACERCA DAS VARIÁVEIS AFERIDAS	102

3.4.	ACERCA DAS REGRESSÕES E ANÁLISE DAS CORRELAÇÕES	103
3.5.	ACERCA DA MODELAGEM.....	104
	CONCLUSÃO	105
	AGRADECIMENTOS	105
	REFERÊNCIAS.....	106
CAPÍTULO 2	111
<i>PROTOCOUNT</i>: UMA FERRAMENTA INOVADORA PARA QUANTIFICAR PARASITOS PROTOZOÁRIOS EM BRÂNQUIAS DE PEIXES	111
	ABSTRACT	112
	RESUMO.....	113
	INTRODUÇÃO.....	114
	MATERIAIS E MÉTODOS	116
	RESULTADOS	117
3.1.	INTERFACE <i>PROTOCOUNT</i>	117
3.2.	UTILIZAÇÃO DO <i>PROTOCOUNT</i>	119
3.3.	SIMULAÇÃO E COMPARAÇÃO DAS ESTIMATIVAS	121
	DISCUSSÃO	124
4.1.	<i>PROTOCOUNT</i>	124
4.2.	SIMULAÇÃO E COMPARAÇÃO DAS ESTIMATIVAS	124
	CONCLUSÃO	125
	AGRADECIMENTOS	125
	REFERÊNCIAS.....	126
CONCLUSÕES GERAIS	129
REFERÊNCIAS DA INTRODUÇÃO	130
GLOSSÁRIO	137
APÊNDICE	138

INTRODUÇÃO

PARASITISMO E SUAS RELAÇÕES

Existem diversas definições sobre o que é parasitismo, porém, uma das mais comumente aceitas afirma que é uma associação entre organismos de espécies distintas na qual se observa, além de associação íntima e duradoura, uma dependência metabólica de grau variado. Porém, ele é muito mais do que isso; é um dos modos de vida mais bem sucedidos entre os organismos vivos.

Levando-se em consideração uma definição ampla de parasitismo, aproximadamente 50% de todas as espécies animais conhecidas se encaixam nesta categoria e, mesmo quando se considera definições mais restritivas, o número de parasitas ainda é enorme, pois praticamente todo metazoário de vida livre abriga pelo menos uma espécie parasita (POULIN, 2000).

Uma vez que vivem em uma íntima relação de dependência com seus hospedeiros, os parasitos desenvolvem uma série de adaptações para assegurar sua permanência no hospedeiro, entre as quais figuram os mecanismos de evasão do sistema imune e os mecanismos de reconhecimento de hospedeiro. Na via contrária, os hospedeiros também se adaptam para resistir ao parasito e reduzir os danos causados por eles. A este processo denomina-se co-evolução e ele é de grande importância na relação parasito-hospedeiro, especialmente para as espécies de parasitos com elevada especificidade.

Porém, essa relação entre parasito e hospedeiro é influenciada não só pelo grau de co-evolução e de adaptações, mas também por diversos outros fatores, tanto de origem intrínseca aos organismos (como idade e nutrição), quanto de origem extrínseca ou ambiental (como temperatura e salinidade). Portanto, essa relação pode ser esquematizada, de uma forma mais ampla, como na Figura A.

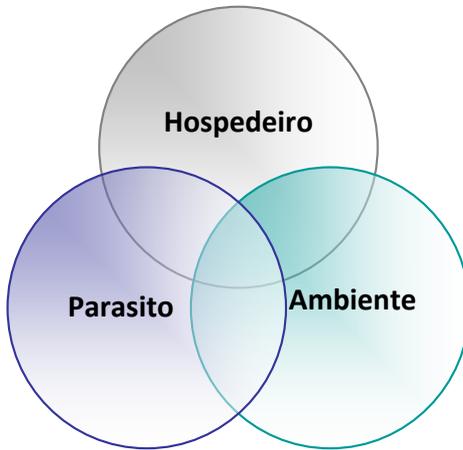


Figura A: Diagrama de Venn com relação hospedeiro-parasita-ambiente

Como foi apresentado, o parasitismo ocorre frequentemente na natureza onde os animais, em geral, apresentam uma grande variedade de parasitos sem, porém, que estes lhes causem danos significativos. Isto acontece devido à existência de um equilíbrio, entre o estado nutricional e fisiológico do animal e o meio ambiente em que está inserido, que impede a manifestação de doenças (ANDRADE *et al.*, 2001, MORAES & MARTINS, 2004; ROBERTS, 1981).

A manifestação de doenças ou de sintomas parasitários depende da interação entre os três conjuntos apresentados anteriormente (parasita, hospedeiro e ambiente). Os parasitos necessitam dos hospedeiros para sua sobrevivência. Em condições normais, num ambiente equilibrado, o número de parasitos por hospedeiro é baixo e o dano causado é mínimo. Entretanto, quando ocorre um desequilíbrio na condição de homeostasia (tanto do ambiente quanto do hospedeiro), que favoreça a proliferação e a infestação parasitária, verifica-se um aumento drástico na intensidade de parasitos por hospedeiro, aumentando substancialmente o dano sofrido pelos hospedeiros. Além disso, elevadas intensidades parasitárias, aliadas às condições ambientais inadequadas, podem causar eventos de mortalidade em massa.

De forma geral, pode-se classificar o processo de doença no hospedeiro da seguinte maneira (GUIDO, 2011):

- Ausência de luta = quando ocorre resistência ou insusceptibilidade do hospedeiro ao parasita; ou, ainda, estado de latência, onde o hospedeiro alberga o parasita, porém não existem condições para o desenvolvimento da doença.
- Reação do hospedeiro = quando existe uma luta entre o hospedeiro e o parasito, que, por sua vez, pode tomar dois caminhos:
 - Relacionamento harmônico = O parasita se instala e se multiplica sem causar danos graves ao hospedeiro – não ocorre sintomatologia.
 - Relacionamento desarmônico = O parasita se instala e se multiplica no hospedeiro provocando alterações orgânicas e funcionais capazes de interferir na homeostasia. Neste caso a reação do hospedeiro pode ter dois desfechos:
 - Incompatibilidade total = Morte do hospedeiro ou do parasita, sendo que, normalmente, a morte do hospedeiro leva à morte do parasita. Quando ocorre a vitória do hospedeiro, na maioria dos casos, não ocorre o desenvolvimento da imunidade ativa.
 - Incompatibilidade parcial = O hospedeiro é incapaz de destruir o parasito, de forma que ele (o parasito) persiste em seu organismo, sempre em relacionamento desarmônico. O hospedeiro pode desenvolver imunidade contra infecções posteriores do mesmo agente.

PARASITOLOGIA E AQUICULTURA

A aquicultura é uma atividade em franca expansão no mundo todo. Enquanto a produção pesqueira global mantém-se estática, a produção aquícola tem apresentando crescimento médio de 6% ao ano (FAO, 2010). Esse crescimento, porém, nem sempre é ordenado e a constante busca por produtividade leva à intensificação dos cultivos.

Em cultivos intensivos os animais, de modo geral, estão constantemente expostos a diversos fatores causadores de estresse, tais como: alta densidade de estocagem, baixa qualidade de água, alterações bruscas de temperatura, diminuição no teor de oxigênio dissolvido, manejo inadequado e alimentação deficiente em nutrientes, entre outros. (MARTINS, 1998; THATCHER & BRITES-NETO, 1994)

Essa exposição crônica a fatores de estresse acarreta um constante desequilíbrio homeostático, que favorece a proliferação de patógenos, especialmente em locais com altas densidades de estocagem, tornando os animais mais suscetíveis às doenças (CECCARELLI et al., 1990; MARTINS & ROMERO, 1996).

Animais doentes têm seu crescimento retardado, são menos atraentes ao consumidor e podem até ter sua comercialização proibida, causando prejuízos econômicos aos produtores (BARKER et al., 2002; RANZANI-PAIVA & SILVA-SOUZA, 2004). Além disso, ambientes com baixa qualidade de água tornam os animais muito mais vulneráveis, de modo que mesmo baixas infestações parasitárias podem matar (MOLNÁR, 1994), culminando em mortalidades massivas, que também causam significativos prejuízos econômicos (CECCARELLI et al., 1990; KLESIUS & ROGERS, 1995; MORAES & MARTINS, 2004; SNIESZKO, 1974).

Faz-se necessário, portanto, forte trabalho de acompanhamento e manejo sanitário em cultivos, como forma de prevenir o surgimento de doenças e evitar potenciais prejuízos. Neste contexto, a parasitologia - ciência que estuda os parasitos, seus hospedeiros e as relações entre eles - é de extrema importância para assegurar um acompanhamento sanitário adequado.

A parasitologia de organismos aquáticos é uma área ainda recente no Brasil, sendo bastante carente de profissionais especializados e de laboratórios equipados, especialmente no tocante a trabalhos de acompanhamento sanitário em cultivos. Há uma grande lacuna de conhecimento na detecção e quantificação de patógenos, por parte dos laboratórios nacionais, que necessita de fortes investimentos e estudos para ser ultrapassada.

Ademais, a falta de precisão e padronização na área ainda é um poderoso entrave para os estudos parasitológicos, uma vez que torna difícil a comparação de dados, impossibilitando o desenvolvimento de parâmetros de avaliação do grau de sanidade e, conseqüentemente, a formulação de estratégias adequadas para o controle sanitário em cultivos.

Agravando este quadro, um dos grupos de patógenos cujos estudos estão menos desenvolvidos na parasitologia de organismos aquáticos no Brasil é dos protozoários.

Na aquicultura mundial, os protozoários possuem lugar de destaque, pois são considerados responsáveis por grandes prejuízos econômicos, estando incluídos entre os principais agentes etiológicos causadores de mortalidade em cultivos.

Para exemplificar, podem ser citados alguns estudos que descrevem a mortalidade em massa causada por protozoários em áreas de cultivo, como os casos de: infestação por *Perkinsus marinus* em *Crassostrea virginica*, na costa leste americana (BURRESON & RAGONE-CALVO, 1996; CROSBY & ROBERTS, 1990); infestação por *Ichthyophthirius multifiliis* em *Oncorhynchus mykiss*, na Dinamarca (BUCHMANN & BRESCIANI, 1997); e a infestação por *Piscinoodinium pillulare* em tambacus, no estado de São Paulo (MARTINS et al., 2001).

No Brasil, os protozoários são importantes agentes etiológicos, em especial nas pisciculturas de água doce. Dentre os principais causadores de prejuízos pode-se citar: o dinoflagelado *Piscinoodinium pillulare*, o ciliado *Ichthyophthirius Multifiliis* e os ciliados do gênero *Trichodina*.

Piscinoodinium pillulare

O dinoflagelado *Piscinoodinium pillulare* (Schäperclaus, 1954) Lom, 1981 é um parasito obrigatório que se fixa às células do hospedeiro através de rizocistos (prolongamentos em forma de raiz), podendo causar necrose de células, pontos de hemorragia no tegumento, hiperplasia das brânquias e fusão das lamelas secundárias, o que resulta em graves dificuldades respiratórias e mortalidade (MARTINS et al., 2001; PAVANELLI et al., 2002). Ele é responsável por importantes danos em pisciculturas em diversos locais no mundo. No Sudeste brasileiro, foram relatadas infestações elevadas que chegaram a causar a morte de mais de 4000 peixes em 24 horas (MARTINS et al., 2001).

Tricodínídeos

Os protozoários ciliados do gênero *Trichodina* (Ehrenberg, 1830) estão entre os principais agentes etiológicos de morbimortalidade em tilápias cultivadas no Brasil, além de serem relatados como parasitos encontrados em cultivos de diversos peixes de importância comercial (AZEVEDO et al., 2006; GHIRALDELLI et al., 2006; MARTINS et al., 2002; TAVARES-DIAS et al., 2001; VARGAS et al., 2000). Eles são ectoparasitos oportunistas, com baixa especificidade de hospedeiro, sendo frequentemente encontrados parasitando peixes marinhos e de água doce (ARTHUR & LOM, 1984; XU et al., 2001). Sua patogenicidade varia de acordo com a imunidade dos peixes, mas podem causar sérios danos, destruindo o epitélio do peixe com sua movimentação rotatória e abrindo portas de entrada para infecções secundárias (AL-RASHEID, 2000; MANCINI et al., 2000). Parasitam a superfície do corpo, nadadeiras e brânquias dos animais, reproduzindo-se rapidamente através de fissão binária (MARTINS et al., 2002; MANCINI et al., 2000).

Ichthyophthirius multifiliis

O ciliado *Ichthyophthirius Multifiliis* (Fouquet, 1876) é um ectoparasito de baixa especificidade, causador da chamada “doença dos pontos brancos”, responsável por perdas na produção de peixes ornamentais e larviculturas (MARTINS *et al.*, 2002; NOGA, 2000; THILAKARATNE *et al.*, 2003). Seu ciclo de vida é composto por três estágios distintos; na fase de trofozoíto/trofote, o parasito fixa-se no epitélio da superfície corporal ou nas brânquias do peixe, formando cistos repletos de indivíduos que quando alcançam cerca de 1 mm se rompem, liberando na água milhares de protozoários na fase de tomonete. Nesta fase, aderem-se ao substrato onde se desenvolvem para posteriormente retornar à coluna d’água sob a forma de terontes infectantes (BUCHMANN *et al.*, 2001). Este parasito pode causar danos significativos, especialmente nas brânquias onde provoca hiperplasia, fusão das lamelas secundárias, edemas e, até, necroses (FERRAZ DE LIMA *et al.*, 1991; MARTINS & ROMERO, 1996).

MÉTODOS UTILIZADOS NA QUANTIFICAÇÃO DE PROTOZOÁRIOS PARASITOS

A quantificação de protozoários é uma tarefa complexa, que, mesmo nos dias de hoje, ainda não foi completamente resolvida. Existe uma vasta gama de métodos de detecção e identificação de protozoários, porém a quantificação precisa de alguns organismos alvos, em uma amostra complexa, continua sendo um desafio metodológico (DAIMS & WAGNER, 2007). Atualmente, é possível dividir as principais abordagens utilizadas no mundo para quantificação de protozoários em: métodos moleculares; métodos estereológicos; métodos de análise de imagens; e métodos de cultura.

Abordagem molecular

Os métodos moleculares baseiam-se, principalmente, na utilização da técnica de amplificação da reação em cadeia da polimerase (PCR). Inicialmente possibilitavam apenas análises qualitativas, sendo utilizados somente na identificação e detecção dos protozoários, que era feita em função da presença ou não de sequências específicas de DNA. Porém, com o desenvolvimento do PCR em tempo real (Real-Time PCR) tornou-se possível fazer, também, análises quantitativas.

Desenvolvida no começo da década de 1990 (HIGUCHI et al., 1992), a técnica de PCR em tempo real explora a relação proporcional entre o ciclo limítrofe, onde é detectada amplificação exponencial, e a quantidade de cópias de um fragmento alvo de ácido nucléico (MONIS et al, 2005). Assim, a amplificação de uma sequência padrão de DNA, juntamente com a sequência da amostra analisada, permite a construção de uma curva padrão de amplificação, que possibilita estimar o número inicial de cópias da sequência de interesse (EDWARDS et al., 2004).

O método de quantificação por PCR em tempo real é considerado bastante confiável e preciso, contanto que tenham sido desenvolvidos marcadores moleculares eficientes para a espécie estudada. Porém, ele é considerado presuntivo, pois detecta uma sequência de DNA e não, necessariamente, um patógeno viável. Além disso, nem sempre o número de cópias de determinada sequência pode ser convertido em número de células, seja porque o genoma do animal ainda é desconhecido (o que ocorre na maioria dos casos) ou porque o número de cópias da sequência no genoma varia conforme a fase de crescimento. Desse modo, não é possível dizer quantas vezes a sequência se repete por indivíduo (LUDWIG & SCHLEIFER, 2000).

Abordagem estereológica

Estereologia é uma área interdisciplinar bastante complexa, que remonta aos séculos XVIII e XIX, quando G. Buffon e A. Delesse começaram a utilizar a união de princípios da geometria (como o princípio de Cavalieri) e da estatística, na tentativa de tentar estimar volumes tridimensionais a partir de seções planas.

Seu nome origina-se da palavra grega para “sólido”, e foi definida, originalmente, como “a interpretação espacial das seções”. Sua versão moderna, entretanto, pode ser definida como “um conjunto de processos que objetivam quantificar as propriedades geométricas de objetos alvos sem fazer suposições sobre suas características inerentes” (REMACHA-TRIVIÑO et al., 2008).

Suas quantificações baseiam-se na análise da superfície de amostras, obtidas por meio de cortes histológicos, contra um sistema teste (que pode ter diversas formas e seguir diversos princípios) de modo aleatório e sistematizado, com o intuito de obter extrapolações não tendenciosas.

Apesar de muito antiga, a estereologia é um campo em constante evolução que, recentemente, experimentou um novo despertar, especialmente na área de neurociência, através do surgimento de novas técnicas, como as de dissector óptico (GUNDERSEN, 1986) e fracionador óptico (WEST et al., 1991), e pela fabricação de novos e sofisticados equipamentos automatizados, que utilizam tecnologia de ponta para garantir quantificações imparciais e eficientes (MAYHEW & GUNDERSEN, 1996; REMACHA-TRIVIÑO et al., 2008).

Mesmo com a evolução destes recursos, ainda é possível fazer as quantificações de forma manual, com o auxílio de acessórios e calibrações especiais para os microscópios.

Abordagem de análise de imagens

Muito recentes, os métodos de análise de imagens tem como base o uso de marcadores fluorescentes, que se ligam às células dos microorganismos, para visualizar e quantificar células com o auxílio de ferramentas de análise de imagens.

Esses marcadores fluorescentes podem se ligar ao DNA, como é o caso do brometo de etídio (RAMSING et al. 1996) e do fluorocromo SYBR Green (SCHMID et al., 2000), ou à antígenos específicos, como no método de imunofluorescência, que usa anticorpos ligados à corantes marcadores. No caso dos marcadores ligantes ao DNA, eles podem fazê-

lo de forma generalista, incapaz de quantificar uma população específica, ou utilizando técnicas sofisticadas, como a *fluorescence in situ hybridization* (FISH), que consegue fazer detecções múltiplas e espécie – específicas (DAIMS & WAGNER, 2007).

Após a marcação, as células podem, então, ser quantificadas com bastante precisão e confiabilidade, tanto através de contagem visuais, utilizando microscópios de epifluorescência, quanto através de processos automatizados, por meio de microscópios motorizados controlados por computador ou citômetros de fluxo.

Abordagem de cultura

Os métodos de cultura são amplamente utilizados na microbiologia e têm como principal fundamento a avaliação da proliferação da espécie estudada, especialmente em condições *in vitro*. Geralmente são coletadas amostras de tecido do hospedeiro, que são incubadas em um meio de cultura específico, para serem, posteriormente, avaliadas.

Muitas vezes, se expressa o resultado da análise em classes de intensidade, porém existem técnicas que permitem a quantificação. Este tipo de método é bastante utilizado na detecção de *Perkinsus marinus*, que é feita através de cultura no meio fluído de tioglicolato de Ray (RAY, 1952; RAY, 1966).

Entretanto, a utilização de métodos de quantificação por cultura geralmente não são adequados para análise de comunidades (DAIMS & WAGNER, 2007) e seus resultados também não correspondem necessariamente aos valores reais de parasitismo (REMACHA-TRIVIÑO et al., 2008). Além disso, a maioria dos microorganismos na natureza ainda não podem ser cultivados (AMANN et al., 1995), como é o caso de muitos protozoários parasitos de peixes.

Quantificação de protozoários ictioparasitos no Brasil

Ao contrário do que ocorre em muitos lugares dos Estados Unidos e da Europa, as técnicas de quantificação de protozoários parasitos no Brasil ainda são bem pouco desenvolvidas.

A maioria dos laboratórios que trabalham com parasitologia de peixes não tem condições de realizar quantificações por meio de métodos moleculares, de análise de imagens e, até mesmo, estereológicos, por falta de equipamentos ou de mão de obra especializada. Além disso, os métodos de cultura estão muito pouco desenvolvidos no que se refere à protozoários parasitos em peixes.

Por causa disso, houve a necessidade de se improvisar e desenvolver uma estratégia diferente para quantificar os protozoários presentes em brânquias de peixes. Surgindo, assim, o método comumente utilizado para quantificação dos protozoários presentes nas brânquias de peixes no Brasil, que se baseia na contagem, com o auxílio de uma câmara de McMaster, dos protozoários presentes em uma amostra, cujo volume de solução de formalina contendo brânquias cuidadosamente raspadas é conhecido (TAVARES-DIAS *et al.*, 2001). O valor de protozoários obtido na contagem da amostra com brânquias é, então, extrapolado para o peixe, em função da diferença entre o volume de solução na amostra e o volume total da solução de formalina contendo as brânquias (GHIRALDELLI *et al.*, 2006; TAVARES-DIAS *et al.*, 2001).

Entretanto, este método estimativo possui diversas restrições quanto a sua precisão. Inicialmente pelo fato de que não há uma padronização nacional do volume utilizado para a contagem e, principalmente, pelo fato de a estimativa basear-se no volume total da solução com as brânquias, que, além de também não ser padronizado, não leva em conta diferenças de tamanho entre brânquias, não evidenciando, portanto, uma real relação com a superfície parasitada.

Por estes motivos, torna-se de extrema importância o desenvolvimento e a padronização de métodos mais precisos e confiáveis, para detecção e quantificação destes protozoários.

Abordagem proposta

Levando em consideração as dificuldades de aplicação dos métodos descritos anteriormente e buscando uma alternativa viável, precisa e confiável para a quantificação de protozoários em nossa realidade, esta Dissertação propõe a utilização de um método onde os resultados quantitativos se baseiem em extrapolações de contagens amostrais, porém, em função de parâmetros de extrapolação. Parâmetros estes que sejam obtidos através de modelos matemáticos e que possuam uma relação direta com a quantidade de brânquia disponível para ser parasitada.

Dessa forma, não será necessária a utilização de equipamentos de custo elevado, já que muitas vezes as coletas e análises parasitológicas de cultivos são realizadas em condições onde, mesmo equipamentos simples, como balanças de precisão, não estão disponíveis.

ESPÉCIE ALVO

Para estabelecer o modelo matemático proposto nesta dissertação, a espécie escolhida foi a tilápia do Nilo, *Oreochromis niloticus* (Linnaeus, 1758), devido à sua grande importância na piscicultura de água doce mundial.

A tilápia do Nilo é um dos peixes mais cultivados na aquicultura mundial (CAVICHIOLO *et al.*, 2002), chegando a mais de 2 milhões de toneladas produzidas ao redor do mundo (FAO, 2009). No Brasil, sua produção alcança 70 mil toneladas anuais e no Estado de Santa Catarina ela corresponde a 47% do total de peixes cultivados (BOSCARDIN, 2008; EPAGRI, 2008).

Suas diversas qualidades zootécnicas, como rápido crescimento, grande rusticidade e boa adaptabilidade renderam-lhe o apelido de “frango d’água” (LITTLE *et al.*, 1994). Além disso, a espécie possui uma incrível capacidade de tolerar diversas condições de estresse, tais como baixa qualidade de água, baixas concentrações de oxigênio, grandes variações de pH e altos valores de salinidades (GALLI, TORLONI, 1986; ZANIBONI FILHO, 2004)

Originária do continente Africano, ela é exótica no Brasil, porém, devido ao seu hábito alimentar fitoplanctófago, sua fácil adaptação ao clima brasileiro e sua fácil reprodução, ela atualmente encontra-se amplamente distribuída pelas bacias hidrográficas do Brasil.

COMPREENDENDO A MODELAGEM MATEMÁTICA

Modelagem matemática é a área do conhecimento que estuda a simulação de sistemas reais a fim de prever o comportamento dos mesmos, sendo empregada em diversos campos de estudo, tais como física, química, biologia, economia e engenharia. Um modelo matemático pode ter várias formas, mas, geralmente, pode ser classificado nas seguintes categorias:

- **Linear ou não linear** - A questão da linearidade é dependente do contexto. Geralmente assume-se que o modelo é linear quando: todos os operadores em um modelo exibem linearidade; em um modelo estatístico, os parâmetros estimados apresentam relação linear; ou, ainda, quando as restrições e funções objetivas em um modelo de programação matemática são representadas somente por equações lineares. Quando os pressupostos de linearidade não podem ser cumpridos, assume-se a não linearidade.

- **Determinístico ou probabilístico (estocástico)** - Considera-se determinístico o modelo no qual o estado de um sistema é definido por causas que podem ser descritas adequadamente sem recorrer a elementos probabilísticos, onde os resultados são determinados de forma exata a partir das condições iniciais. Ao passo que no modelo probabilístico existe aleatoriedade, sendo os estados das variáveis definidos por distribuições de probabilidade.
- **Estático ou dinâmico** - Um modelo estático não leva em consideração o tempo, ao passo que o modelo dinâmico o faz.
- **Discreto ou contínuo** - Modelos discretos são aqueles cujas variáveis possuem distribuições de probabilidade discretas, enquanto modelos contínuos possuem variáveis com distribuições de probabilidade contínuas.

Para melhor compreensão das ferramentas de modelagem e das técnicas aplicadas na modelagem, alguns conceitos não triviais devem ser esclarecidos, entre eles os de:

Bootstrap

Bootstrapping é um termo de origem inglesa, que se originou na década de 1880 como um acessório para ajudar a calçar botas e, gradualmente, adquiriu uma coleção de significados metafóricos adicionais. O tema comum a todos esses significados é a realização de um processo sem ajuda externa, mas com etapas de facilitação interna.

Na estatística é um método computacional utilizado para designar medidas de exatidão às estimativas amostrais (EFRON & TIBSHIRANI, 1994). Sua idéia básica consiste na re-amostragem de um conjunto de dados, diretamente ou via um modelo ajustado, a fim de criar réplicas dos dados e a partir das quais se pode avaliar a variabilidade de quantidades de interesse, sem usar cálculos analíticos. Pode ser aplicado quando existe, ou não, um modelo probabilístico bem definido para os dados e em situações paramétricas ou não paramétricas.

Método dos mínimos quadrados (MMQ)

O Método dos Mínimos Quadrados é uma técnica de otimização matemática, que busca encontrar o melhor ajuste para um conjunto de dados através do processo de minimizar a soma dos quadrados das diferenças entre os valores estimados e os dados observados (resíduos).

É amplamente aplicado em modelos de regressões, onde se tenta minimizar a soma dos quadrados dos resíduos da regressão, de forma a maximizar o grau de ajuste do modelo aos dados observados.

Aprendizagem de máquina (Machine learning)

A aprendizagem de máquina é um sub-campo da inteligência artificial dedicado ao desenvolvimento de algoritmos e técnicas que permitam ao computador aprender, isto é, que permitam ao computador aperfeiçoar seu desempenho em alguma tarefa, baseando-se em um raciocínio indutivo, que extrai regras e padrões a partir de conjuntos de dados empíricos.

O principal foco da aprendizagem de máquina é desenvolver sistemas capazes de reconhecer automaticamente padrões complexos e fazer decisões inteligentes baseadas nos dados. O objetivo principal de um sistema de aprendizagem é generalizar a partir de sua experiência, para isso ele usa como base um conjunto de dados de treinamento, que são usados para extrair padrões de distribuição probabilística que permitam a ele produzir respostas úteis para novos conjuntos de dados.

Conjunto de treinamento e conjunto de teste

No campo de aprendizagem de máquina, muitas vezes os dados são divididos em conjuntos menores para que sejam realizados diversos tipos de análises e processos. Na nomenclatura utilizada, conjunto de treinamento é definido como o conjunto de dados utilizado para o processo de aprendizagem, enquanto o conjunto teste é definido como o conjunto utilizado para testes de validação e comparação.

Qualidade de ajuste (Goodness of fit)

A qualidade de ajuste (*Goodness of fit*) ou aderência de um modelo estatístico descreve o quão bem ele se ajusta a um conjunto de observações. Medidas de aderência costumam resumir a discrepância entre os valores estimados (esperados) pelo modelo e os valores observados, podendo ser usadas em testes de hipótese.

Sobre-ajuste (overfitting)

Na estatística, o sobre-ajuste ocorre quando um modelo descreve erros aleatórios ou ruídos ao invés de uma real relação entre os dados (STATSOFT, 2011). O sobre-ajuste geralmente aparece em modelos

muito complexos que, ao invés de aprender a generalizar a partir de um conjunto de dados de treinamento, aprendem a imitar aquele conjunto de dados de treinamento. Isso resulta em modelos com baixa capacidade preditiva, visto que o modelo se ajusta muito bem ao conjunto de dados de origem, mas tem pouca flexibilidade para conjuntos de dados desconhecidos.

Poda (pruning)

A poda é uma técnica dentro da aprendizagem computacional que reduz os tamanhos de árvores de decisão, através da remoção de seções com pouco poder classificatório. A idéia principal da poda é melhorar a exatidão preditiva, eliminando dados errôneos e reduzindo o sobreajuste, ao mesmo tempo em que produz estruturas menos complexas e, conseqüentemente, de entendimento mais fácil (STATSOFT, 2011).

Validação cruzada (cross-validation)

Validação cruzada é uma técnica para verificar a forma como os resultados de uma análise estatística irão generalizar para um conjunto de dados independentes (STATSOFT, 2011), sendo usada principalmente para estimar a exatidão de um modelo preditivo na prática. Ela segue o mesmo princípio de dividir os dados em conjunto de treinamento e conjunto de teste, sendo este processo repetido diversas vezes usando diferentes partições, para reduzir a variabilidade.

Erro de validação cruzada generalizada (GCV error)

É uma medida de aderência que leva em consideração não só o erro residual, mas também a complexidade do modelo, através de uma penalidade c (cujo valor geralmente está entre 2 e 3) para adição de novas funções base (STATSOFT, 2011).

Limiar (threshold)

Na estatística, limiar é o valor mínimo de alguma quantidade, seja para definir níveis de aceitação/rejeição ou para estabelecer valores que devem ser atingidos. Em modelos matemáticos, os limiares são normalmente utilizados como limitantes de precisão, com o objetivo de reduzir a complexidade e evitar sobreajuste.

FERRAMENTAS DE MODELAGEM COMPUTACIONAL

Com o desenvolvimento da informática, a modelagem computacional experimentou um grande salto nas últimas décadas e hoje em dia é cada vez mais comum o uso de modelos de aprendizagem de máquina, bem como modelos de regressão, nas mais diversas áreas do conhecimento. Esses modelos, que até pouco tempo atrás demandavam uma capacidade de processamento raramente disponível, hoje estão presentes em muitas de nossas tarefas cotidianas, e processos como *data mining* tem ganhado espaço tanto no meio acadêmico como no corporativo.

Muito desse crescimento se deve à capacidade que os modelos apresentam de estimar pontos de funções a partir de um conjunto mínimo de informações, mesmo quando o fator de complexidade técnica e outras causas de influência são desconhecidos (SIKKA et al., 2010).

Existem várias técnicas dentro da modelagem computacional, com diferentes tipos de algoritmos e variadas abordagens, entre as quais se podem citar:

Redes Neurais Artificiais (Artificial Neural Networks - ANN)

Redes neurais são sistemas computacionais inspirados pela estrutura e/ou aspectos funcionais das redes neurais biológicas. Sua origem teórica remonta à década de 40 do século XX e sua estrutura é composta por “neurônios” artificiais (unidades de processamento) que são interligados em forma de rede, seguindo os seguintes princípios (STATSOFT, 2011):

- Cada “neurônio” recebe um número de dados de entrada (seja dos dados originais ou dos dados de saída de outros “neurônios” da rede);
- Cada entrada de dados vem através de uma conexão que tem uma força (peso); que corresponde à eficiência sináptica de um neurônio biológico;
- Cada neurônio tem um único valor de limiar;
- É feita a soma ponderada dos dados de entrada e os limiares são subtraídos para formar o sinal de ativação do “neurônio” (também conhecido como potencial pós-sináptico);
- O sinal de ativação passa por uma função de ativação para gerar os dados de saída do “neurônio”.

As redes neurais artificiais vêm sendo utilizadas com sucesso em uma gama extraordinária de problemas, desde problemas de classificação e controle, até reconhecimento de padrões e modelagem de relações complexas entre dados de entrada e saída. As chaves de seu sucesso estão em seu poder de modelar funções extremamente complexas, sem as restrições da linearidade, e sua facilidade de utilização, devido ao seu apelo intuitivo.

K-nearest neighbors (k-NN)

O algoritmo utilizado na técnica de *k-nearest neighbors* está entre os mais simples da aprendizagem de máquina. Seu funcionamento baseia-se na classificação de um objeto (ou do valor de um objeto) através da votação majoritária de seus vizinhos, ou seja, o objeto (ou valor) é designado como pertencente à classe (ou valor) mais comum entre seus K vizinhos mais próximos (sendo K um número inteiro positivo, geralmente pequeno). Outra abordagem utiliza K igual a todo o conjunto amostral e atribui, aos votos, pesos de acordo com a distância Euclidiana (no caso de variáveis contínuas), seguindo o pressuposto intuitivo de que objetos mais próximos são potencialmente similares.

Nesta técnica de modelagem a escolha do valor de K é essencial para a qualidade do modelo; valores muito baixos de K geram variâncias elevadas, ao passo que valores muito altos de K causam tendenciosidade (STATSOFT, 2011). Portanto, para auxiliar na escolha do valor ótimo de K utiliza-se a técnica de validação cruzada.

Support vector machine (SVM)

Support vector machines é um conjunto de métodos de aprendizagem baseados no conceito de “planos de decisão”, ou seja, planos que separam conjuntos de objetos com diferentes classes (STATSOFT, 2011). O algoritmo original de SVM foi inventado por Vladimir Vapnik e era utilizado apenas para a tarefa de classificação binária e não probabilística. Atualmente ele sofreu diversas modificações e já conta, também, com um método para regressões.

Uma modelagem *Support vector machines* constrói um hiperplano vetorial (ou um conjunto de hiperplanos) em um grande espaço dimensional (ou infinito), ou seja, ela delimita espaços multidimensionais que podem assumir qualquer forma possível de ser descrita por vetores. Esses hiperplanos são utilizados para tarefas de classificação e regressão, separando conjuntos de dados, de forma que,

quanto maior a distância de um hiperplano para o ponto externo mais próximo, menor é o erro de generalização.

Multivariate Adaptive Regression Splines (Marsplines)

Multivariate adaptive regression splines (FRIEDMAN, 1991) é uma forma de análise de regressão não paramétrica, que pode ser considerada uma extensão dos modelos lineares, com o diferencial de automaticamente modelar não-linearidades e interações.

Assim como os modelos lineares, ela tem como objetivo principal prever valores de uma variável contínua dependente a partir de um conjunto de variáveis preditoras (independentes). Ela, porém, não faz nenhuma suposição ou imposição sobre a relação funcional (linear, logística, etc.) entre as variáveis dependentes e independentes. Ao invés disso, a Marsplines constrói essa relação a partir de coeficientes e funções base (originadas dos próprios dados), dividindo o conjunto de dados de entrada em partições, cada uma com sua própria equação de regressão.

Dessa forma, o algoritmo Marsplines funciona como múltiplas regressões lineares onde são estimados pontos de ruptura (nós), que definem uma “região de aplicação” para determinada equação linear simples. Além disso, nele são aplicadas técnicas de ajuste e correção, como método dos mínimos quadrados, a validação cruzada e a poda.

Em resumo, a implementação de um modelo Marsplines segue os seguintes passos (STATSOFT, 2011):

1. Começa-se com o modelo mais simples, envolvendo somente uma função base constante;
2. Procuram-se funções base para cada possível nó em cada uma das variáveis, sendo adicionadas ao modelo aquelas que minimizem o erro preditivo (com base nos mínimos quadrados);
3. O passo dois é repetido até que se atinja o limiar (valor estabelecido *a priori* pelo operador);
4. Aplica-se um processo de poda, removendo as funções base que menos contribuírem para a aderência dos dados (com base nos mínimos quadrados).

De forma mais simplificada, pode-se, ainda, dividir o processo em duas etapas: *forward pass*, onde o modelo ganha complexidade, e *backward pass*, onde é feita a poda para reduzir o sobre-ajuste.

OBJETIVOS

Ao desenvolver este trabalho científico buscou-se contribuir para o desenvolvimento e aprimoramento dos métodos de quantificação de protozoários em peixes. O estudo apresentado na dissertação procurou desenvolver ferramentas capazes de estimativas confiáveis e precisas, mesmo para laboratórios em condições de precariedade de equipamentos e mão de obra qualificada. Assim, para desenvolver o estudo proposto foram estabelecidos os seguintes objetivos:

Objetivo Geral

Desenvolver uma ferramenta para quantificação de protozoários parasitos em brânquias de peixes, tendo a tilápia do Nilo como espécie alvo.

Objetivos Específicos

- Padronizar os protocolos de coleta de amostras para quantificação;
- Padronizar os protocolos de quantificação de protozoários;
- Desenvolver um método que possibilite inferir sobre a quantidade total de protozoários parasitos presentes nas brânquias, a partir de contagens feitas em amostras, utilizando como parâmetro de extrapolação o peso seco de tecido branquial;
- Desenvolver um modelo preditivo para estimar variáveis utilizadas como parâmetro de extrapolação no método;
- Desenvolver um software para auxiliar na aplicação do método de quantificação.

CAPÍTULO 1

DESENVOLVIMENTO DE UM MODELO PREDITIVO PARA ESTIMAR PARÂMETROS UTILIZADOS NA QUANTIFICAÇÃO DE PROTOZOÁRIOS PARASITOS DE PEIXES

ABSTRACT

A predictive model was developed using Multivariate Adaptive Regression Splines (MARSplines) to estimate “gill wet weight”, “gill tissue dry weight” and “gill tissue dry weight on a 1,5 ml sample”, having fish weight and fish width as predictors. For its development, 129 Nile Tilapia were anesthetized, biometrically measured and photographed. Then they were euthanized and had their gills removed, weighed and stored in numbered cups with 200 ml of 10% formalin solution. Data for the following variables were then obtained from the 129 sample fish: fish weight; fish’s length, height, and relative surface area, obtained by image analysis; gills wet weight, after one week immersion in formalin solution; gill’s tissue dry weight; gill’s tissue dry weight on sample, obtained from the mean of five 1,5 ml homogenized samples. Correlation indexes were verified by multiple regressions and by bootstrapping of 10.000 samples. Fish weight and fish length were chosen as predictor variables due their measuring practicality. Data were then submitted to MARSplines using STATISTICA 8.0 and the model equations for “gills wet weight”, “gill’s tissue dry weight” and “gill’s tissue dry weight on a 1,5 ml sample” were obtained. MARSplines modeling results had their goodness of fit verified by R^2 and residuals analysis. The statistical significance of the estimated parameters was analyzed by pairwise T test. Both goodness of fit and statistical significance of estimated parameters analyses sustain that MARSplines modeling results were robust, providing conditions to develop a software for protozoan parasites quantification on fish gills.

Keywords: Aquaculture. Math Modeling. Multivariate Adaptive Regression Splines (MARSplines). Nile tilapia. Protozoa.

RESUMO

Este trabalho buscou desenvolver um modelo preditivo, com o objetivo de auxiliar na quantificação de protozoários parasitos de peixes. Para isto, 129 tilápias do Nilo foram anestesiadas, avaliadas biometricamente e fotografadas. Depois de eutanasiadas, tiveram suas brânquias removidas, pesadas e armazenadas em 200 ml de solução formalina 10%. Foram aferidos os dados das seguintes variáveis: peso, comprimento, altura, área superficial relativa, através de análise das imagens; peso da brânquia hidratada, após uma semana de imersão em solução formalina; peso seco de tecido branquial; e peso seco de tecido branquial em uma amostra de 1,5 ml, obtido da média de cinco amostras homogeneizadas. Os índices de correlação foram verificados por meio de regressão múltipla e de *bootstrap* com 10.000 amostras. O peso e comprimento do peixe foram escolhidos como variáveis predictoras. Os dados foram submetidos à MARSplines, utilizando STATISTICA 8.0 (Statsoft), e os resultados da modelagem passaram por análise dos resíduos e de aderência. A significância estatística dos parâmetros estimados foi verificada por meio de Teste T de Student para dados pareados. Ambas as análises sustentam que os resultados da modelagem feita com MARSplines são bastante robustos. Os protocolos de coleta e aferição de variáveis provaram ser eficientes e o modelo preditivo subsidiou o posterior desenvolvimento de um software para estimar a quantidade total de protozoários nas brânquias de peixes, auxiliando na expansão, padronização, precisão e confiabilidade das análises quantitativas feitas por laboratórios nacionais de parasitologia de peixes.

Palavras-chave: Aqüicultura. Modelagem matemática. Protozoários. *Multivariate Adaptive Regression Splines* (MARSplines). Tilápia do Nilo.

INTRODUÇÃO

A eficiência do acompanhamento sanitário de cultivos aquícolas, a nível nacional, é definida pela conjunção de três fatores: profissionais qualificados, laboratórios adequadamente equipados e métodos práticos e confiáveis para detecção e quantificação de patógenos.

Tanto a formação de profissionais qualificados quanto a estruturação adequada dos laboratórios dependem de políticas públicas e de incentivos privados ou governamentais, que estão fora do escopo deste artigo. Contudo, através dele é possível contribuir para o desenvolvimento de métodos viáveis para detecção e quantificação de patógenos. É neste contexto que este trabalho se presta ao intuito de tornar mais precisa e confiável a quantificação de protozoários em peixes, através do desenvolvimento, aprimoramento e padronização de métodos para sua quantificação.

Ao contrário do que ocorre em muitos lugares dos Estados Unidos e da Europa, o controle sanitário de protozoários parasitos de peixes no Brasil ainda é bastante incipiente (TAVARES-DIAS *et al.*, 2009 a; LEONARDO *et al.*, 2006). A maioria dos trabalhos ainda são exclusivamente qualitativos (MARTINS *et al.*, 2002; RANZANI-PAIVA *et al.*, 2005; EIRAS *et al.*, 2008; EIRAS *et al.*, 2007), avaliando somente a presença do patógeno, enquanto os que fazem alguma avaliação quantitativa utilizam parâmetros e métodos de quantificação diferentes, tornando difícil a comparação de dados (TAVARES-DIAS *et al.*, 2009 b; TAVARES-DIAS *et al.*, 2001; TAVARES-DIAS *et al.*, 2000; GHIRALDELLI *et al.*, 2006; JERÔNIMO *et al.*, 2011, BRACCINI *et al.*, 2007; BACHMANN *et al.*, 2007).

A quantificação precisa de protozoários ainda apresenta desafios metodológicos (DAIMS & WAGNER, 2007), especialmente em uma amostra complexa como a brânquia de um peixe, onde se podem encontrar protozoários parasitos fixados ao tecido ou em suspensão. Porém, é de grande importância para o estudo da sanidade em organismos aquíticos, sendo necessária para o estabelecimento de parâmetros de avaliação do grau de enfermidade.

Atualmente, podem-se dividir as principais abordagens utilizadas na quantificação de protozoários ao redor do mundo em: métodos moleculares; métodos estereológicos; métodos de análise de imagens e métodos de cultura.

Os métodos moleculares baseiam-se principalmente na utilização da técnica de amplificação da reação em cadeia da polimerase (PCR). Inicialmente, eram utilizados somente na identificação e detecção dos

protozoários em função da presença de sequências específicas de DNA, porém, com o desenvolvimento do PCR em tempo real (*Real-Time PCR*) (HIGUCHI et al., 1992) tornou-se possível fazer, também, a quantificação dos protozoários em função do número de cópias amplificadas de uma sequência específica de DNA (MONIS et al., 2005).

O método de quantificação por PCR em tempo real é considerado bastante confiável e preciso, porém, nem sempre o número de cópias pode ser convertido em número de células (LUDWIG & SCHLEIFER, 2000). Além disso, ele também apresenta desvantagens como seu custo, que é extremamente proibitivo para a maioria dos laboratórios de parasitologia aquícola no País, e a necessidade de marcadores moleculares bem desenvolvidos e eficientes para a espécie estudada.

Os métodos estereológicos baseiam-se na análise da superfície de amostras, obtidas por meio de cortes histológicos, contra um sistema teste (que pode ter diversas formas). A estereologia é uma área bastante complexa, fundamentada em diversos princípios de geometria e estatística, e muito discutida ainda, pois apresenta diversas abordagens para os problemas da estimação de quantidades e do cálculo de volumes (MAYHEW & GUNDERSEN, 1996). Sua versão moderna utiliza equipamentos com tecnologia de ponta para fazer uma análise tridimensional do objeto de estudo, o que garante quantificações imparciais e altamente eficientes, sendo muito usados na área de neurociência (REMACHA-TRIVIÑO et al., 2008; MAYHEW & GUNDERSEN, 1996)

O método de quantificação por estereologia é uma alternativa bastante interessante, cujo custo varia de muito elevado à relativamente baixo. Infelizmente, ela necessita, invariavelmente, de equipamentos e calibrações especiais e demanda profissionais especializados para uma quantificação adequada. Estas condições ainda não fazem parte da realidade da maioria dos laboratórios que trabalham com sanidade de organismos aquáticos no Brasil.

Os métodos de análise de imagens baseiam-se no uso de marcadores fluorescentes, que se ligam às células dos microorganismos, para visualizar e quantificar células com o auxílio de ferramentas de análise de imagens.

Os marcadores se ligam ao DNA (RAMSING et al., 1996) ou a antígenos bastante específicos, como na técnica de FISH (DAIMS & WAGNER, 2007), de forma que, após a marcação, as células podem ser quantificadas com precisão através de contagem visuais ou através de

processos automatizados, utilizando microscópios motorizados ou citômetros de fluxo.

Os métodos de análise de imagens são, portanto, bastante poderosos, combinando técnicas moleculares e imunológicas com equipamentos de última tecnologia para fazer quantificações precisas e específicas. Entretanto, o custo dos equipamentos necessários, bem como dos corantes utilizados, tornam sua utilização, na maioria dos laboratórios brasileiros, extremamente proibitiva para a atual realidade da parasitologia de peixes nacional, especialmente em aplicações rotineiras.

Os métodos de cultura, por sua vez, baseiam-se na avaliação da proliferação da espécie alvo de estudo, especialmente *in vitro*. Geralmente, as amostras de tecido do hospedeiro são coletadas e incubadas em um meio de cultura, para que sejam, posteriormente, avaliadas. Muitas vezes, o resultado da análise é expresso em classes de intensidade, porém existem técnicas que permitem a quantificação. Um exemplo desse tipo de método é o meio fluído de tioglicolato de Ray (RAY, 1952; 1966), usado na detecção e quantificação de *Perkinsus marinus*.

Entretanto, a utilização de métodos de quantificação por cultura geralmente não é adequada para análise de comunidades (DAIMS & WAGNER, 2007), além disso, a maioria dos microorganismos na natureza ainda não podem ser cultivados (AMANN et al., 1995), como é o caso de muitos protozoários parasitos de peixes. Ademais, este tipo de método apresenta como desvantagens a necessidade de um meio de cultura específico, a demanda de tempo para incubação e o fato de a quantificação não corresponder, necessariamente, ao parasitismo real (REMACHA-TRIVIÑO et al., 2008).

No Brasil, os laboratórios de parasitologia de peixes que trabalham com acompanhamento sanitário de cultivos ainda são em pequeno número e carentes de profissionais e equipamentos. Boa parte dos laboratórios não tem uma estrutura adequada, sendo raros os laboratórios que podem contar com termocicladores, microscópios de epifluorescência, ou equipamentos para estereologia computadorizada. A situação é ainda mais crítica quando são avaliadas as condições das coletas em campo que, muitas vezes, ocorrem em locais afastados e em condições precárias, onde até mesmo balanças de precisão são equipamentos raros.

Dessa forma, um dos métodos comumente utilizados para quantificação de protozoários parasitas de brânquias de peixes baseia-se na utilização de uma câmara de McMaster para contar os protozoários

em um determinado volume de solução contendo brânquias cuidadosamente raspadas, para posteriormente extrapolar-se a quantidade de protozoários do volume quantificado para o volume total. (TAVARES-DIAS *et al.*, 2001; GHIRALDELLI *et al.*, 2006). Este método estimativo possui, entretanto, sérias restrições quanto a sua precisão, especialmente para protozoários que ficam fixos nas brânquias. Inicialmente pelo fato de que não há uma padronização no volume utilizado para a contagem e, principalmente, pelo fato de a estimativa basear-se no volume total da solução com as brânquias, que, além de também não ser padronizado, não leva em conta diferenças de tamanho entre brânquias, não evidenciando, portanto, uma real relação com a superfície parasitada. Torna-se, então, de extrema importância o desenvolvimento e a padronização de métodos mais precisos e confiáveis para quantificação de protozoários parasitos de peixes no Brasil.

Com esse objetivo e buscando uma alternativa viável para a quantificação de protozoários em nossa realidade, foi desenvolvido, neste estudo, um método de quantificação de protozoários em brânquias de tilápias do Nilo, baseado em extrapolações de contagens amostrais em função de parâmetros estimativos obtidos através de modelos matemáticos, que posteriormente será apresentado sob a forma de um software. Tendo em vista que softwares podem ser ferramentas intuitivas, portanto de fácil utilização, e que modelos matemáticos não dependem de equipamentos sofisticados, podendo ser bastante confiáveis e precisos.

Para desenvolver o software, primeiramente foi desenvolvido um modelo preditivo capaz de estimar os parâmetros utilizados nas extrapolações. Assim, buscou-se com o modelo aqui desenvolvido estimar o peso seco total de tecido branquial e o peso seco de tecido branquial presente nas amostras de 1,5 ml (compostas por solução com raspado de suas brânquias) usadas para quantificar os protozoários no microscópio.

A partir destes dados, é possível inferir sobre o valor total de protozoários parasitos presentes nas brânquias, com base no valor de protozoários quantificados nas amostras, usando o peso seco como padrão de extrapolação. Tal qual: se em n amostras de x g de brânquias quantificadas foram quantificados y protozoários, em um total de w g de brânquias espera-se ter z protozoários.

Este método baseia-se em um princípio de amostragem e inferência estatística (COSTA, 2008) e tem como pressupostos as condições de que:

- A. Dada uma mesma solução homogênea com raspado de brânquias, amostragens de mesmo volume resultam em quantidades equivalentes de tecido branquial
- B. Dada uma mesma brânquia, sabendo-se a quantidade de protozoários presentes em uma amostra de tecido branquial com peso conhecido, é possível extrapolar essa quantidade para o peso total do tecido da brânquia
- C. Dadas soluções com raspado de brânquias de diferentes peixes, uma vez fixo o volume da solução e o volume da amostragem, a variação no peso seco encontrado será exclusivamente em função do tamanho e peso da brânquia.

Tais pressupostos são embasados pela lógica indutiva da inferência estatística e probabilística (COSTA, 2008) e por testes realizados experimentalmente. A escolha do peso seco como padrão estimativo para o cálculo da quantidade de protozoários presentes nas brânquias baseia-se no fato de o peso seco ser a medida que melhor representa a quantidade de material do tecido branquial, local onde os protozoários se fixam e se alimentam, visto que o cálculo preciso da área superficial das brânquias é extremamente difícil.

A tilápia do Nilo, *Oreochromis niloticus* (Linnaeus, 1758) foi a espécie escolhida para o desenvolvimento do modelo preditivo, devido à sua grande importância na piscicultura de água doce tanto mundial (CAVICHIOLO *et al.*, 2002; FAO, 2009), quanto nacional (BOSCARDIN, 2008; EPAGRI, 2008).

Este capítulo, portanto, objetiva desenvolver um modelo preditivo para estimar parâmetros de peso seco, que serão posteriormente utilizados no desenvolvimento de um software para quantificação de protozoários parasitos de brânquias de peixes. Neste processo, ele também busca estabelecer protocolos de coleta e aferição para que possam ser desenvolvidos modelos semelhantes para outras espécies de peixe.

MATERIAIS E MÉTODOS

1.1. Coleta dos animais

Todas as coletas seguiram o mesmo processo, que está representando na Figura 1.

A primeira coleta foi realizada no dia 29 de outubro de 2009 na Fundação 25 de Julho, localizada no município de Joinville (26° 18' 16'' S, 48° 50' 44'' W).

As tilápias do Nilo, coletadas por meio de tarrafa, foram previamente separadas em cinco classes de peso, cada uma correspondendo a um intervalo de 50 g em uma escada de 0 à 250 g, sendo coletados cinco exemplares para cada classe. Os animais obtidos foram anestesiadas com eugenol (5 gotas L⁻¹), fotografados sobre fundo azul ao lado de uma régua e pesados em uma balança eletrônica (Apêndice A). Em seguida foram eutanasiados por aprofundamento anestésico e tiveram as brânquias extraídas, pesadas individualmente em balança de precisão e fixadas em formalina 10% em potes numerados com volume fixo de 200 ml.

A segunda coleta foi realizada, também, na Fundação 25 de Julho em Joinville, nos dias 10 e 11 de maio de 2010. Nesta foram coletadas 94 tilápias, sendo cada uma delas submetida aos mesmos procedimentos realizados na primeira coleta, entretanto, desta vez os animais foram separados em 10 classes de peso, cada um correspondendo a 100 g em uma escala de 0 à 1.000 g.

Uma terceira coleta foi realizada no dia 19 de maio de 2010, desta vez em laboratório, com peixes oriundos da Piscicultura Panamá, localizada no município de Paulo Lopes. Foi coletado um total de 11 peixes, com pesos variando entre 46,4 e 129,2 g. Os procedimentos realizados foram os mesmos, com a exceção de que a balança eletrônica utilizada para pesar os peixes foi do modelo Marte e a balança de precisão utilizada para pesar as brânquias foi do modelo Bel Mark série M.

De forma geral, buscou-se que os 130 peixes amostrados representassem diferentes classes de peso, de modo que dentro de cada classe houvessem repetições nos pesos amostrados, para garantir que as correlações entre as variáveis não tivessem como origem uma única amostra. Um dos animais coletados foi descartado por apresentar peso de 2,975 kg, valor muito distante dos demais animais amostrados e dos valores de peso encontrados na produção de tilápias.

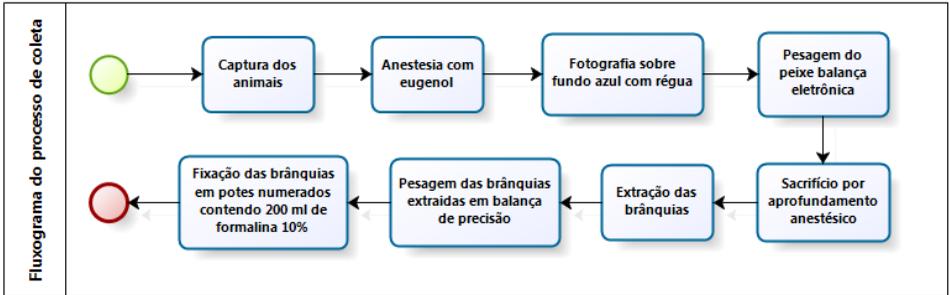


Figura 1: Fluxograma do processo de coleta dos animais.

1.2. Aferição das variáveis para a modelagem

As Figuras 2 e 3 descrevem os processos de aferição das variáveis de forma simplificada e detalhada, respectivamente.

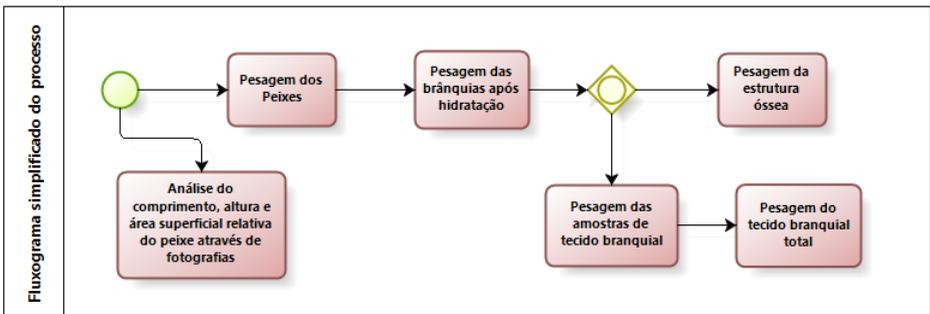


Figura 2: Fluxograma simplificado do processo de aferição das variáveis para modelagem.

2.2.1. Peso do peixe

Foram utilizados os mesmo valores de peso que foram mensurados durante as coletas.

2.2.2. Comprimento, altura e área superficial relativa do peixe (ASR)

Os valores de comprimento total do peixe, altura máxima do peixe e área superficial relativa foram obtidos por meio de análises das fotografias dos animais realizadas com auxílio do software Adobe Photoshop CS3.

Para realizar essas análises foi definida, primeiramente, uma escala customizada de medição, que utilizou como base para cálculo a unidade lógica de 1 cm definida a partir da régua que foi fotografada junto aos peixes. Em seguida, foi selecionada toda a área da fotografia correspondente ao peixe e foram gravadas as medições.

O comprimento total do peixe também foi mensurado durante a coleta com o auxílio de uma régua, porém, as medições feitas nas fotografias provaram-se mais acuradas, especialmente por calcularem valores com várias casas decimais de precisão.

O comprimento total foi definido como a medida da boca até o fim da nadadeira caudal do peixe, enquanto a altura máxima foi definida como a medida de um extremo ao outro no eixo vertical do peixe. A área superficial relativa (ASR) corresponde à área superficial plana do peixe capturada na fotografia.

2.2.3. Peso da brânquia hidratada (PBH)

Para que fossem encontrados os valores de peso das brânquias hidratadas, cada uma das 105 brânquias coletadas nas duas últimas coletas foi mantida em solução de formalina 10% por duas semanas, sendo novamente pesadas, em balança de precisão Bel Mark série M, após esse período. Esse procedimento foi realizado para garantir uma padronização nas pesagens das brânquias, de forma que não houvesse variância causada pelo uso de balanças distintas nas diferentes coletas. Infelizmente as 25 primeiras brânquias coletadas não passaram pelo mesmo procedimento, porém, todas as 25 tiveram o seu peso corrigido para a modelagem em função da média de porcentagem do ganho de peso na hidratação, cujo valor encontrado foi de 41%.

2.2.4. Peso seco ósseo (PSO)

Dando continuidade ao processamento das amostras, as brânquias foram meticulosamente raspadas sobre placas de Petri, conforme protocolo de Jerônimo et al. (2011), tomando-se cuidado para separar todo o tecido branquial da estrutura óssea. Todo o tecido branquial foi devolvido ao seu pote de origem com formalina 10%. A parcela de tecido cartilaginoso foi removida e a estrutura óssea foi colocada em placas de Petri de vidro, forradas com papel alumínio, para secagem em estufa a 200 °C por 30 min. Posteriormente, cada estrutura óssea foi pesada, em balança de precisão Bel Mark série M, para que fossem obtidos os valores de peso seco ósseo.

2.2.5. Peso seco de tecido branquial em amostra de 1,5 ml (PSTA)

Para a avaliação do peso seco de tecido em uma amostra de 1,5 ml de solução de formalina 10% com raspado de brânquias, primeiramente foram pesados individualmente, em balança de precisão Bioprecisa FA2104N, 650 tubos Eppendorf de 2,0 ml, cinco para cada peixe coletado.

Em seguida, cada um dos potes com raspado de tecido branquial teve seu conteúdo despejado em um Becker de vidro graduado, com atenção para que parte do tecido não permanecesse dentro do pote. O volume da solução foi checado e caso não correspondesse ao volume estipulado de 200 ml, medidas de correção eram tomadas para assegurar a precisão do volume, tomando-se cuidado para não interferir no conteúdo de tecido branquial.

Para a retirada das alíquotas, a solução foi devidamente homogeneizada por agitação, sendo retiradas, com auxílio de um pipetador automático LABMATE Soft 10 ml, cinco alíquotas de 1,5 ml, que foram pipetadas nos tubos Eppendorf previamente pesados e numerados. Com o intuito de manter o volume fixo e reduzir o erro provocado por uma mudança na diluição, para cada alíquota retirada, o volume de 1,5 ml era repostado com a mesma quantidade de formalina 10%. Após a retirada das cinco alíquotas, cada solução com raspado de brânquias foi devolvida ao seu respectivo pote.

Para o processo de secagem, os eppendorfs foram centrifugados, numa centrifuga modelo Eppendorf 5804R, por 10 minutos a 14.000 RPM, tiveram o sobrenadante descartado e foram mantidos em estufa a 75 °C overnight (Apêndice A), sendo posteriormente pesados em balança de precisão Bioprecisa FA2104N para serem obtidos os valores de peso seco de tecido branquial em amostra de 1,5 ml.

Devido à escala de precisão desta variável, foram utilizadas quatro casas decimais e todas as pesagens foram feitas na presença de material dessecante (sílica gel). Além disso, a temperatura ambiente durante as pesagens foi registrada.

2.2.6. Peso seco total do tecido branquial (PST)

Para mensurar esta variável, o conteúdo de cada pote foi filtrado em uma malha de 150 μm , de forma que todo o tecido branquial de cada peixe fosse separado e colocado em placas de Petri de vidro, previamente numeradas, forradas com papel alumínio e pesadas em balança de precisão Bel Mark série M.

O tecido foi vistoriado em busca de qualquer sujeira ou material não pertencente ao tecido branquial, como cartilagens ou pedaços de escamas que, caso encontrado, foi prontamente removido.

Para secagem (Apêndice A), as placas de Petri foram mantidas a 100 °C *overnight*, sendo os valores de peso seco total do tecido branquial obtidos após pesagem em balança de precisão Bel Mark série M.

1.3. Desenvolvimento do modelo

As Figuras 4 e 5 descrevem o processo de desenvolvimento do modelo preditivo de forma simplificada e detalhada, respectivamente.

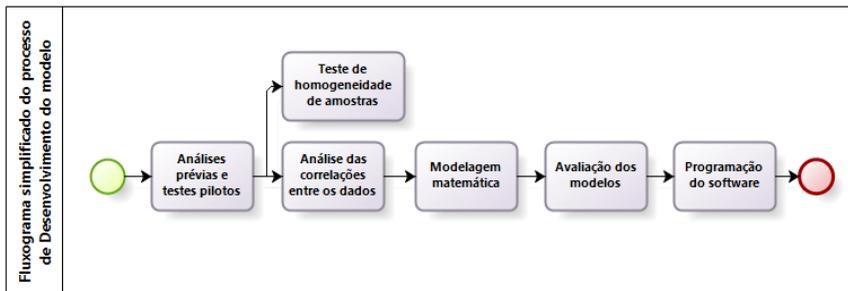


Figura 4: Fluxograma simplificado do processo de desenvolvimento do modelo.

2.3.1. Análises prévias

Foram feitas diversas análises prévias para verificar o comportamento dos dados e garantir o cumprimento dos pré-requisitos pertinentes ao desenvolvimento do modelo. Nessas análises prévias foram utilizadas 10 amostras com o intuito de verificar o peso seco total da brânquia, reduzindo assim o n amostral para as análises de PSO, PST e PSTA. A variável “peso seco total da brânquia” foi abandonada, pois se mostrou inadequada para ser usada na modelagem, uma vez que sua aferição impossibilitava a aferição da variável PST, que é, por sua vez, muito mais representativa da área utilizada para o parasitismo.

2.3.2. Teste de homogeneidade de amostras

Um dos primeiros passos para o desenvolvimento do modelo foi verificar a validade de um dos pré-requisitos para o modelo: a condição de que dada uma mesma solução homogênea com raspado de brânquias, amostragens de mesmo volume resultariam em quantidades equivalentes de tecido branquial.

Para tanto foi realizado um teste piloto, onde 20 amostras de 1,5 ml, retiradas de uma mesma solução de raspado de brânquias, foram secas em estufa e o peso seco do tecido branquial mensurado em balança de precisão Bioprecisa FA2104N. Dado que este teste piloto, apesar de rudimentar, apontou que os valores de peso seco do tecido branquial foram muito próximos entre as amostras, resultando em desvios padrões bastante reduzidos, foi possível dar continuidade ao processo de desenvolvimento do modelo.

Como estratégia para avaliar a homogeneidade das amostras de 1,5 ml retiradas da solução de raspado de brânquias, cada uma das cinco alíquotas retiradas foi assumida como sendo um tratamento, sendo os diferentes peixes considerados como repetições, para que pudesse ser realizada uma análise de variância (ANOVA) para medidas repetidas, comparando as médias dos valores obtidos em cada uma das alíquotas, tanto através do software STATISTICA 8.0 (StatSoft), quanto por meio de *bootstrap* de 100.000 amostras com o software *Resampling Procedures* (David C. Howell)

Além disso, cada conjunto de cinco repetições teve sua média e desvio padrão calculados sendo, posteriormente, calculada a média dos desvios dos conjuntos de repetições e o coeficiente de variação médio.

2.3.3. Correlação dos dados

Para que uma variável possa ser descrita por outra de forma adequada, é importante que exista uma forte correlação entre as duas, sendo que o coeficiente de determinação (R^2) é uma das formas mais utilizadas para indicar a força de uma correlação e o grau explicativo de um modelo de regressão.

Dessa forma, com o auxílio dos pacotes computacionais STATISTICA 8.0 (StatSoft) e Excel 2007 (Microsoft), foi realizada uma regressão linear múltipla com todas as variáveis aferidas, além de regressões lineares e potenciais para cada par de variáveis. Foram testados, também, outros tipos de regressões, como logarítmica e exponencial, porém as que melhor descreveram as relações entre as variáveis foram as regressões lineares e potenciais.

Os coeficientes de determinação (R^2) e as estimativas das equações obtidas nas regressões lineares e de potência foram avaliados e comparadas em busca das equações que melhor descrevessem o comportamento de cada uma das variáveis analisadas. Na análise das estimativas, além da análise dos resíduos, foi calculado o qui-quadrado e feito um teste T de Student para dados pareados, comparando as médias dos valores observados com a dos valores estimados.

As regressões que apresentaram os melhores resultados foram submetidas a um ajuste pelo método dos quadrados mínimos e seus pares de variáveis tiveram seus coeficientes de correlação linear de Pearson (r) calculados por meio de *bootstraps* com 10.000 amostras, utilizando o software *Resampling Procedures* (David C. Howell), de maneira que foi possível obter o intervalo de 99% de confiança para R^2 .

2.3.4. Modelagem

Foram testadas e comparadas, com auxílio do software STATISTICA 8.0, diversas ferramentas de modelagem para os dados, de forma que cada uma delas passou por uma análise das previsões e dos resíduos.

Foram testadas as ferramentas de aprendizagem de máquina (*machine learning*) *k-Nearest Neighbors* e *Support Vector Machines* (SVM), bem como o a ferramenta de *Neural Network* do STATISTICA 8.0, *STATISTICA Automated Neural Networks* (SANN), porém as melhores estimativas e menores resíduos, especialmente para valores alvos pequenos, foram obtidas com *Multivariate Adaptive Regression Splines* (MARSplines).

Os dados de todas as variáveis foram submetidos ao método de MARSplines (FRIEDMAN, 1991), onde foi avaliada a importância relativa de cada uma das variáveis preditoras. Após esse processo, foram selecionadas, dentre as variáveis com os melhores resultados, aquelas com a maior praticidade para serem aferidas, sendo, portanto, as variáveis “peso do peixe” e “comprimento” escolhidas como preditoras. Os dados foram, então, novamente submetidos ao método MARSplines, obtendo-se as equações do modelo para as variáveis dependentes PBH, “Peso seco de tecido branquial em amostra de 1,5 ml” e “Peso seco total do tecido branquial”.

Os resultados da modelagem por MARSplines foram avaliados através da análise dos resíduos e da qualidade de ajuste (*goodness of fit*), onde foram verificados os R^2 e os valores dos erros absolutos e relativos. A significância estatística de cada um dos parâmetros estimados foi avaliada utilizando-se Teste T de Student para dados pareados, comparando os valores estimados aos valores observados, tanto por meio do software STATISTICA 8.0, quanto pelo software *Resampling Procedures*, onde se utilizou um *bootstrap* de 100.000 repetições.

Fluxograma detalhado do processo de desenvolvimento do modelo

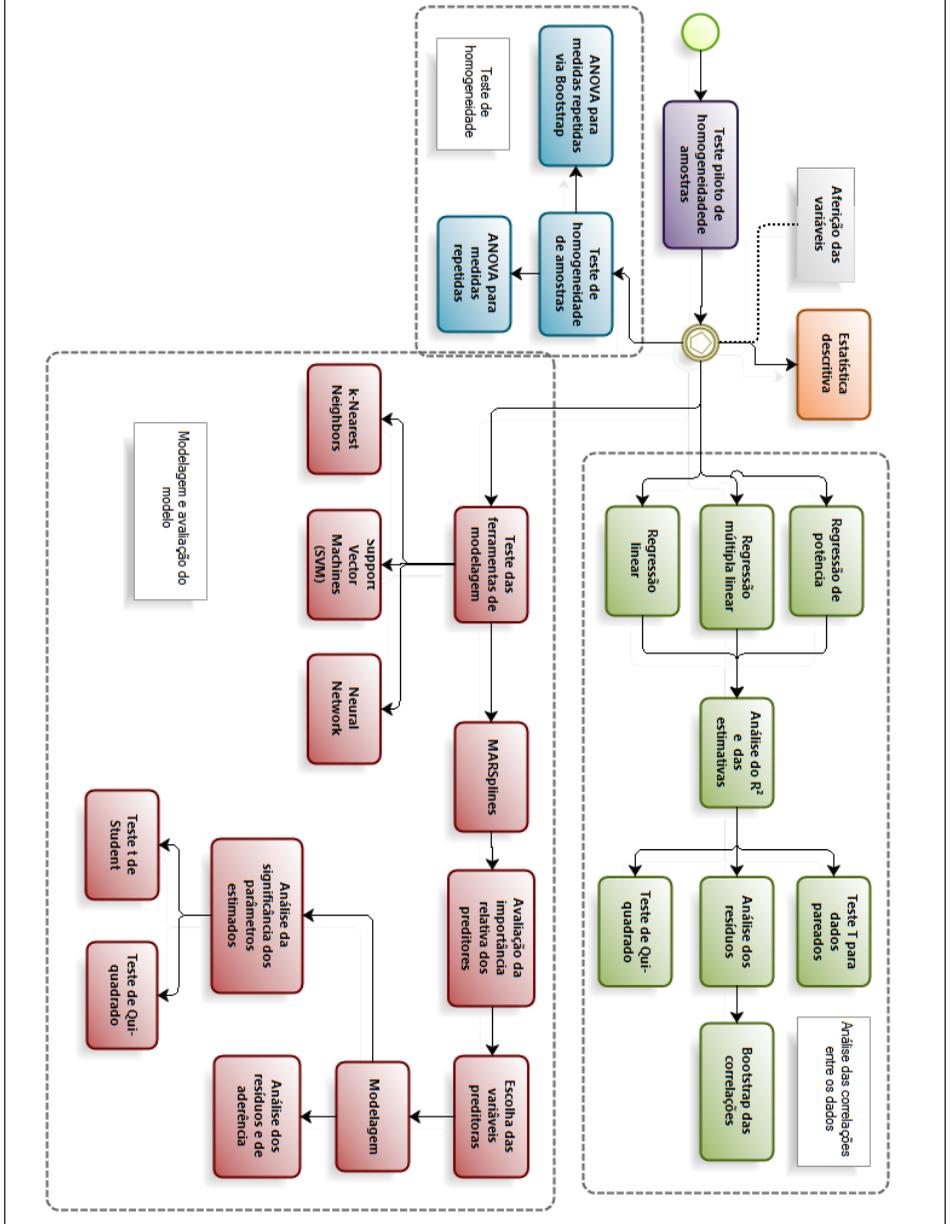


Figura 5: Fluxograma detalhado do processo de desenvolvimento do modelo

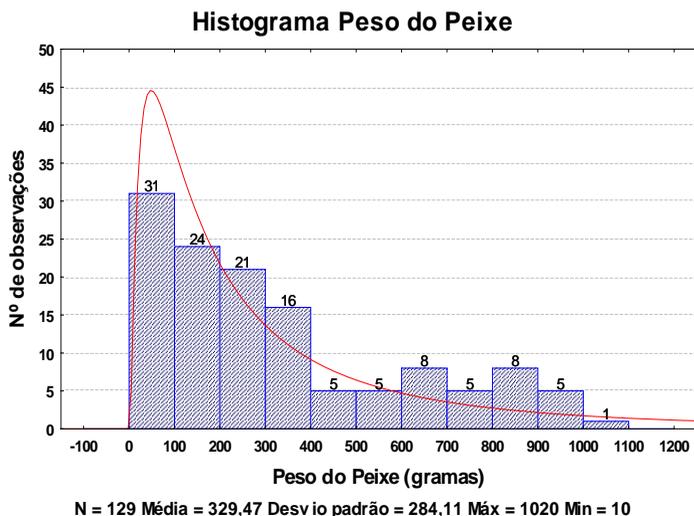
RESULTADOS

2.1. Estatística descritiva

3.1.1. Peso do peixe

Constituídos por um n amostral de 129 peixes, os valores de peso apresentaram média e desvio padrão de $329,47 \pm 284,11$ g, sendo o valor mínimo igual a 10 e o máximo igual a 1.020 g. Com os dados aferidos foram construídos um histograma (Figura 6.a), representando as classes de peso amostradas e suas freqüências de observação, e um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes (Figura 6.b).

Os dados apresentaram distribuição log-normal, sendo que cada classe teve, pelo menos, cinco repetições, com exceção da classe correspondente ao peso acima de 1 kg, com somente um representante. O maior desvio padrão observado foi referente à classe com pesos entre 400 e 500 g, enquanto o menor desvio encontrado nas quatro primeiras classes de peso; devendo-se tanto ao maior número de observações, quanto à maior homogeneidade de peso apresentada pelos peixes durante essa etapa de cultivo.



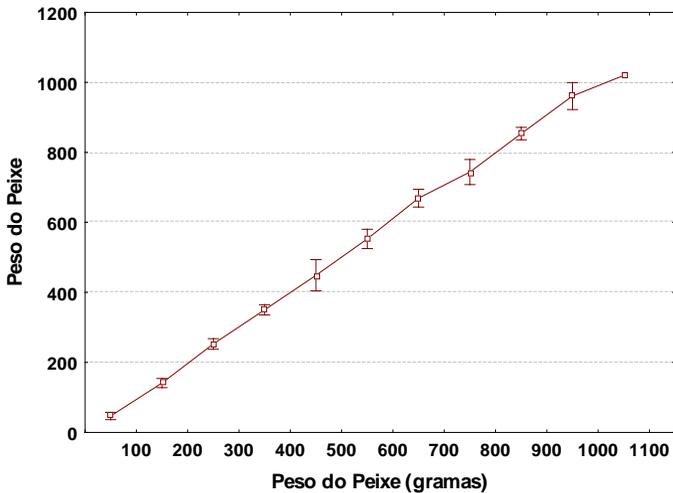


Figura 6: a) Histograma do peso do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

3.1.2. Comprimento

Constituídos por um n amostral de 128 peixes, devido ao fato de uma das fotografias ter sido extraviada, os valores de comprimento apresentaram média e desvio padrão de $26,07 \pm 9,25$ cm, sendo o valor mínimo igual a 9,25 e o máximo igual a 44,49 cm. Com os dados aferidos foram construídos um histograma (Figura 7.a), representando as classes de comprimento amostradas e suas freqüências de observação, e um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes (Figura 7.b).

Os valores de comprimento apresentaram um padrão de distribuição normal, apesar de serem oriundos dos mesmos peixes. Tal fato é devido a uma menor variação em comparação com o peso do peixe, o que é, também, refletido nos valores de intervalo de confiança.

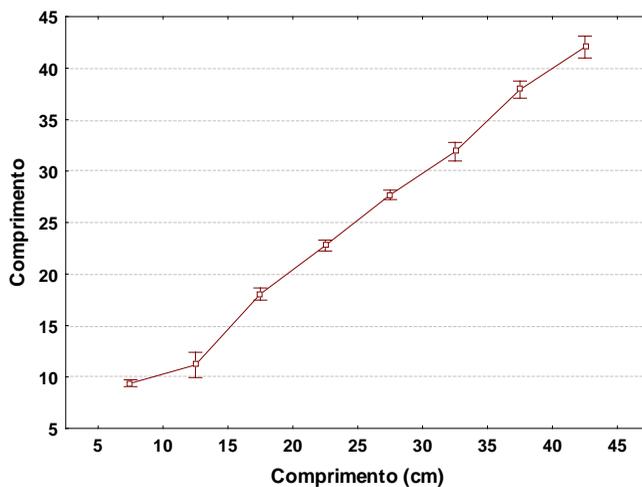
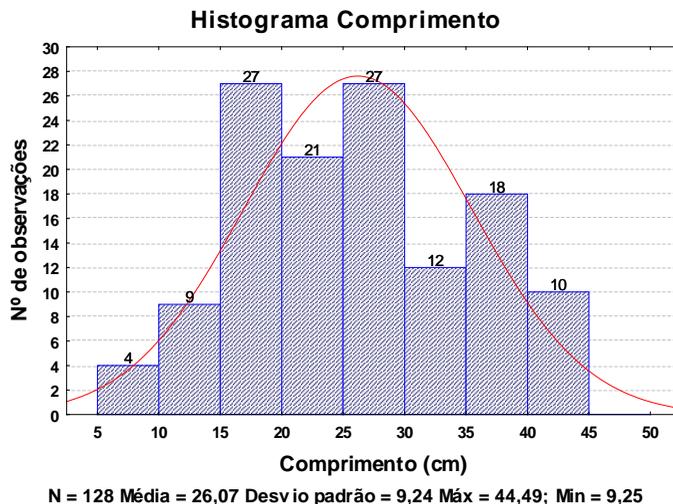


Figura 7: a) Histograma do comprimento do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

3.1.3. Altura

Constituídos, também, por um n amostral de 128 peixes, os valores de altura apresentaram média e desvio padrão de $10,38 \pm 4,22$ cm, sendo o valor mínimo igual a 3,02 e o máximo igual a 19,77 cm. Com os dados aferidos também foram construídos um histograma

(Figura 8.a), representando as classes de altura amostradas e suas freqüências de observação, e um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes (Figura 8.b).

Assim como os valores de comprimento, os valores de altura também apresentaram uma distribuição normal, bem como intervalos de confiança bastante reduzidos.

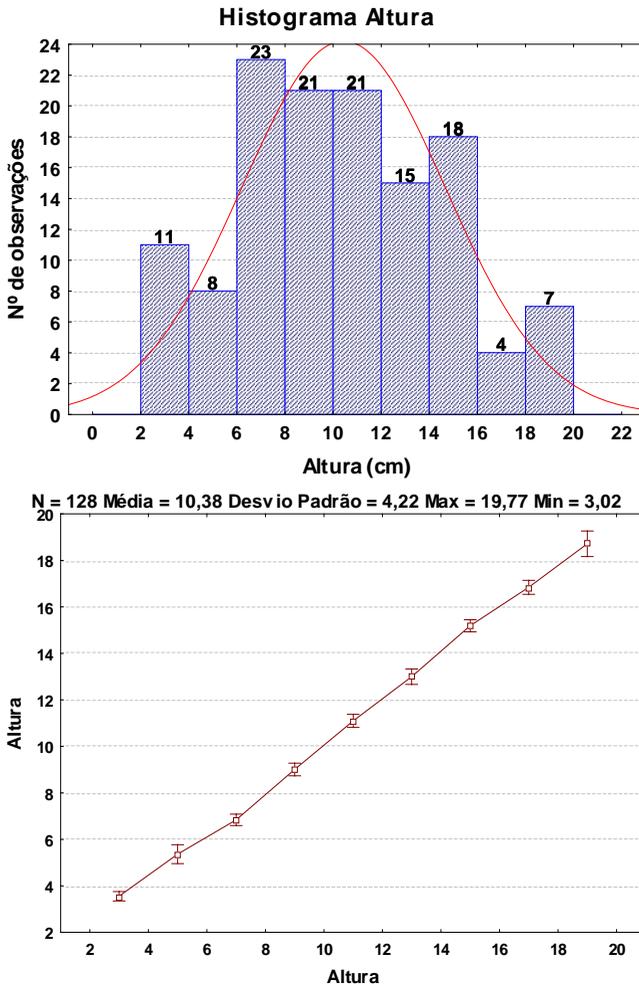
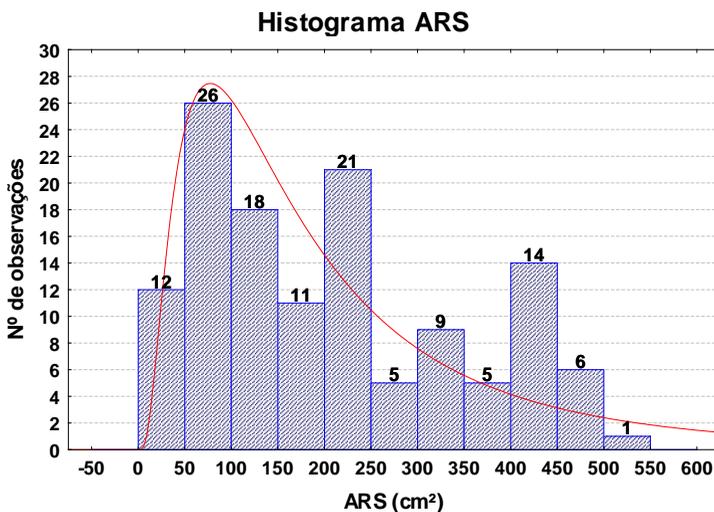


Figura 8: a) Histograma da altura do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

3.1.4. Área superficial relativa (ASR)

Com um n amostral de 128 peixes, os valores de área superficial apresentaram média e desvio padrão de $209,37 \pm 139,37 \text{ cm}^2$, sendo o valor mínimo igual a 19,79 e o máximo igual a $513,79 \text{ cm}^2$. Foram construídos um histograma (Figura 9.a), representando as classes de área superficial relativa amostradas e suas freqüências de observação, além de um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes (Figura 9.b).

A distribuição dos valores de área superficial relativa enquadrou-se como log-normal, sendo a classe com valores entre 350 e 400 cm^2 aquela que apresentou o maior desvio padrão.



N = 128 Média = 209,37 Desvio Padrão = 139,37 Max = 513,79 Min = 19,79

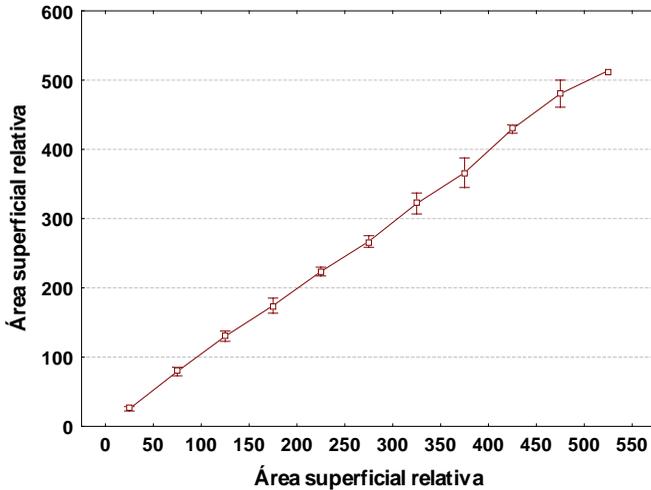


Figura 9: a) Histograma da área superficial relativa do peixe com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

3.1.5. Peso da brânquia hidratada (PBH)

Com um n amostral de 129 peixes, os valores de PBH apresentaram média e desvio padrão de $14,64 \pm 11,86$ g, sendo o valor mínimo igual a 0,68 e o máximo igual a 45,7 g. Foram, também, construídos um histograma (Figura 10.a), representando as classes de PBH amostradas e suas freqüências de observação, e um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes amostradas (Figura 10.b).

Os valores de PBH (Figura 10.a) apresentam distribuição log-normal muito próxima daquela apresentada pelos valores de peso do peixe (Figura 6.a). Os intervalos de confiança foram, em geral, reduzidos, com exceção das classes que apresentaram número de observações abaixo de quatro, cujo desvio padrão foi mais elevado.

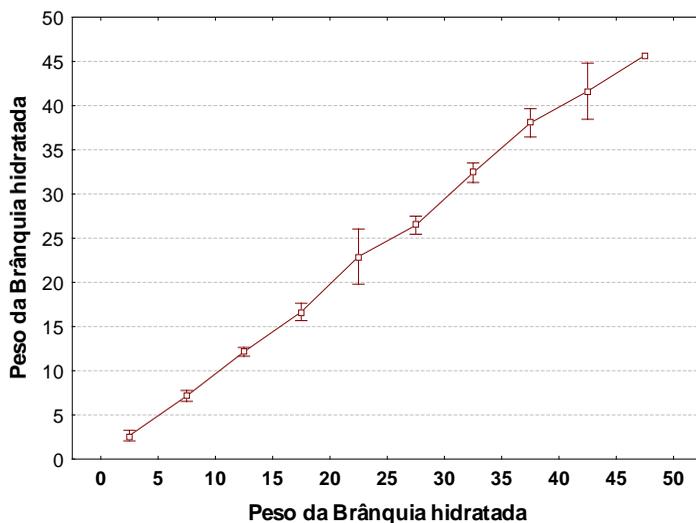
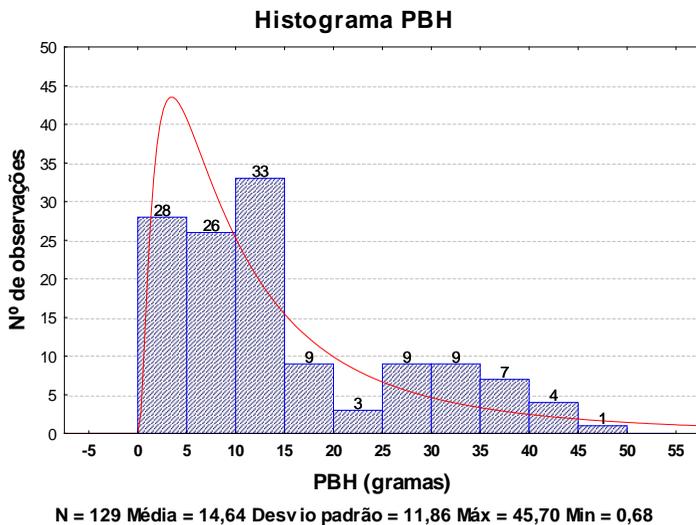


Figura 10: a) Histograma do peso da brânquia hidratada com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

3.1.6. Peso seco ósseo (PSO)

Com um n amostral de 119 peixes, devido ao uso de 10 amostras em análises prévias, os valores de PSO apresentaram média e desvio

padrão de $1,95 \pm 1,97$ g, sendo o valor mínimo igual a 0,024 e o máximo igual a 7,67 g. Foi feito um histograma (Figura 11.a), representando as classes de PSO amostradas e suas freqüências de observação, e um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes (Figura 11.b).

Os valores de PSO encontrados ficaram bastante concentrados nas duas primeiras classes, ao passo que as duas últimas classes apresentaram um desvio padrão elevado.

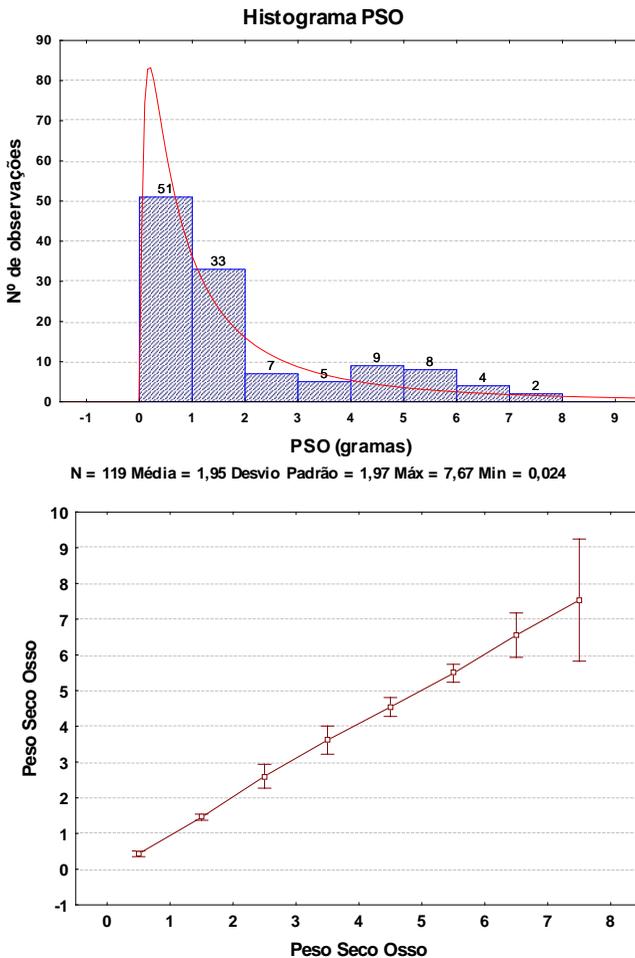
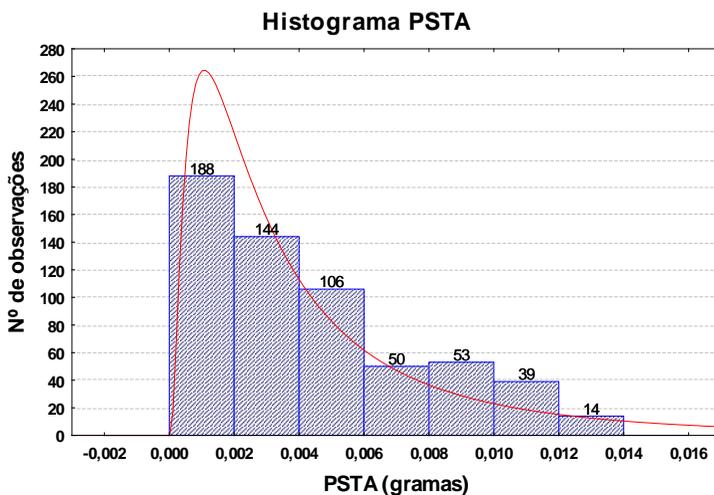


Figura 11: a) Histograma do peso seco ósseo com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

3.1.7. Peso seco de tecido branquial em amostra de 1,5 ml (PSTA)

Devido ao fato de terem sido feitas cinco repetições para cada um dos 119 peixes, sendo uma das repetições de um dos peixes impossível de mensurar na balança de precisão disponível, o n amostral total de PSTA foi de 594. Seus valores apresentaram média e desvio padrão de $0,0044 \pm 0,0034$ g, sendo o valor mínimo igual a 0,0001 e o máximo igual a 0,0134 g. Foi feito um histograma (Figura 12.a), representando as classes de PSTA amostradas e suas freqüências de observação, e um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes (Figura 12.b).

Os valores de PSTA apresentaram uma distribuição log-normal. Como é possível ver no gráfico 7.b, o intervalo de confiança das classes amostrais reflete um desvio padrão bastante baixo, dando suporte à validação do pressuposto A.



N = 594 Média = 0,0044 Desvio Padrão = 0,0034 Máx = 0,0134 Min = 0,0001

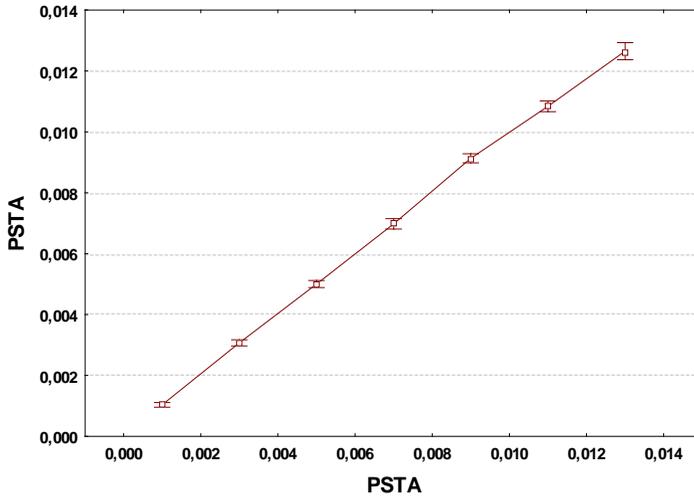


Figura 12: a) Histograma do peso seco de tecido branquial na amostra (PSTA) com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

3.1.8. Peso seco total do tecido branquial (PST)

Com um n amostral de 118, devido às análises prévias e a perda de uma amostra, os valores de PST apresentaram média e desvio padrão de $0,54 \pm 0,42$ g, sendo o valor mínimo igual a 0,019 e o máximo igual a 1,64 g. Foi feito um histograma (Figura 13.a), representando as classes de PST amostradas e suas freqüências de observação, e um gráfico de média e erro, representando a média e intervalo de confiança (95%) para cada uma das classes (Figura 13.b).

Os valores de PST também apresentaram distribuição próxima da log-normal, sendo a quarta classe aquela que apresentou o maior desvio padrão.

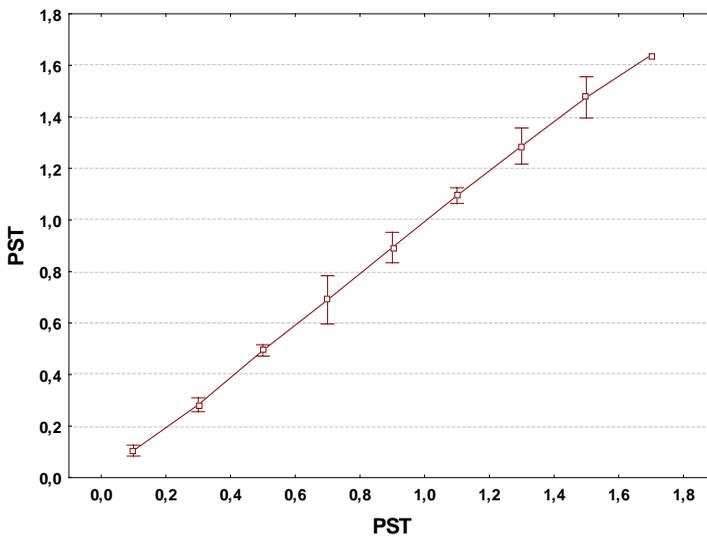
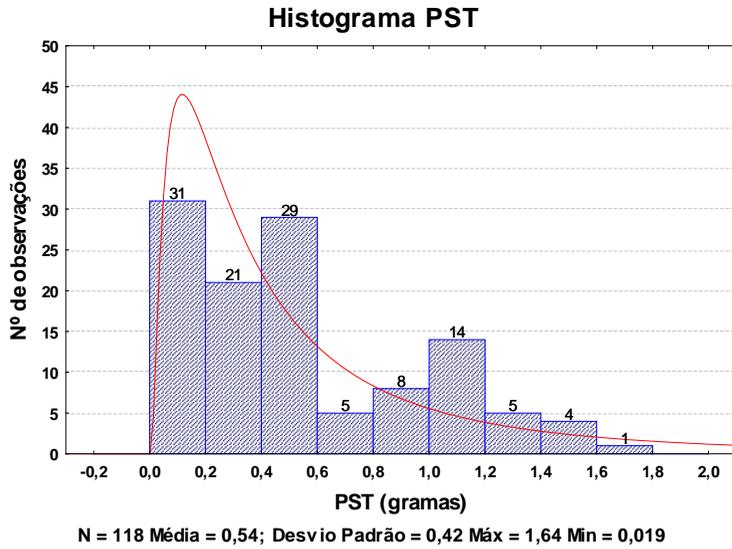


Figura 13: a) Histograma do peso seco de tecido branquial (PST) com estatística descritiva b) médias com intervalo de confiança (95%) das classes.

2.2. Desenvolvimento do modelo

3.2.1. Teste de homogeneidade de amostras

Como anteriormente ressaltado na estatística descritiva, os intervalos de confiança apresentados na Figura 12.b são, por si só, um forte indicativo do baixo desvio padrão entre as repetições. Além disso, a média dos desvios dos conjuntos de repetições foi 0,00064, apresentando um coeficiente de variação de 14,7% em relação a média de 0,0044.

A análise de variância, realizada com as cinco alíquotas de 1,5 ml retiradas da solução de raspado de brânquias, esteve de acordo com os pressupostos de normalidade ($p = 0,99$) e mostrou que não houve diferença estatística ($p = 0,99991$) entre as alíquotas (Figura 14). Além disso, a ANOVA para medidas repetidas realizada via *bootstrap* de 100.000 amostras (Figura 15) também resultou em $p = 0,99$.

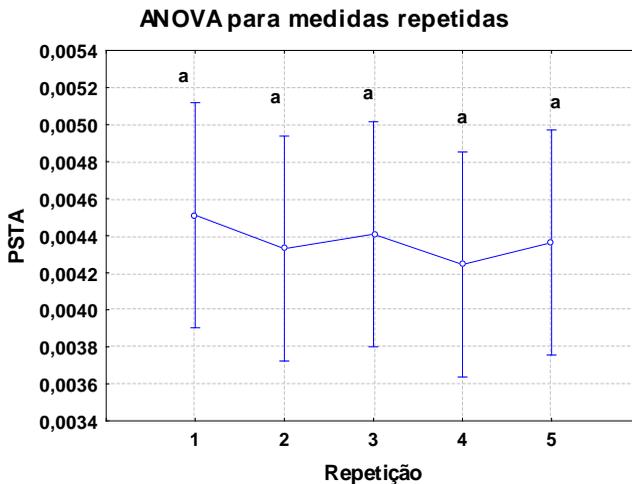


Figura 14: Resultado da ANOVA para medidas repetidas comparando as médias e intervalos de confiança (95%) das repetições das alíquotas de 1,5ml.

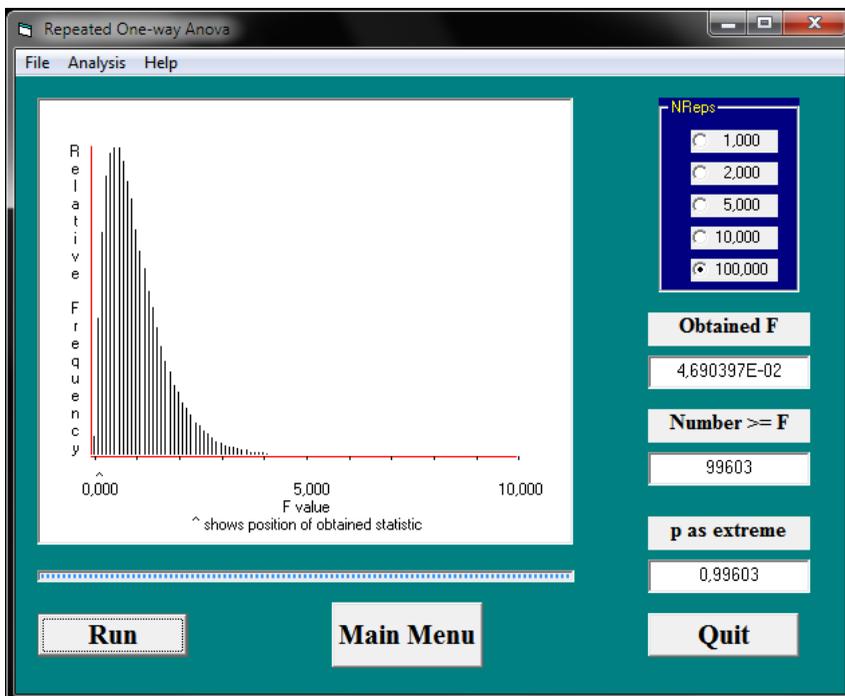


Figura 15: Captura de tela com resultado de ANOVA via *bootstrap* de 100.000 repetições comparando as repetições das alíquotas de 1,5ml.

3.2.2. Regressões e análise das correlações

Devido a grande quantidade de resultados originados nas análises das regressões, serão apresentados aqui somente aqueles com maior relevância.

3.2.2.1. Regressão múltipla

A regressão múltipla feita *a priori*, utilizando todas as variáveis dependentes e independentes aferidas, resultou em um R^2 múltiplo de 0,96 (Tabela 1), com erro padrão da estimativa igual a 2,34.

As variáveis independentes mais significativas para a regressão foram “peso do peixe” e “ASR”, sendo a variável “peso do peixe” aquela que apresentou os maiores valores de significância e correlação (Tabela 3 e 5).

A análise de variância (Tabela 4) mostra que o componente da regressão foi bastante significativo, ao passo que o componente residual foi baixo, indicando que a regressão é capaz de explicar a maior parte da variação dos dados.

Os resultados da regressão múltipla feita *a priori* foram bastante animadores, tanto pelo resultado da análise de variância, quando pelas elevadas correlações apresentadas entre as variáveis independentes e as variáveis dependentes (Tabela 5).

A matriz de correlações está disposta na Tabela 5 e o resumo da estatística descritiva, das variáveis utilizadas na regressão múltipla, pode ser visto na Tabela 2.

Tabela 1: Resumo dos resultados da estatística da regressão. Florianópolis, junho de 2011.

Estatística	Valor
R Múltiplo	0,9809
R ² Múltiplo	0,9622
R ² Ajustado	0,9609
F (4,112)	712,8941
p	0,0000
Erro padrão da estimativa	2,3440

Tabela 2: Estatística descritiva para as variáveis utilizadas na regressão múltipla. Florianópolis, junho de 2011.*

Variáveis	Estatística Descritiva		
	Médias	Desvio	N
Peso do peixe	343,9694	283,4091	117
ASR	220,5490	140,1169	117
Altura	10,7006	4,2328	117
Comprimento	26,8109	9,2370	117
PBH	15,2324	11,8476	117
PSO	1,9233	1,9285	117
PSTA	0,0044	0,0033	117
PST	0,5311	0,4160	117

* ASR = Área superficial relativa; PBH = Peso da brânquia hidratada; PSO = Peso seco ósseo; PSTA = Peso seco de tecido branquial na amostra de 1,5 ml; PST = Peso seco total do tecido branquial.

Tabela 3: Distribuição dos resultados da regressão múltipla para a variável peso da brânquia hidratada (PBH). Florianópolis, junho de 2011.*

N=117	Beta	EroPadrão deBeta	B	EroPadrão deB	t (112)	p
Intersecção			0,705718	1,789833	0,39429	0,694115
Peso do Peixe	0,471387	0,105534	0,019706	0,004412	4,46669	0,000019
ASR	0,653791	0,192649	0,055281	0,016290	3,39368	0,000954
Altura	-0,123824	0,086047	-0,346582	0,240844	-1,43903	0,152930
Comprimento	-0,021378	0,129548	-0,027421	0,166162	-0,16502	0,869223

* Os valores em vermelho indicam significância estatística; ASR = Área superficial relativa.

Tabela 4: Análise de variância dos componentes da regressão múltipla para a variável peso da brânquia hidratada (PBH). Florianópolis, junho de 2011.*

Efeito	Análise de Variância: PBH				
	Soma dos Quadrados	gl	Quadrado Médio	F	p
Regressão	15667,10	4	3916,775	712,8940	0,00
Resíduo	615,35	112	5,494		
Total	16282,45				

* Os valores em vermelho indicam significância estatística.

Tabela 5: Matriz de correlações obtidas na regressão múltipla. Florianópolis, junho de 2011.*

Variáveis	Correlações							
	Peso do Peixe	ASR	Altura	Comprimento	PBH	PSO	PSTA	PST
Peso do Peixe	1,00	0,98	0,93	0,95	0,98	0,95	0,95	0,96
ASR	0,98	1,00	0,97	0,98	0,97	0,94	0,96	0,97
Altura	0,93	0,97	1,00	0,97	0,93	0,89	0,92	0,92
Comprimento	0,95	0,98	0,97	1,00	0,95	0,91	0,94	0,94
PBH	0,98	0,97	0,93	0,95	1,00	0,98	0,95	0,97
PSO	0,95	0,94	0,89	0,91	0,98	1,00	0,93	0,94
PSTA	0,95	0,96	0,92	0,94	0,96	0,93	1,00	0,97
PST	0,96	0,97	0,92	0,94	0,97	0,94	0,97	1,00

* ASR = Área superficial relativa; PBH = Peso da brânquia hidratada; PSO = Peso seco ósseo; PSTA = Peso seco de tecido branquial na amostra de 1,5 ml; PST = Peso seco total do tecido branquial.

3.2.2.2. Regressões individuais

Tendo como base os resultados da regressão múltipla, verificamos que a variável independente de maior relevância foi o peso do peixe. Assim, serão apresentados aqui somente os resultados para esta variável.

Na Figura 16 estão representadas as curvas de regressão linear, para cada uma das variáveis dependentes PBH, PSO, PSTA e PST, tendo como variável independente o peso do peixe. Nela é possível verificar as equações e o R^2 para cada uma das regressões individuais.

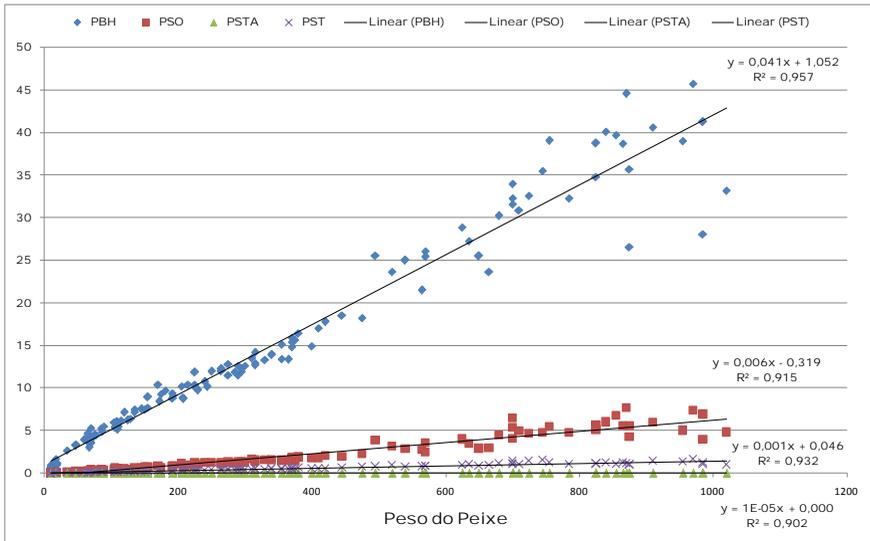


Figura 16: Curvas de regressão linear, onde X = peso do peixe e Y = peso da brânquia hidratada (PBH), peso seco ósseo (PSO), peso seco do tecido branquial na amostra (PSTA) e peso seco total do tecido branquial (PST).

Na Figura 17 são apresentadas as curvas de regressão potencial, para cada uma das variáveis dependentes PBH, PSO, PSTA e PST, usando o peso do peixe como variável independente. Nela, também, é possível verificar as equações e o R^2 para cada uma das regressões individuais.

Pelo comportamento das curvas das Figuras 16 e 17, observa-se que, mesmo na regressão potencial, os dados apresentam comportamento bastante próximo da linearidade. Neste caso, optou-se, então, pelo princípio da parcimônia, tratando os dados como lineares.

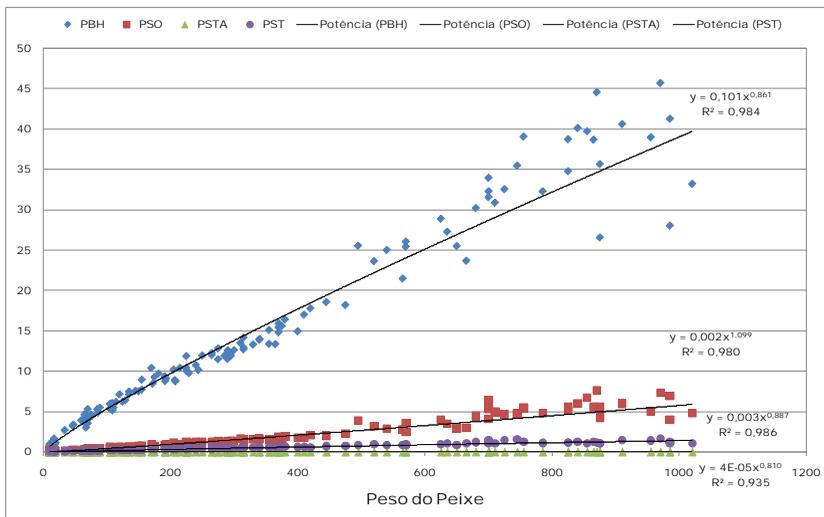


Figura 17: Curvas de regressão potencial, onde X = peso do peixe e Y = peso da brânquia hidratada (PBH), peso seco ósseo (PSO), peso seco do tecido branquial na amostra (PSTA) e peso seco total do tecido branquial (PST).

Observa-se mudança na dispersão dos dados a partir do momento em que os animais atingiram a faixa de 500 g de peso, especialmente para a variável PBH. Essa mudança reflete-se na correlação entre o peso do peixe e o peso das brânquias, que na faixa de peso entre 10 e 500 g apresenta $R^2 = 0,973$ ($n = 97$), e na faixa de peso entre 500 g e 1 kg, apresenta $R^2 = 0,528$ ($n = 32$). Tal fato indica a necessidade de uma modelagem que considere possíveis variações na correlação em diferentes momentos da curva de regressão.

Tabela 6: Comparação entre os valores de média e desvio padrão observados e estimados, para cada uma das variáveis dependentes. Florianópolis, junho de 2011.*

Variáveis	Real		Estimado (linear)		Estimado (potência)	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
PBH	14,56	11,90	14,56	11,65	14,21	10,89
PSO	1,95	1,97	1,95	1,89	1,28	1,15
PSTA	0,0044	0,0033	0,0044	0,0031	0,0043	0,0030
PST	0,54	0,42	0,54	0,41	0,52	0,39

* PBH = Peso da brânquia hidratada; PSO = Peso seco ósseo; PSTA = Peso seco de tecido em amostra de 1,5 ml de solução; PST = Peso seco total do tecido branquial.

Apesar dos menores valores de correlação, as regressões lineares apresentaram melhores estimativas (Tabela 6), especialmente em relação à média e variância, sendo, portanto, definidas como as de maior relevância para serem apresentadas.

3.2.2.3. Análise das estimativas

3.2.2.3.1. Peso da brânquia hidratada (PBH) linear

Na análise das estimativas foi possível constatar que o erro padrão da regressão foi 2,458. Os resíduos das estimativas apresentaram média igual a 0% e ficaram dispersos conforme mostra a Figura 18. Ambas as análises, de qui-quadrado ($\chi^2 = 35,63$) e o teste T pareado (Tabela 7), tiveram como resultado $p = 1,0$, indicando que os valores estimados são estatisticamente semelhantes aos valores observados.

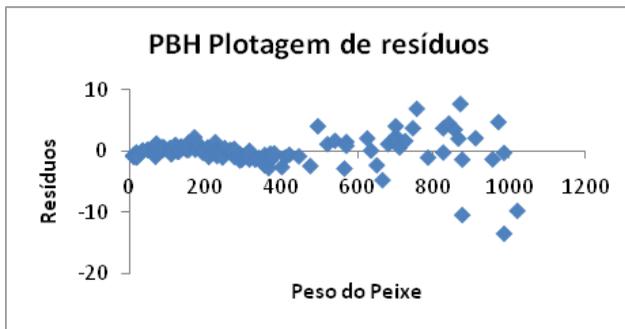


Figura 18: Plotagem dos resíduos da estimativa de peso da brânquia hidratada (PBH).

Tabela 7: Resultados do Teste T pareado comparando os valores estimados de peso da brânquia hidratada (PBH) com os observados. Florianópolis, junho de 2011.

	<i>PBH Real</i>	<i>PBH Estimado</i>
Média	14,56099838	14,56099838
Variância	141,6911577	135,6964321
Observações	129	129
Correlação de Pearson	0,97861719	
Hipótese da diferença de média	0	
gl	128	
Stat t	-4,21594E-15	
P(T<=t) uni-caudal	0,5	
t crítico uni-caudal	1,656845227	
P(T<=t) bi-caudal	1	
t crítico bi-caudal	1,978670823	

3.2.2.3.2. Peso seco ósseo (PSO) linear

Na análise das estimativas foi possível constatar que o erro padrão da regressão foi 0,577. Os resíduos das estimativas também apresentaram média igual a 0% e ficaram dispersos conforme mostra a Figura 19. Novamente, a análise de qui-quadrado ($\chi^2 = 28,92$) e o teste T pareado (Tabela 8) tiveram como resultado $p = 1,0$.

Entretanto, como algumas estimativas apresentaram valores negativos, o uso desta regressão foi considerado inválido.

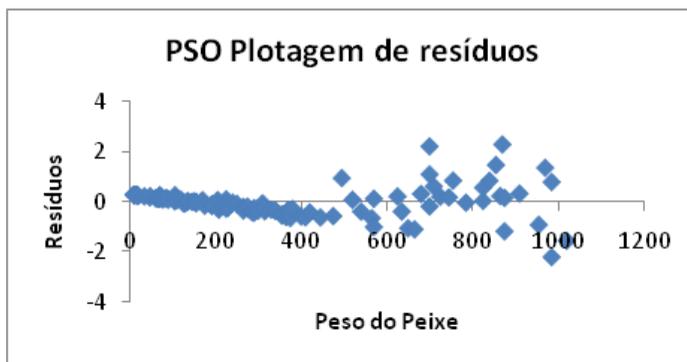


Figura 19: Plotagem dos resíduos da estimativa de peso seco ósseo (PSO).

Tabela 8: Resultados do Teste T pareado comparando os valores estimados de peso seco ósseo (PSO) com os observados. Florianópolis, junho de 2011.

	<i>PSO Real</i>	<i>PSO Estimado</i>
Média	1,952008403	1,952008403
Variância	3,889454534	3,558827675
Observações	119	119
Correlação de Pearson	0,956553203	
Hipótese da diferença de média	0	
gl	118	
Stat t	-5,73473E-15	
P(T<=t) uni-caudal	0,5	
t crítico uni-caudal	1,657869523	
P(T<=t) bi-caudal	1	
t crítico bi-caudal	1,980272226	

3.2.2.3.3. Peso seco do tecido branquial na amostra (PSTA) linear

Na análise das estimativas constatou-se que o erro padrão da regressão foi de 0,001. Os resíduos das estimativas apresentaram média igual a 0% e ficaram dispersos conforme mostra a Figura 20. Ambas análises, de qui-quadrado ($\chi^2 = 0,029$) e o teste T pareado (Tabela 9), tiveram como resultado $p = 1,0$, reforçando a semelhança estatística entre os valores estimados e observados.

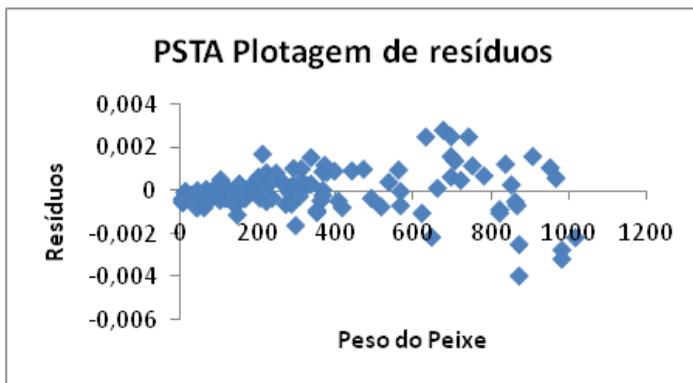


Figura 20: Plotagem dos resíduos da estimativa de peso seco do tecido branquial na amostra (PSTA).

Tabela 9: Resultados do Teste T pareado comparando os valores estimados de peso seco do tecido branquial na amostra (PSTA) com os observados. Florianópolis, junho de 2011.

	<i>PSTA Médio</i>	<i>PSTA Estimado</i>
Média	0,004371429	0,004371429
Variância	1,0856E-05	9,80113E-06
Observações	119	119
Correlação de Pearson	0,950172627	
Hipótese da diferença de média	0	
gl	118	
Stat t	-8,20591E-15	
P(T<=t) uni-caudal	0,5	
t crítico uni-caudal	1,657869523	
P(T<=t) bi-caudal	1	
t crítico bi-caudal	1,980272226	

3.2.2.3.4. Peso seco total do tecido branquial (PST) linear

Na análise das estimativas foi possível constatar que o erro padrão da regressão foi 0,11. Os resíduos das estimativas apresentaram, novamente, média igual a 0% e ficaram dispersos conforme mostra a Figura 21. A análise de qui-quadrado ($\chi^2 = 1,71$) e o teste T pareado (Tabela 10) mais uma vez tiveram como resultado $p = 1,0$, reforçando a semelhança estatística entre os valores estimados e observados.

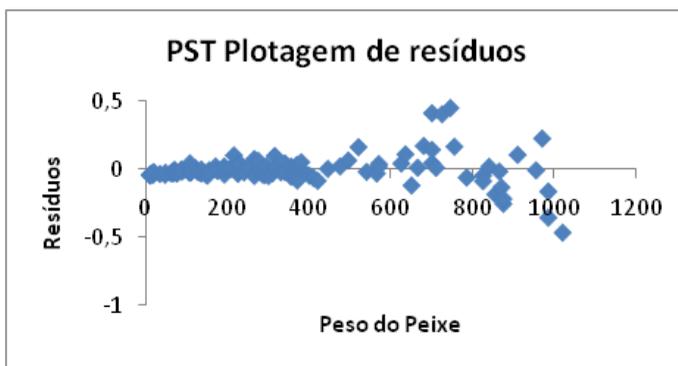


Figura 21: Plotagem dos resíduos da estimativa de peso seco total do tecido branquial (PST).

Tabela 10: Resultados do Teste T pareado comparando os valores estimados de peso seco total do tecido branquial (PST) com os observados. Florianópolis, junho de 2011.

	<i>PST Real</i>	<i>PST Estimado</i>
Média	0,537381017	0,537381017
Variância	0,176228751	0,164321286
Observações	118	118
Correlação de Pearson	0,965625067	
Hipótese da diferença de média	0	
gl	117	
Stat t	6,18164E-15	
P(T<=t) uni-caudal	0,5	
t crítico uni-caudal	1,657981659	
P(T<=t) bi-caudal	1	
t crítico bi-caudal	1,980447532	

3.2.2.4. Análise das correlações

Os resultados do *bootstrap* com 10.000 amostras para PBH (Figura 22) evidenciaram uma correlação linear de Pearson (r) igual a 0,979, com intervalo de confiança de 99% entre 0,962 e 0,993, e erro padrão de 0,006.

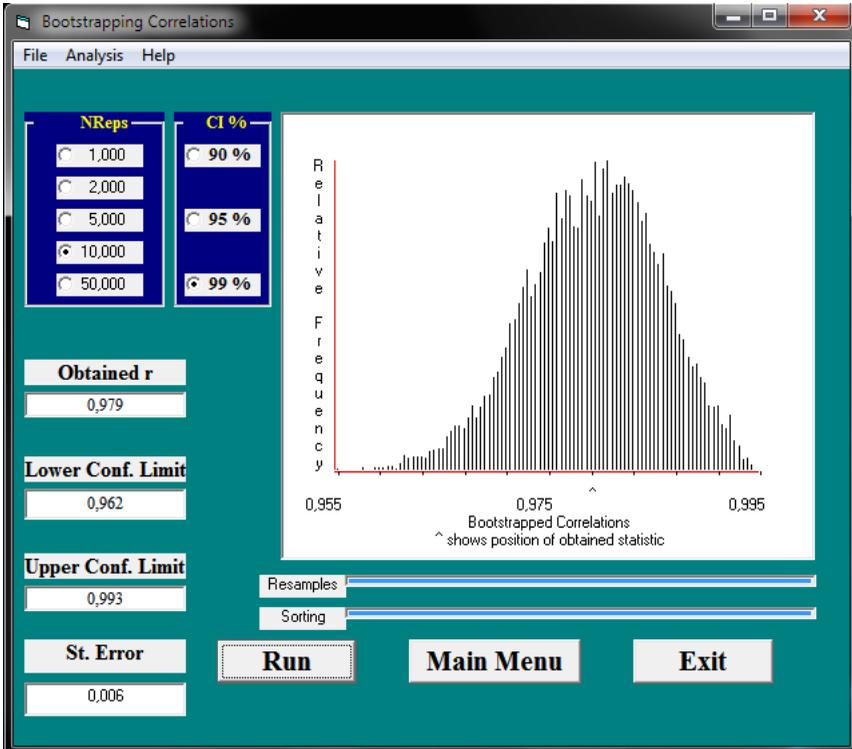


Figura 22: Captura de tela com resultado da análise de correlação por *bootstrap* entre peso do peixe e peso da brânquia hidratada (PBH).

Os resultados para PSO (Figura 23) evidenciaram uma correlação linear de Pearson (r) igual a 0,957, com intervalo de confiança de 99% entre 0,933 e 0,976, e erro padrão de 0,008.

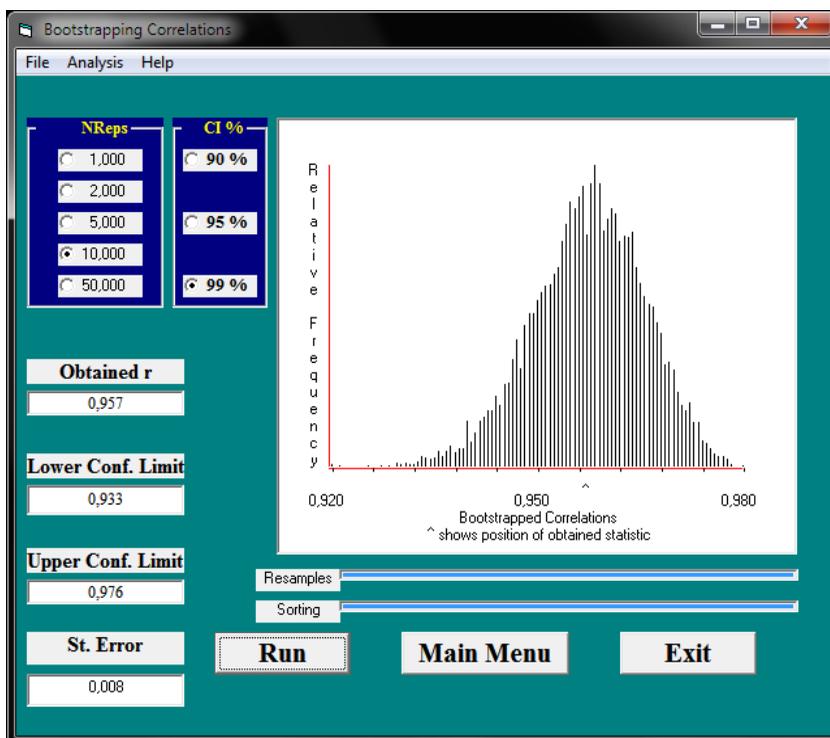


Figura 23: Captura de tela com resultado da análise de correlação por *bootstrap* entre peso do peixe e peso seco ósseo (PSO).

Na análise de correlação entre peso do peixe e PSTA (Figura 24), os resultados mostraram uma correlação linear de Pearson (r) igual a 0,950, com erro padrão de 0,010 e intervalo de confiança de 99% entre 0,923 e 0,975.

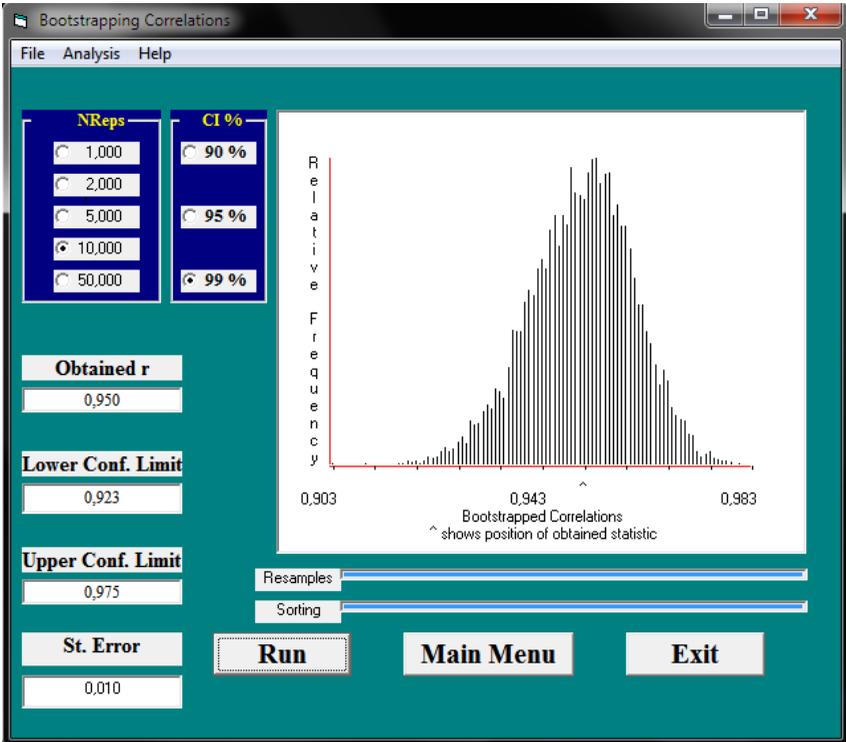


Figura 24: Captura de tela com resultado da análise de correlação por *bootstrap* entre peso do peixe e peso seco do tecido branquial na amostra (PSTA).

Por fim, a análise de correlação entre peso do peixe e PST (Figura 25) mostrou uma correlação linear de Pearson (r) igual a 0,966, com erro padrão de 0,009 e intervalo de confiança de 99% entre 0,943 e 0,986.

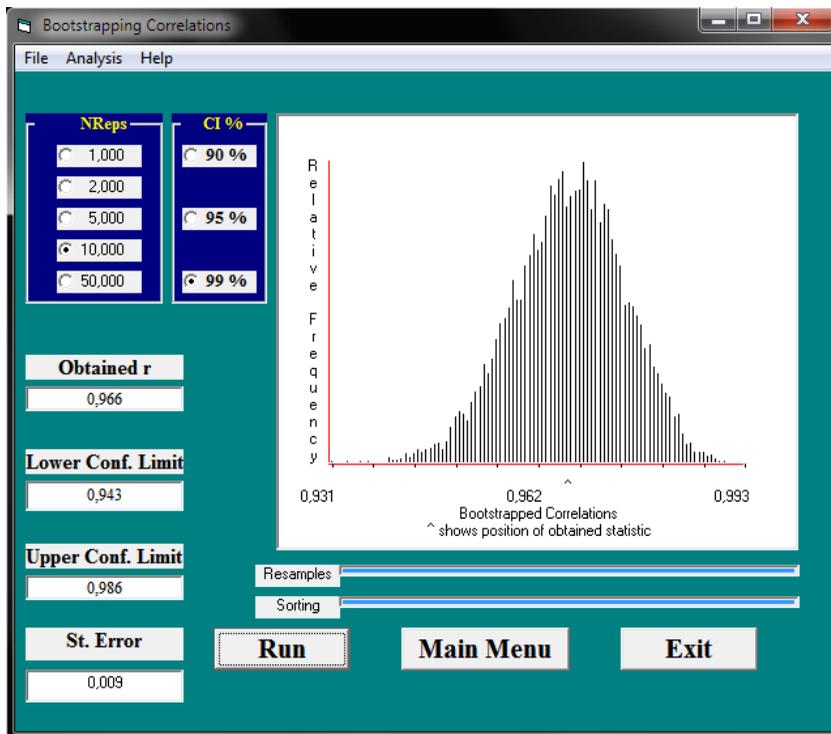


Figura 25: Captura de tela com resultado da análise de correlação por *bootstrap* entre peso do peixe e peso seco do tecido branquial (PST).

Os valores apresentados pelas análises de correlação por *bootstrap* comprovaram, com grande robustez, que a correlação entre a variável independente “peso do peixe” e cada uma das variáveis dependentes (PBH, PSO, PSTA e PST) é muito forte, chegando, como no caso da variável PBH, a ser quase perfeita em seu intervalo superior de confiança. Além disso, os valores de erro padrão foram consideravelmente baixos.

3.2.3. Modelagem

O método escolhido para modelagem foi o de *multivariate adaptive regression splines* (MARSplines), tanto em função de ter fornecido as melhores estimativas e menores resíduos, quanto devido à sua capacidade adaptativa, capaz de modelar funções que expressam diferentes variâncias em diferentes regiões do espaço preditivo.

3.2.3.1. Análise prévia

Na análise prévia, realizada utilizando-se todas as variáveis independentes, foi constatado que as variáveis preditoras de maior relevância foram “peso do peixe” e “altura”, ao passo que a variável ASR não foi referenciada em nenhuma das funções base (Tabela 11).

Tabela 11: Resultado da análise prévia para importância dos preditores. Florianópolis, junho de 2011.

Dependentes	Número de referência para cada preditor
	Referências (para funções base)
Peso do Peixe	2
ASR	0
Altura	3
Comprimento	1

Este resultado serviu como subsidio para excluir a utilização da variável ASR no modelo preditivo, especialmente devido à dificuldade de sua mensuração sem ferramentas adequadas e conhecimento específico.

Foi testado o uso concomitante das outras três variáveis como preditoras no modelo, porém levando-se em consideração a diferença mínima nas estimativas, optou-se por utilizar somente as duas variáveis com maior praticidade de aferição, portanto “peso do peixe” e “comprimento” foram as variáveis escolhidas.

3.2.3.2. MARSplines

O modelo desenvolvido através da MARSplines é constituído de quatro funções base e possui cinco termos, sendo três variáveis dependentes a serem estimadas (PBH, PSTA, PST) e duas variáveis predictoras independentes (Peso do Peixe, Comprimento). O limiar utilizado para complexidade da modelagem foi de 0,0005 e como forma de evitar sobre-ajuste (*overfitting*) foi aplicado ao modelo um processo de poda, para limitar a complexidade do modelo através da redução do número de funções base. As demais especificações do modelo podem ser vistas na Tabela 12.

Tabela 12: Especificações do modelo desenvolvido através da MARSplines. Florianópolis, junho de 2011.

Especificações do Modelo	Sumário do Modelo
	Value
Independentes	2
Dependentes	3
Nº de Termos	5
Nº de funções base	4
Ordem de interações	1
Penalidade	2,000000
Limiar	0,000500
Erro GCV	3,700247
Poda	Sim

A estatística descritiva da regressão (Tabela 13) mostra que os valores estimados pelo modelo, para cada uma das variáveis dependentes, apresentaram médias idênticas àquelas dos valores observados num nível de até seis casas decimais. Além disso, a média dos resíduos foi igual a 0 e valores de R^2 para todas as variáveis dependentes ficaram acima de 0,9.

Tabela 13: Distribuição do resultado da estatística descritiva da MARSplines. Florianópolis, junho de 2011.

Estatística da Regressão	Estatística da Regressão		
	PBH	PSTA	PST
Média (observados)	15,23238	0,004365	0,531104
Desvio Padrão (observados)	11,84761	0,003285	0,416003
Média (estimados)	15,23238	0,004365	0,531104
Desvio Padrão (estimados)	11,71301	0,003164	0,405287
Média (resíduo)	0,00000	0,000000	0,000000
Desvio Padrão (resíduo)	1,78080	0,000884	0,093811
R-quadrado	0,97741	0,927620	0,949147
R-quadrado ajustado	0,97639	0,924360	0,946856

Os coeficientes utilizados no modelo, bem como seus nós, podem ser vistos na Tabela 14, ao passo que na Tabela 15 observa-se a importância de cada preditor no modelo.

Tabela 14: Distribuição dos coeficientes do modelo desenvolvido através da MARSplines. Florianópolis, junho de 2011.*

Coeficientes, nós e funções base	Coeficiente do Modelo				
	Coeficientes PBH	Coeficientes PTSA	Coeficientes PST	Nós Peso do Peixe	Nós Comprimento
Raiz	13,45900	0,004426	0,522068		
Term. 1	1,52091	0,000346	0,043098		28,57500
Term. 2	-0,33583	-0,000016	-0,001973		28,57500
Term. 3	0,00832	0,000002	0,000368	330,0000	
Term. 4	-0,01987	-0,000012	-0,001423	330,0000	

* As células destacadas indicam funções base do tipo max (0, independente-nó), se não max (0, nó-independente)

Tabela 15: Importância dos preditores no modelo desenvolvido através da MARSplines. Florianópolis, junho de 2011.

Dependentes	Número de referências para cada preditor*
	Referências (para funções base)
Peso do Peixe	2
Comprimento	2

* Número de vezes que cada preditor é referenciado (usado)

As equações desenvolvidas pelo modelo para estimar cada uma das três variáveis dependentes (Figuras 26, 27 e 28) foram:

3.2.3.2.1. Peso da Brânquia Hidratada

Peso da Brânquia Hidratada = $1,34590046943541e+001 + 1,52091129068729e+000 * \max(0; \text{Comprimento} - 2,85750000000000e+001) - 3,35832682589586e-001 * \max(0; 2,85750000000000e+001 - \text{Comprimento}) + 8,32181912220444e-003 * \max(0; \text{Peso do Peixe} - 3,30000000000000e+002) - 1,98665391104146e-002 * \max(0; 3,30000000000000e+002 - \text{Peso do Peixe})$

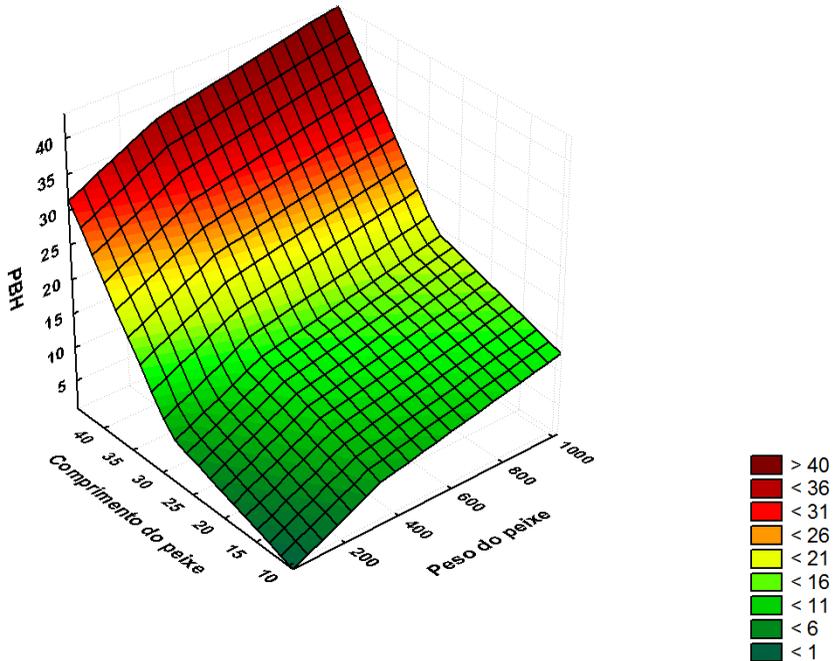


Figura 26: Plotagem da função para PBH, onde X = peso do peixe, Y = comprimento do peixe e Z = peso da brânquia hidratada.

3.2.3.2.2. Peso seco de tecido branquial em amostra de 1,5 ml

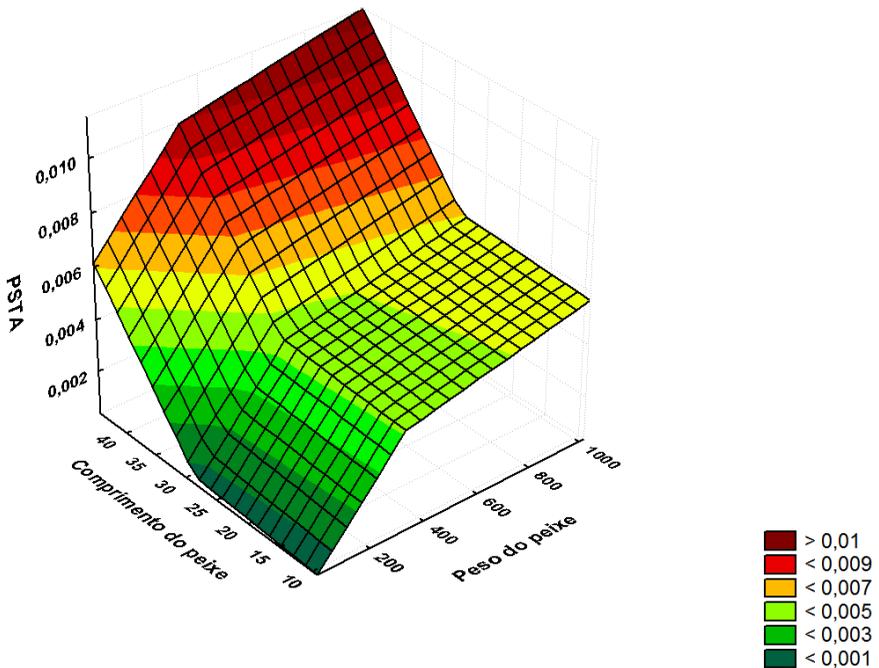
$$\begin{aligned} \text{PSTA} = & 4,42639564167902\text{e-}003 + 3,45853322900816\text{e-} \\ & 004 * \max(0; \text{Comprimento} - 2,85750000000000\text{e}+001) - \\ & 1,57333427311875\text{e-}005 * \max(0; 2,85750000000000\text{e}+001 - \\ & \text{Comprimento}) + 2,23249818114776\text{e-}006 * \max(0; \text{Peso do Peixe} - \\ & 3,30000000000000\text{e}+002) - 1,18525810550014\text{e-}005 * \max(0; \\ & 3,30000000000000\text{e}+002 - \text{Peso do Peixe}) \end{aligned}$$


Figura 27: Plotagem da função para PSTA, onde X = peso do peixe, Y = comprimento do peixe e Z = Peso seco de tecido branquial em amostra de 1,5 ml.

3.2.3.2.3. Peso seco total do tecido branquial

$$\begin{aligned} \text{PST} = & 5,22067839879816\text{e-}001 + 4,30982060784010\text{e-} \\ & 002 * \max(0; \text{Comprimento} - 2,85750000000000\text{e}+001) - \\ & 1,97271452915104\text{e-}003 * \max(0; 2,85750000000000\text{e}+001 - \\ & \text{Comprimento}) + 3,67934219484082\text{e-}004 * \max(0; \text{Peso do Peixe} - \\ & 3,30000000000000\text{e}+002) - 1,42338860538150\text{e-}003 * \max(0; \\ & 3,30000000000000\text{e}+002 - \text{Peso do Peixe}) \end{aligned}$$

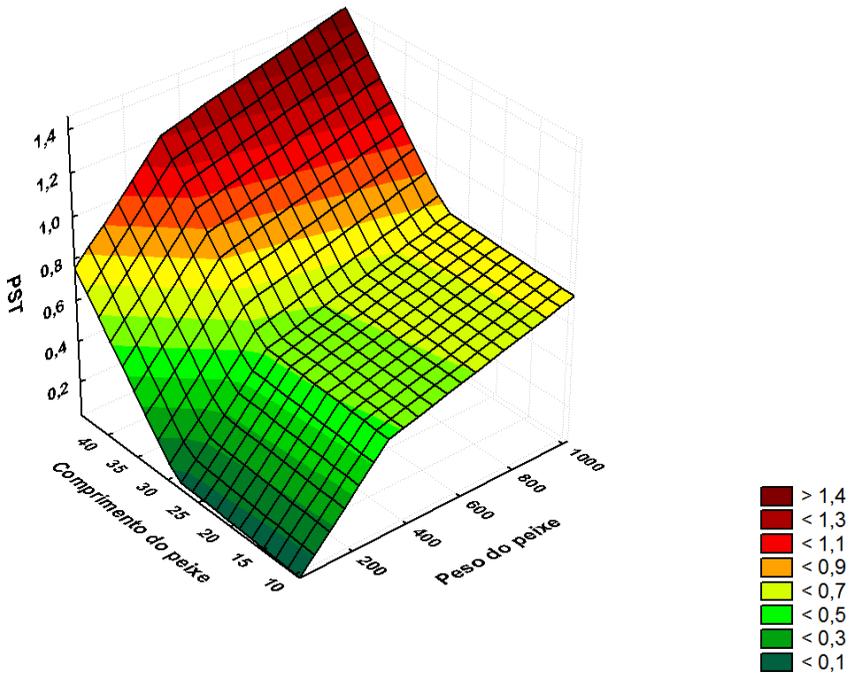


Figura 28: Plotagem da função para PST, onde X = peso do peixe, Y = comprimento do peixe e Z = Peso seco total do tecido branquial.

3.2.3.3. Análise residual e de qualidade de ajuste (*goodness of fit*)

Os resultados da estatística descritiva da regressão (Tabela 12) foram por si só bastante robustos. Porém, para assegurar a relevância e robustez do modelo, foi feita uma série de testes, dentre eles um teste de qualidade de ajuste e uma análise dos resíduos.

3.2.3.3.1. Peso da brânquia hidratada (PBH)

A análise da qualidade de ajuste resultou em um coeficiente de correlação de 0,99 (Tabela 16), com erro quadrático médio relativo igual a 0,007 e erro absoluto médio relativo igual a 0,067. O elevado valor do coeficiente de correlação, que chegou próximo à correlação perfeita, e os baixíssimos valores de erro relativo, demonstraram com bastante confiabilidade que o modelo é capaz de estimativas muito precisas.

Tabela 16: Resumo da qualidade de ajuste para a equação de peso da brânquia hidratada (PBH). Florianópolis, junho de 2011.

Itens calculados	Resumo da qualidade de ajuste
	Variável PBH
Erro quadrático médio	3,144147
Erro absoluto médio	1,041172
Erro quadrático médio relativo	0,007033
Erro absoluto médio relativo	0,066990
Coefficiente de correlação	0,988639

Nas Figuras 29 a. e 29 b., estão representadas, respectivamente, a curva de regressão entre os valores observados vs. estimados de PBH e a plotagem dos resíduos das estimativas de PBH.

A Figura 29 b. deixa bem evidente que os maiores resíduos se concentram nos maiores valores de PBH. Apenas dez estimativas apresentaram resíduos quadráticos maiores do que 10, sendo todas elas relativas à peixes com mais de 500 g de peso, o que reflete, então, a redução de correlação anteriormente verificada para esta faixa de peso do peixe.

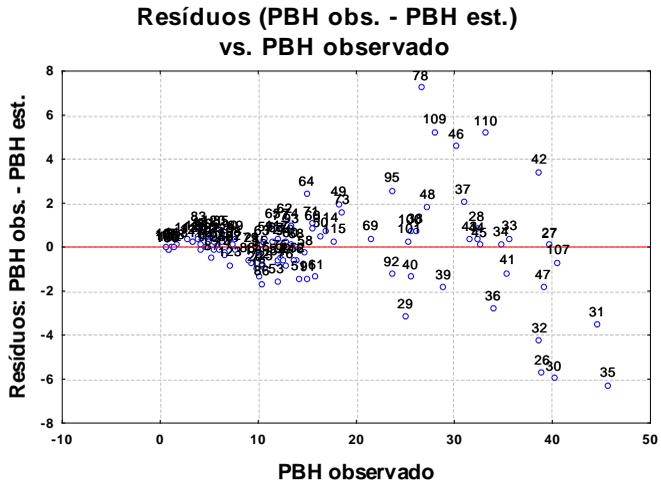
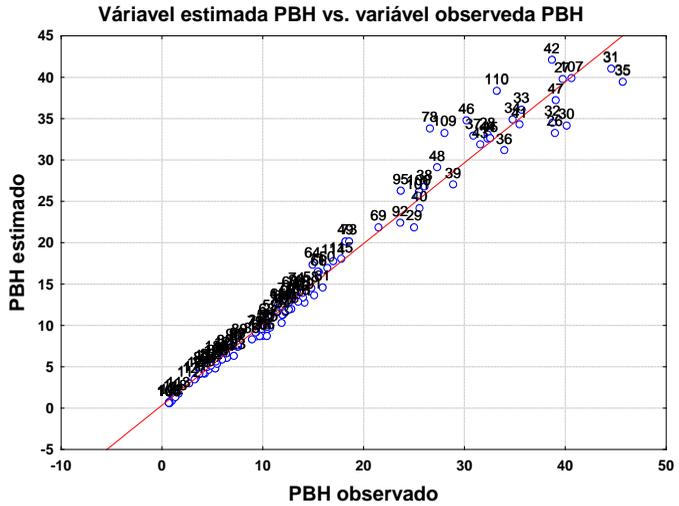


Figura 29: a) Curva de regressão entre os valores observados vs. estimados de peso da brânquia hidratada (PBH) b) plotagem dos resíduos das estimativas de peso da brânquia hidratada (PBH).

3.2.3.3.2. Peso seco do tecido branquial na amostra (PSTA)

A análise da qualidade de ajuste para a equação de PSTA resultou em um coeficiente de correlação de 0,96 (Tabela 17), com erro quadrático médio relativo igual a 0,049 e erro absoluto médio relativo igual a 0,169.

Apesar do baixo valor de erro quadrático médio (0,000001), os valores de erros relativos demonstraram que o erro de estimativa para PSTA foram maiores do que aqueles encontrados para PBH, diferença que também foi verificada no menor valor do coeficiente de correlação.

Ainda assim os valores de erros relativos ficaram dentro das casas decimais (não ultrapassando 0,2), sendo considerados muito baixos, especialmente levando-se em consideração a escala de valores utilizada na aferição e o fato de serem dados de origem biológica.

Tabela 17: Resumo da qualidade de ajuste para a equação de peso seco do tecido branquial na amostra (PSTA). Florianópolis, junho de 2011.

Itens calculados	Resumo da qualidade de ajuste
	Variável PSTA
Erro quadrático médio	0,000001
Erro absoluto médio	0,000621
Erro quadrático médio relativo	0,049035
Erro absoluto médio relativo	0,169037
Coefficiente de correlação	0,963130

Na curva de regressão entre os valores médios observados e os valores estimados de PSTA (Figura 30 a.) foi possível notar que a dispersão dos dados de PSTA é um pouco maior em comparação à dispersão dos dados de PBH. O mesmo ocorre em relação à dispersão dos resíduos (Figura 30 b.)

Entretanto, os resíduos na Figura 30 b. já não colocaram tanto em evidência a influência dos peixes acima de 500 g na dispersão.

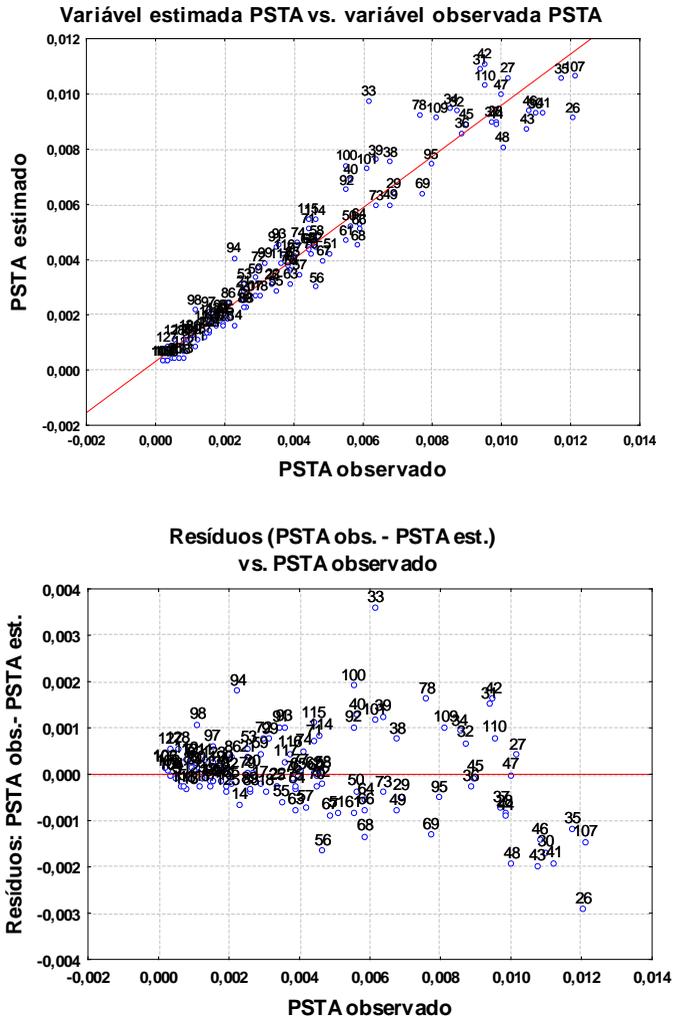


Figura 30: a) Curva de regressão entre os valores observados vs. estimados de peso seco do tecido branquial na amostra (PSTA) b) plotagem dos resíduos das estimativas de peso seco do tecido branquial na amostra (PSTA).

3.2.3.3.3. Peso seco do tecido branquial (PST)

A análise da qualidade de ajuste para a equação de PST resultou em um coeficiente de correlação de 0,97 (Tabela 18), com erro quadrático médio relativo igual a 0,016 e erro absoluto médio relativo igual a 0,095.

Pode-se constatar que a correlação bastante elevada resultou em baixos valores de erro. Novamente, apesar de maiores do que erro quadrático médio e absoluto médio, os valores de erros relativos foram bem baixos, mantendo-se abaixo de 0,1, ao passo que o coeficiente de correlação foi bastante elevado.

Tabela 18: Resumo da qualidade de ajuste para a equação de peso seco do tecido branquial (PST). Florianópolis, junho de 2011.

Itens calculados	Resumo da qualidade de ajuste
	Variável PST
Erro quadrático médio	0,008725
Erro absoluto médio	0,051921
Erro quadrático médio relativo	0,015566
Erro absoluto médio relativo	0,094864
Coefficiente de correlação	0,974242

A curva de regressão (Figura 31 a.) e a dispersão dos resíduos (Figura 31 b.) mostram que a maior dispersão dos dados concentrou-se, novamente, nos valores mais elevados. Ela foi mais evidente do que a observada na distribuição dos valores de PSTA ressaltando, assim, a influência da diferença de correlação apresentada pelos peixes com peso acima de 500 g.

Apesar da dispersão aparentemente forte dos resíduos nos dados de peso acima de 500 g, apenas nove pontos apresentaram resíduos maiores do que 0,1.

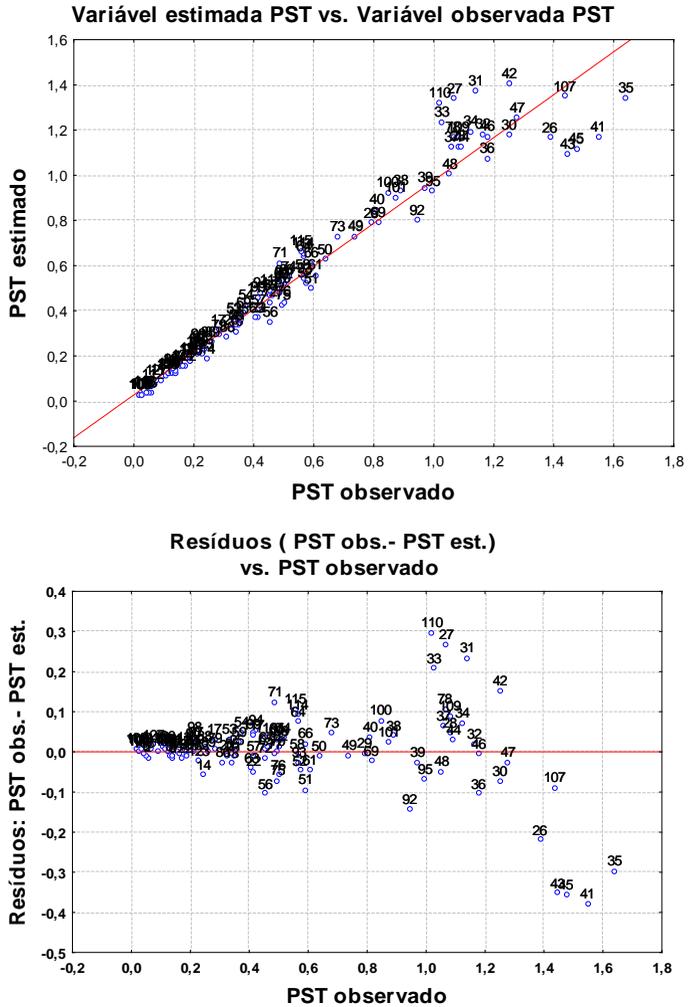


Figura 31: a) Curva de regressão entre os valores observados vs. estimados de peso seco do tecido branquial (PST) b) plotagem dos resíduos das estimativas de peso seco do tecido branquial (PST).

3.2.3.4. Significância dos parâmetros estimados

Como último passo de validação do modelo, foi testada a significância dos parâmetros estimados, comparando os valores estimados com os valores observados.

3.2.3.4.1. Peso da brânquia hidratada (PBH)

O teste T de Student para amostras dependentes, assim como a análise de qui-quadrado ($\chi^2 = 13,7$), resultaram em $p = 1,0$ na comparação entre os valores estimados e observados de PBH, demonstrando igualdade estatística entre eles. Além disso, a avaliação do tamanho amostral necessário para encontrar uma diferença entre as duas médias, com intervalo de confiança de 99% e poder estatístico de 95%, foi igual a infinito, o que reforça a comprovação de igualdade entre as variáveis. A Figura 32 deixa bem evidente a precisão da estimativa através dos intervalos de confiança idênticos.

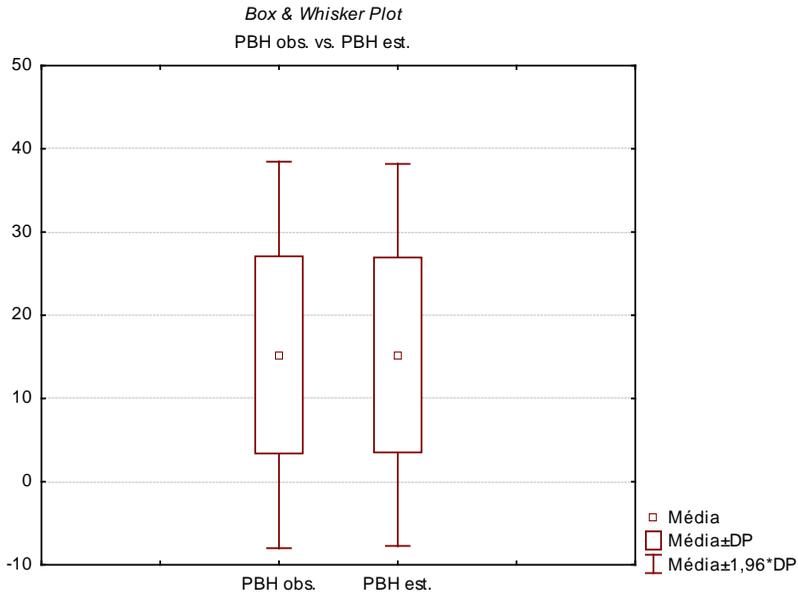


Figura 32: Resultado do teste T de Student para amostras pareadas para PBH. obs. = observados; est. = estimados; DP = desvio padrão.

3.2.3.4.2. Peso seco de tecido branquial na amostra (PSTA)

Novamente, ambas as análises resultaram em $p = 1,0$, ou seja, valores estatisticamente iguais para os dados observados e estimados. O valor de qui-quadrado foi igual 0,018 e o resultado do teste T de Student, representado na Figura 33, mostrou uma variação mínima entre o intervalo de confiança das duas variáveis. A avaliação do tamanho amostral necessário para encontrar uma diferença entre as duas médias, novamente foi igual a infinito, mesmo intervalo de confiança de 99% e poder estatístico de 95%.

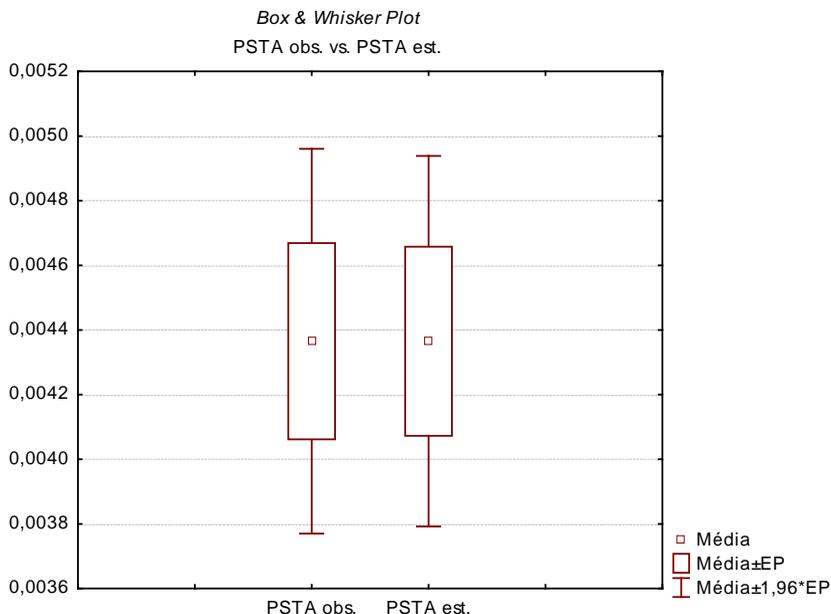


Figura 33: Resultado do teste T de Student para amostras pareadas para PSTA. obs. = observados; est. = estimados; EP = erro padrão.

3.2.3.4.3. Peso seco total de tecido branquial (PST)

A análise de qui-quadrado ($\chi^2 = 0,98$) e o teste T de Student para amostras dependentes resultaram em $p = 1,0$, mais uma vez demonstrou igualdade estatística entre valores observados e estimados. A avaliação do tamanho amostral necessário para encontrar uma diferença entre as duas médias foi, também, igual a infinito, considerando-se um intervalo de confiança de 99% e poder estatístico de 95%. A Figura 34 mostra que ambas as variáveis foram estatisticamente iguais, apresentando uma variação mínima nos intervalos de confiança.

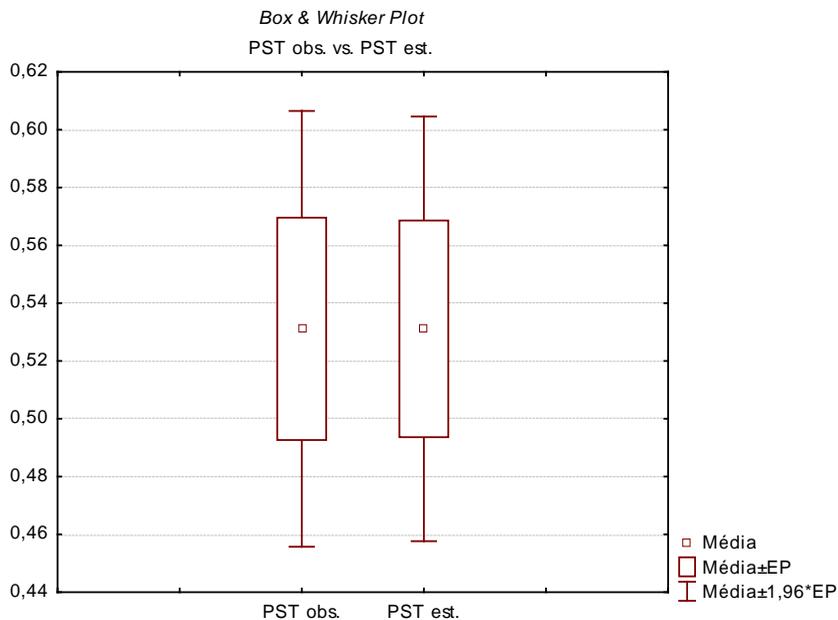


Figura 34: Resultado do teste T de Student para amostras pareadas para PST. obs. = observados; est. = estimados; EP = erro padrão.

DISCUSSÃO

3.1. Uma questão de lógica

Uma vez que este trabalho aborda a questão da lógica, é pertinente discutir um pouco sobre este tema.

A lógica é uma ciência bastante complexa que abrange todas as áreas do conhecimento e a própria ciência em si. Existem diversas correntes de pensamento e diferentes abordagens para a questão da lógica, porém, pode-se estabelecer duas abordagens principais: a lógica clássica e a anti-clássica (COSTA, 2008). Não cabe a este trabalho discutir os detalhes das diferentes formas de lógica anti-clássica, entretanto, convém refletir sobre a lógica clássica e os princípios de indução e dedução.

Dedução é o processo de inferir partindo do geral (universal) para o particular, ou seja, conhecendo-se previamente uma lei geral, aplica-se-a para algo particular. Esse tipo de raciocínio é mais aceito pela lógica clássica, sendo considerado como “válida”, entretanto não oferece nenhum conhecimento novo, posto que se baseia em algo previamente conhecido.

A indução, por sua vez, é o processo de inferir sobre o geral partindo de algo particular, ou seja, generalizar para o todo algo observado em uma amostra. Devido à sua natureza, a indução é muito mais susceptível de refutação, sendo muitas vezes considerada pela lógica clássica como “não válida”. Porém, “não haveria ciência empírica se os cientistas procurassem empregar unicamente formas válidas de inferências” (COSTA, 2008).

Chega-se, então, no problema da indução ou problema de Hume, onde não se pode justificar que determinado comportamento irá ocorrer unicamente pelo fato de ele ter ocorrido anteriormente (HUME, 1748). Felizmente, existem vários tipos de inferências indutivas e nem todos estão fadados à invalidade, entre eles pode-se citar a inferência estatística e a inferência probabilística.

Entende-se por inferência estatística a estimativa de parâmetros, o teste de hipóteses e a teoria da decisão. Ela está presente em praticamente todas as ciências empíricas, bem como na tecnologia, tanto em suas formas puras como aplicadas e, conforme discute Costa (2008) “constitui a maior evidência de que existe uma lógica indutiva”

A inferência probabilística considera uma relação de probabilidade entre a conjunção das premissas e a conclusão, ou seja, ela aceita um argumento como correto caso a probabilidade de ele ocorrer seja alta, sem, entretanto, afirmar que ele é absoluto ou imutável.

Assim, o principal diferencial das inferências estatística e probabilística reside no fato de ambas considerarem um valor de probabilidade para a ocorrência de um evento, ao invés da dicotomia do sim ou não. Segundo Costa (2008), ao atribuir-se uma probabilidade, “deixar-se-ia o problema da inferência propriamente dita de lado, pois, dada a probabilidade de um argumento, é uma questão pessoal e pragmática se alguém quer ou não recorrer a ele”.

Com base nestes argumentos, pode-se, portanto, considerar pertinente a abordagem lógica utilizada neste trabalho.

3.2. Acerca do processo de coleta

Os principais fatores limitantes da amostragem foram a disponibilidade de animais em todas as faixas de peso, especialmente para as repetições com peso homogêneo, e a disponibilidade de pessoal para processar o material. Porém, mesmo assim, buscou-se uma amostra suficientemente grande e com animais que representassem as diversas classes de tamanho.

Neste trabalho deparou-se, também, com as dificuldades, experimentadas pela maioria dos laboratórios no Brasil, nas coletas de campo, especialmente em relação à equipamentos de mensuração. Sendo assim, buscou-se ao máximo não depender das condições de coleta para as aferições. A variável “peso do peixe”, porém, foi aferida em campo, sendo sua qualidade a principal fonte de variâncias não intrínsecas.

3.3. Acerca das variáveis aferidas

É pertinente notar que o grande número de variáveis (dependentes e independentes) aferidas durante o processo de coleta teve o intuito de garantir a melhor exploração possível dos dados.

Considerou-se o peso seco total de tecido branquial (PST) como a variável dependente que melhor representa o espaço físico disponível para parasitismo. Ela foi escolhida por ser uma medida que reflete a quantidade total de brânquia de forma direta e confiável, pois não sofre a influência de variações na hidratação.

Entretanto, mesmo variáveis não diretamente relacionadas com o parasitismo, como o peso seco ósseo, foram aferidas para o caso de eventualmente serem pertinentes em uma estimativa indireta.

Da mesma forma, foram aferidas diversas variáveis independentes, mas somente o “peso do peixe” e “comprimento do peixe” foram consideradas as variáveis preditoras de maior relevância. Esta escolha baseou-se na proposta do trabalho, que considera importante a possibilidade de aplicação do modelo em condições de precariedade de instrumentos de aferição, como em situações de coleta a campo, e, também, em laboratórios com profissionais pouco qualificados. Por isso, foi atribuído a elas um peso maior do que para as variáveis de difícil mensuração (como a “área superficial relativa”) e variáveis com possíveis imprecisões na aferição manual (como a variável “altura”, cujo ponto ideal de medição varia conforme o peixe).

Outra questão cuja discussão é relevante diz respeito aos volumes utilizados como padrão para a variável PSTA.

O volume de 1,5 ml, utilizado como padrão para amostra neste trabalho, teve como base 10 repetições de contagens na câmara de McMaster e foi escolhido devido à capacidade prática de aferição da variável PSTA, uma vez que o volume de tecido branquial em uma amostra de 0,15 ml, além de muito influenciado pelo diâmetro da ponteira (que impede a passagem de pedaços maiores de raspado de brânquias), é virtualmente impossível de ser mensurado, visto que os valores de PSTA num volume 10 vezes maior já demandavam uma balança de precisão com 4 casas decimais.

Já o volume de 200 ml, utilizado como padrão para solução fixadora, foi escolhido por ser um volume adequado mesmo para potes pequenos, além de ser capaz de envolver e fixar, com eficiência, brânquias de animais com mais de 1 kg.

3.4. Acerca das regressões e análise das correlações

Apesar de não terem sido utilizadas na modelagem final, as regressões e análises realizadas com os dados foram essenciais para melhor compreender o comportamento das variáveis.

Os resultados obtidos nas análises das correlações foram responsáveis pelo suporte necessário para assegurar a validade da modelagem e auxiliar na escolha da ferramenta mais adequada para modelar os dados.

3.5. Acerca da modelagem

A escolha da *Multivariate adaptive regression splines* (MARSplines) mostrou-se bem adequada para modelagem dos dados utilizados neste trabalho. Os valores de erro quadrático médio relativo para as variáveis estimadas (Tabela 15, 16 e 17), os elevados valores de correlação (Tabela 15, 16 e 17) e os diversos testes de aderência e significância (páginas 89 a 98), mostram que sua capacidade adaptativa e seus mecanismos de controle de sobre-ajuste, são eficientes em fornecer um modelo bastante preciso e robusto.

É necessário destacar, ainda, que a MARSplines apresenta uma série de vantagens entre as quais pode-se citar:

- Seus modelos são mais flexíveis do que modelos de regressões lineares e mais simples do que modelos de neural networks.
- Não faz qualquer pressuposto quanto à relação funcional das variáveis dependentes e independentes
- É capaz de adaptar-se à diferentes correlações na curva de dados, uma vez que divide o espaço amostral em partições, cada uma com sua equação de regressão própria
- Seus modelos não requerem preparação prévia dos dados (contendo até mesmo o efeito de *outliers*)
- Consegue lidar com grandes conjuntos de dados
- Apresenta baixa tendenciosidade.

Além disso, Sikka et al. (2010) em estudo comparando o uso de diferentes técnicas de modelagem para estimativa de pontos de função, relataram MARSplines como sendo aquela que apresentou os menores valores de erro médio relativo e os melhores valores de correlação.

CONCLUSÃO

Este trabalho buscou, desde seu princípio, ser bastante criterioso em todas as etapas do processo, primando pela máxima rigidez, especialmente nos processos das aferições, de forma a minimizar qualquer erro sistemático. Buscou, também, detalhar da melhor forma possível todo o processo de desenvolvimento do modelo preditivo, com o intuito de garantir a reprodutibilidade tanto para a mesma espécie quanto para diferentes espécies de peixe.

Além disso, houve um grande esforço para testar diferentes tipos de modelagem e no sentido de assegurar a robustez das correlações e do modelo preditivo. O desenvolvimento de vários testes permitiu garantir a utilização do modelo mais preciso e confiável possível.

O resultado deste esforço e rigidez foi um modelo preditivo com grande exatidão e confiabilidade, onde todos os coeficientes de correlação ficaram acima de 0,9 e todos os erros quadráticos relativos ficaram abaixo de 0,05, tendo como únicas exigências os valores de peso e comprimento do peixe.

A partir do modelo preditivo foi possível desenvolver o software, batizado de *ProtoCount*, contribuindo com a quantificação de protozoários parasitos em tilápias do Nilo no Brasil. Além disso, os protocolos de coleta e aferição de variáveis provaram ser eficientes, abrindo caminho para o desenvolvimento de modelos preditivos para outras espécies de peixes.

AGRADECIMENTOS

Os autores agradecem a CAPES e CNPq pelo auxílio e bolsas de Mestrado e Produtividade em Pesquisa (302493/2010-7), e também a Roberto Hoppe (Fundação 25 de Julho) e Juan Esquivel (Piscicultura Panamá) pela doação dos animais e colaboração na captura.

REFERÊNCIAS

- AMANN, R. I.; LUDWIG, W.; SCHLEIFER, K. H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. **Microbiological Reviews**, v. 59, n. 1, p. 143-69, 1995.
- BACHMANN, F.; GREINERT, J. A.; BERTELLI, P. W.; FILHO, H. H. da S.; LARA, N. O. T. de; GHIRALDELLI, L.; MARTINS, M. L. Parasitofauna de *Pimelodus maculatus* (Osteichthyes: Pimelodidae) do rio Itajaí-Açu em Blumenau, Estado de Santa Catarina, Brasil. **Acta Scientiarum. Biological Sciences**, v. 29, n. 1, p. 109-114, 2007.
- BOSCARDIN, N. R. Produção Aqüícola. In: OSTRENSKY, A.; BORGHETTI, J.R.; SOTO, D. (organizadores). **Aqüicultura no Brasil: o desafio é crescer**. Brasília, DF: 2008, p.27-72.
- BRACCINI, G. L.; VARGAS, L.; RIBEIRO, R. P.; DIGMAYER, L. A. F. M. Ectoparasitos de tilápia do Nilo (*Oreochromis niloticus*), das linhagens Chitralada e GIFT, em diferentes densidades e alimentadas com dois níveis de proteína. **Acta Scientiarum. Animal Sciences**, v. 29, n. 4, p. 441-448, 2007.
- CAVICHIOLO, F.; VARGAS, L.; RIBEIRO, R. P.; MOREIRA, H. L. M.; LEONARDO, J. M. Níveis de suplementação de vitamina C na ração sobre a ocorrência de ectoparasitas, sobrevivência e biomassa em alevinos de tilápia do Nilo (*Oreochromis niloticus* L.). **Acta Scientiarum**, v. 24, n. 4, p. 957-964, 2002.
- COSTA, N. C. A. da. **Lógica Indutiva e Probabilidade**. 3ª ed. São Paulo: Hucitec, 2008.
- DAIMS, H.; WAGNER, M. Quantification of uncultured microorganisms by fluorescence microscopy and digital image analysis. **Applied microbiology and biotechnology**, v. 75, n. 2, p. 237-48, 2007.
- EIRAS, J C; CHEMES, S. B.; CASABIANCA, O. A.; TAKEMOTO, R. M. ; ROSSI, L. M. ; PAVANELLI, G. C. An unusual infection of *Henneguya* sp. (Myxozoa , Myxosporea , Myxobolidae) in the gills of *Hoplosternum littorale* Hancock , 1928 (Osteichthyes , Callichthyidae) from Middle Paraná River , Santa Fe , Argentina. **Canadian Journal of Zoology**, v. 28, p. 22-26, 2008.

EIRAS, J.C.; ABREU, P. C.; ROBALDO, R.; PEREIRA JÚNIOR, J. *Myxobolus platanus* n. sp. (Myxosporea, Myxobolidae), a parasite of *Mugil platanus* Günther, 1880 (Osteichthyes, Mugilidae) from Lagoa dos Patos, RS, Brazil. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 59, n. 4, p. 895-898, 2007.

EPAGRI, Empresa de pesquisa Agropecuária e Extensão Rural de Santa Catarina. **Desempenho da pesca e da Aqüicultura – Ano 2005-2006**. Instituto Cepa - Centro de Socioeconômica e Planejamento Agrícola, 2008. Disponível em: <http://cepa.epagri.sc.gov.br/Publicacoes/sintese_2006/pesca_2006.pdf> . Acesso em: 20 abr 2008.

FAO, Food and Agriculture Organization Fisheries Department. **The State of World Fisheries and Aquaculture**. Rome, FAO, 2010.

FRIEDMAN, J. Multivariate adaptive regression splines. **Annals of Statistics**, v. 19, p. 1-141, 1991.

GHIRALDELLI, L.; MARTINS, M.L.; JERÔNIMO, G.T.; YAMASHITA, M.M.; ADAMANTE, W.B. Ectoparasites communities from *Oreochromis niloticus* cultivated in the State of Santa Catarina, Brazil. **Journal of Fisheries and Aquatic Science**, New York, v. 1, n. 2, p:181-190, 2006

HIGUCHI, R.; DOLLINGER, G.; WALSH, P.S.; GRIFFITH R. Simultaneous amplification and detection of specific DNA sequences. **Biotechnology**, New York, v. 10, p. 413–417, 1992.

HUME, D. Investigação sobre o Entendimento Humano. 1748. In: **Os Pensadores**, 2ª ed., Abril Cultural, São Paulo, 1980.

JERÔNIMO, G.T.; SPECK, G.M.; CECHINEL, M.M.; GONÇALVES, E.L.T. E MARTINS, M.L. Seasonal Variation on the Ectoparasitic Communities of Nile Tilapia Cultured in Three Regions in Southern Brazil. **Brazilian Journal of Biology**, V. 71, n. 2, p. 1-9, 2011.

LEONARDO, J. M. L. O.; PEREIRA, J. V.; KRAJEVIESKI, M. E. Ocorrência de ectoparasitas e estacionalidade em alevinos de tilápia-do-nilo (*Oreochromis niloticus*) após a reversão sexual, na região noroeste do Paraná. **Iniciação Científica CESUMAR**, p. 185-191, 2006.

LUDWIG W., SCHLEIFER K.H. How quantitative is quantitative PCR with respect to cell counts. **Systematic and Applied Microbiology**, v. 23, p. 556-562, 2000.

MARTINS, MAURÍCIO LATERÇA; ONAKA, E. M.; MORAES, F. R. D.; et al. Recent studies on parasitic infections of freshwater cultivated fish in the state of São Paulo , Brazil. **Acta Scientiarum**, p. 981-985, 2002.

MAYHEW, T. M.; GUNDERSEN, H. J. If you assume, you can make an ass out of u and me': a decade of the disector for stereological counting of particles in 3D space. **Journal of anatomy**, v. 188, p. 1-15, 1996.

MONIS, P. T.; GIGLIO, S.; KEEGAN, A. R.; ANDREW THOMPSON, R. C. Emerging technologies for the detection and genetic characterization of protozoan parasites. **Trends in parasitology**, v. 21, n. 7, p. 340-6, 2005.

RAMSING, N. B.; FOSSING, H.; FERDELMAN, T. G.; ANDERSEN, F.; THAMDRUP, B. Distribution of bacterial populations in a stratified fjord (mariager fjord, denmark) quantified by in situ hybridization and related to chemical gradients in the water column. **Applied and environmental microbiology**, v. 62, n. 10, p. 3915, 1996.

RANZANI-PAIVA, M. J. T.; FELIZARDO, N. N.; LUQUE, J. L. Parasitological and hematological analysis of Nile tilapia *Oreochromis niloticus* Linnaeus , 1757 from Guarapiranga reservoir , São Paulo State, Brazil. **Acta Scientiarum. Biological Sciences**, v. 29, p. 231-237, 2005.

RAY, S. M. A culture technique for the diagnosis of infections with *Dermocystidium marinum*. Mackin, Owen, and Collier, in oysters. **Science**, v. 166, p. 360-361, 1952.

_____. A review of the culture method for detecting *Dermocystidium marinum*, with suggested modifications and precautions. **Proceedings of the National Shellfisheries Association**, v. 54, p. 55-69, 1966.

REMACHA-TRIVIÑO, A.; BORSAY-HOROWITZ, D.; DUNGAN, C.; GUAL-ARNAU, X.; GÓMEZ-LEON, J.; VILLAMIL, L.; GÓMEZ-CHIARRI, M. Numerical Quantification of *Perkinsus marinus* in the American Oyster *Crassostrea virginica* (Gmelin, 1791) (Mollusca: Bivalvia) by Modern Stereology. **Journal of Parasitology**, v. 94, n. 1, p. 125-136, 2008.

SIKKA, G., KAUR, A. & UDDIN, M. Estimating Function points: Using Machine Learning and Regression Models. In: INTERNATIONAL CONFERENCE ON EDUCATION TECHNOLOGY AND COMPUTER, 2, 2010, Shanghai, p. 52-56.

TAVARES-DIAS, M.; BRITO, M. L. S.; LEMOS, J. R. G. Protozoários e metazoários parasitos do cardinal *Paracheirodon axelrodi* Schultz, 1956 (Characidae), peixe ornamental proveniente de exportador de Manaus, Estado do Amazonas, Brasil. **Acta Scientiarum. Biological Sciences**, v. 31, n. 1, p. 23-28, 2009a.

TAVARES-DIAS, M.; MARTINS, MAURÍCIO L.; MORAES, FLÁVIO R. Fauna parasitária de peixes oriundos de “pesque-pague” do município de Franca, São Paulo, Brasil. I. Protozoários. **Revista Brasileira de Zoologia**, v. 18, p. 67 - 79, 2001.

TAVARES-DIAS, M.; MARTINS, MAURÍCIO LATERÇA; MORAES, FLÁVIO RUAS; KRONKA, N. Fator de condição e relação hepato e esplenossomática em teleósteos de água doce naturalmente parasitados. **Acta Scientiarum**, v. 22, n. 2, p. 533-537, 2000.

TAVARES-DIAS, M.; RAPHAEL, J.; LEMOS, G.; MARTINS, L.; JERÔNIMO, G. T. Metazoan and protozoan parasites of freshwater ornamental fish from Brazil. In: TAVARES-DIAS, M. (Organizador). **Manejo e sanidade de peixes em cultivo**. Macapá: Embrapa Amapá, 2009b, p. 469-494.

CAPÍTULO 2

PROTOCOLOUNT: UMA FERRAMENTA INOVADORA PARA QUANTIFICAR PARASITOS PROTOZOÁRIOS EM BRÂNQUIAS DE PEIXES¹

ABSTRACT

Protozoan parasites are a major cause of cultured fish mortality, but there are surprisingly few data with quantitative measurements, especially in Brazil, where most studies are qualitative. There are few reliable methods for quantifying parasite protozoa, but either they require expensive equipments or highly specialized professionals. Both cases still far from the reality of most fish pathology laboratories in Brazil. With that in mind, and looking forward to the development of quantitative parameters associated with disease outbreaks, we developed *ProtoCount* software that uses a predictive model for protozoan parasite estimation on cultured fish gills. *ProtoCount* was developed using JDK 6 JAVA language, with Netbeans 6.9.1 for integrated development environment (IDE) and JAVA Swing for interface. It uses three simple data (fish weight, fish width and the number of protozoa observed in a sample) to estimate the total protozoa intensity on gills. The predictive model inside *ProtoCount* uses MARSplines modeling equations to predict “gill’s tissue total dry weight” and “gill’s tissue dry weight on sample” data, which is then combined to protozoan counting data from a sample to infer over the total protozoa intensity on gills. *ProtoCount*’s only requirements are measuring fish’s weight and width, scraping gill’s tissue into a 200 ml formalin solution and counting protozoa mean burden on 1,5 ml samples using MacMaster’s chamber. Each 1,5 ml sample is composed by the mean protozoa on ten 0,15 ml homogenized samples. Thus, *ProtoCount* stands as a practical, yet very precise, tool for quantifying protozoan parasites on gills.

Keywords: Aquaculture. Math modeling. Protozoa. Nile tilapia. Software.

RESUMO

Protozoários parasitos são grandes causadores de mortalidade em peixes cultivados, porém, existe um número pequeno de trabalhos com análises quantitativas sobre o assunto no Brasil. São conhecidos métodos para quantificar protozoários parasitos, entretanto eles geralmente requerem equipamentos caros e/ou mão de obra especializada. Almejando uma melhoria na confiabilidade e precisão dos trabalhos quantitativos com protozoários parasitos de peixes, este trabalho objetivou desenvolver um software, que utiliza um modelo preditivo para estimar protozoários parasitos em brânquias de tilápia do Nilo. No desenvolvimento do software, que foi batizado de *ProtoCount*, foi utilizado *Java Development Kit* (JDK) 6, tendo Netbeans 6.9.1 como ambiente de desenvolvimento e Java Swing como API de interface. O modelo preditivo embutido no *ProtoCount* utiliza equações modeladas por MARSplines para prever, a partir do peso e comprimento do peixe, os valores de “peso seco total de tecido branquial” e “peso seco de tecido branquial em amostra de 1,5ml”. Eles são, então, combinados aos dados da contagem de protozoários na amostra para inferir sobre a quantidade total de protozoários nas brânquias. Os únicos requisitos para utilização do *ProtoCount* são a aferição do peso e comprimento do peixe, a raspagem do tecido branquial em 200 ml de solução formalina e a contagem, numa câmara de McMaster, da média de protozoários em amostras de 1,5 ml. Os resultados obtidos com o programa *ProtoCount* demonstraram que ele é uma ferramenta prática, e ao mesmo tempo precisa, para quantificar protozoários parasitos em brânquias.

Palavras-chave: Aqüicultura. Modelagem matemática. Protozoários. Tilápia do Nilo. Software.

INTRODUÇÃO

Os protozoários são importantes agentes etiológicos, responsáveis por grandes mortalidades e prejuízos econômicos em cultivos de peixes ao redor do mundo (Buchmann and Bresciani, 1997; Burreson and Ragone-Calvo, 1996 Crosby And Roberts, 1990; Martins et al., 2001). Porém, ao contrário do que ocorre em muitos lugares do mundo, o controle sanitário de parasitos de peixes no Brasil ainda é bastante incipiente (Tavares-Dias *et al.*, 2009; Leonardo *et al.*, 2006), sendo os protozoários um dos grupos de patógenos cujos estudos são menos desenvolvidos.

A maioria dos trabalhos com protozoários de peixes no Brasil ainda são exclusivamente qualitativos (Martins *et al.*, 2002; Ranzani-Paiva et al., 2005; Eiras et al., 2008; Eiras *et al.*, 2007), avaliando somente a presença do patógeno. Parte disso deve-se ao fato de a quantificação precisa de protozoários ainda apresentar desafios metodológicos (Daims and Wagner, 2007), especialmente em amostras complexas, como a brânquia de um peixe. A outra parte deve-se à carência de profissionais especializados e, principalmente, de laboratórios equipados. Boa parte dos métodos de quantificação de protozoários requer equipamentos sofisticados e caros, como o método molecular de PCR *real time*, ou mão de obra especializada, como nos métodos estereológicos (Higuchi et al., 1992; Daims and Wagner, 2007; Mayhew and Gundersen, 1996; Monis et al, 2005; Remacha-Triviño et al., 2008;).

Entretanto, é preciso ressaltar que estudos quantitativos são de grande importância para a pesquisa da sanidade em organismos aquáticos, sendo necessários para o estabelecimento de parâmetros de avaliação do grau de enfermidade.

Pensando na importância do acompanhamento sanitário em cultivos, como forma de prevenir o surgimento de doenças e evitar potenciais prejuízos, e levando em consideração as dificuldades da aplicação de métodos de quantificação sofisticados, objetivou-se desenvolver ferramentas para quantificação de protozoários capazes de estimativas confiáveis e precisas, mesmo em condições de precariedade de equipamentos e mão de obra qualificada. A melhor alternativa

C = comprimento do peixe; QPA = Quantidade de protozoários contados na câmara de McMaster em amostra de 1,5 ml; QTP = Quantidade total de protozoários nas brânquias do peixe; MARSplines = *Multivariate adaptive regression splines*; P = peso do peixe; PBH = Peso da brânquia hidratada; PSO = Peso seco ósseo; PST = Peso seco total do tecido branquial; PSTA = Peso seco de tecido em uma amostra de 1,5 ml de solução formalina 10% com raspado de brânquias; VA = volume da amostra quantificada em câmara de McMaster; VT = volume total da solução de formalina contendo raspado de tecido branquial

encontrada, durante o desenvolvimento deste estudo, foi a criação de um software computacional de fácil utilização e que pudesse ser amplamente disponibilizado. Com ele seria possível estimar a quantidade total de protozoários em brânquias de peixes, através de um método de estimação (Costa, 2008), por extrapolação de amostragem probabilística e inferência estatística.

Este estudo tomou como base o método desenvolvido por Tavares-Dias et al. (2001), que utiliza uma câmara de McMaster para contar os protozoários presentes em uma amostra com 1,5 ml de raspado de brânquias em solução formalina, extrapolando o valor encontrado para o peixe. Porém, foram feitas modificações na sua abordagem de estimativa, de forma que se passou utilizar a relação entre peso seco de tecido branquial na amostra (PSTA) e o peso seco de tecido branquial total (PST) como padrão de extrapolação, ao invés da relação entre o volume da amostra (VA) e o volume total da solução (VT).

Essa modificação levou em consideração que o volume de solução de formalina não possui relação direta com o tamanho da brânquia, e, conseqüentemente, com a área disponível para parasitismo. Ao passo que o peso seco, além de possuir relação direta com quantidade de tecido disponível para parasitismo, não sofre influência de variações na hidratação, sendo uma medida muito mais confiável.

Para assegurar a praticidade do método, uma vez que a aferição dos valores de peso seco de tecido branquial é trabalhosa e requer equipamentos de precisão, foi desenvolvido, através de MARSplines (Friedman, 1991), um modelo preditivo capaz de estimar com bastante confiabilidade e precisão os valores de PSTA e PST, utilizando como preditores apenas o peso e o comprimento do peixe, ou seja, duas variáveis de fácil mensuração.

Portanto, com o auxílio do modelo preditivo, os únicos equipamentos necessários para aplicar o método utilizado no software proposto seriam: microscópio, régua e balança comum. Além disso, o único requisito para sua utilização é o cumprimento das seguintes condições:

- a) O peso e comprimento do peixe devem ser aferidos;
- b) As brânquias do peixe devem ser raspadas e fixadas em 200 ml de solução de formalina;
- c) Deve-se contar, com o auxílio de uma câmara de McMaster, o número de protozoários observados em amostras de 1,5 ml, retiradas aleatoriamente da solução contendo raspado de brânquias após homogeneização.

A espécie escolhida para versão inicial do software foi a tilápia do Nilo, *Oreochromis niloticus* (Linnaeus, 1758), devido à sua grande importância na piscicultura de água doce tanto mundial (Cavichiolo *et al.*, 2002; FAO, 2009), quanto nacional (Boscardin, 2008; EPAGRI, 2008).

MATERIAIS E MÉTODOS

Partindo dos resultados obtidos no desenvolvimento do modelo preditivo para estimar as variáveis “Peso da brânquia hidratada”, “Peso seco de tecido em amostra de 1,5 ml” e “Peso seco total do tecido branquial” em tilápia do Nilo, foi desenvolvido um software baseado na linguagem de programação Java (Sun Microsystems).

Foi utilizada a versão 6 do *Java Development Kit* (JDK), com auxílio do ambiente integrado de desenvolvimento (IDE) Netbeans versão 6.9.1., para desenvolver uma aplicação do tipo desktop. Na renderização dos componentes foi utilizada a interface de programação de aplicações (API) *Swing*, também foi incluída a API Apache POI, responsável pelo processamento das planilhas de Excel (.xls).

No desenvolvimento do software foram inseridas as seguintes variáveis: Comprimento do peixe (C); Peso do peixe (P); Quantidade de protozoários contados na câmara de McMaster em amostra de 1,5 ml (QPA); Peso da Brânquia Hidratada (PBH); Peso seco de tecido branquial em amostra de 1,5 ml (PSTA); Peso seco total do tecido branquial (PST) e Quantidade total de protozoários nas brânquias do peixe (QTP). Onde as três primeiras correspondem às variáveis preditoras utilizadas na entrada de dados e as quatro últimas correspondem às variáveis estimadas resultantes.

Para realização dos cálculos, dentro do software foram embutidas as seguintes equações:

$$\begin{aligned} \text{Peso da Brânquia Hidratada} &= 1,34590046943541e+001 + \\ &1,52091129068729e+000*\max(0; C-2,85750000000000e+001) - \\ &3,35832682589586e-001*\max(0; 2,85750000000000e+001-C) + \\ &8,32181912220444e-003*\max(0; P-3,30000000000000e+002) - \\ &1,98665391104146e-002*\max(0; 3,30000000000000e+002-P) \end{aligned}$$

$$\begin{aligned} \text{Peso seco de tecido branquial em amostra de 1,5 ml} &= \\ &4,42639564167902e-003 + 3,45853322900816e-004*\max(0; C- \\ &2,85750000000000e+001) - 1,57333427311875e-005*\max(0; \\ &2,85750000000000e+001-C) + 2,23249818114776e-006*\max(0; P- \end{aligned}$$

$$3,3000000000000000e+002) - 1,18525810550014e-005*\max(0; 3,3000000000000000e+002-P)$$

$$\begin{aligned} \text{Peso seco total do tecido branquial} = & 5,22067839879816e-001 + \\ & 4,30982060784010e-002*\max(0; C-2,8575000000000000e+001) - \\ & 1,97271452915104e-003*\max(0; 2,8575000000000000e+001-C) + \\ & 3,67934219484082e-004*\max(0; P-3,3000000000000000e+002) - \\ & 1,42338860538150e-003*\max(0; 3,3000000000000000e+002-P) \end{aligned}$$

Quantidade total de protozoários nas brânquias do peixe (QTP) =
PST * QPA / PSTA

Após o desenvolvimento do software, os valores de QTP estimados pelo programa e pelo método original proposto por Tavares-Dias et al. (2001) foram simulados e comparados por meio do teste T de Student para dados pareados.

O cálculo utilizado para estimar os valores de QTP pelo método original foi:

$$QTP = VT * QPA / VA$$

Onde,

$$VT = 200$$

$$VA = 1,5$$

RESULTADOS

O software, batizado de *ProtoCount*, vem dentro de um arquivo de compactação (.rar) composto por três partes principais, um arquivo executável tipo java (.jar), uma pasta “lib”, contendo as bibliotecas externas utilizadas no programa, e um arquivo de parâmetros de configuração (.ini) com os valores das constantes utilizadas nas equações. Além disso, acompanham o software um arquivo do Excel (.xls) contendo um exemplo de entrada de dados e um arquivo *readme* (.txt) contendo informações sobre o programa.

3.1. Interface *ProtoCount*

A interface do programa (Figura i) é composta por pelos seguintes itens:

- Área de “Entrada de dados”: com três caixas de texto onde o usuário pode inserir os valores das variáveis predictoras “Peso

do Peixe” e “Comprimento total do peixe”, além da “Quantidade de protozoários na amostra de 1,5 ml”.

- Botão “calcular”: dá início ao processamento dos dados, onde a partir das equações embutidas no software e dos valores na entrada de dados são estimadas as variáveis “Peso da brânquia hidratada”, “Peso seco de tecido em amostra de 1,5 ml”, “Peso seco total do tecido branquial” e a “Quantidade total de protozoários nas brânquias do peixe”.
- Área de “Dados processados”: onde são exibidos os resultados obtidos do cálculo das equações.
- Menu “Arquivo”: onde se encontra o comando de “Importar Planilha” (Figura ii)
- Menu “Ajuda”: onde se encontra o atalho para a janela “Sobre” do software (Figura iii), contendo informações de contato dos autores.

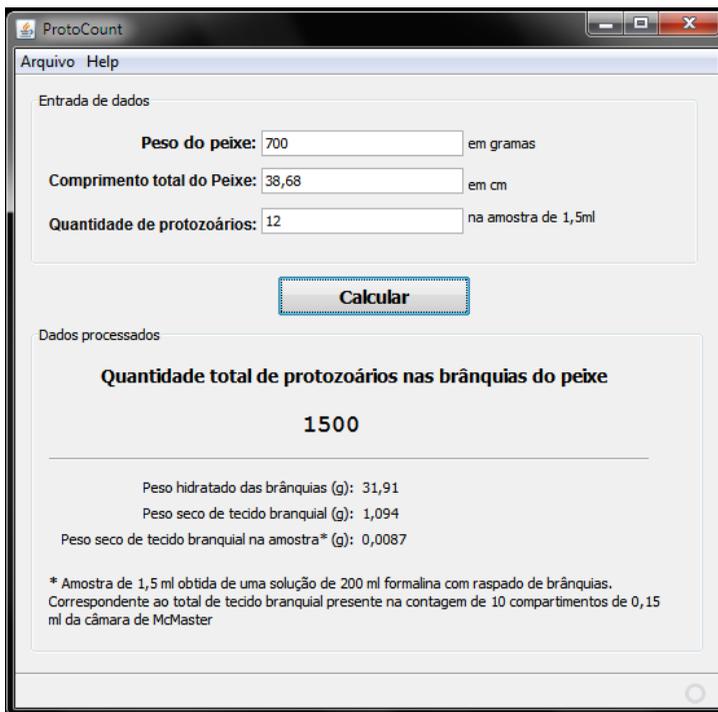


Figura xxxv: Captura de tela com a interface do software *ProtoCount*.

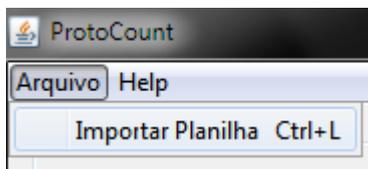


Figura xxxvi: Captura de tela com detalhe do menu “Arquivo” e do comando de “Importar Planilha”.

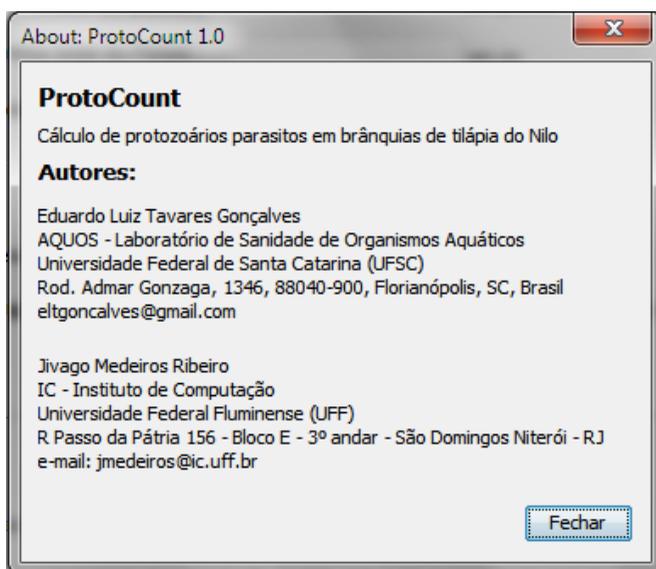


Figura xxxvii: Captura de tela com as informações contidas na janela “Sobre” do software.

3.2. Utilização do *ProtoCount*

O programa é dividido, basicamente, em três etapas de funcionamento:

1) Entrada de dados

É a primeira etapa do processo, onde o usuário digita na interface do programa os valores para as variáveis “Peso do peixe”, “Comprimento do peixe” e “Quantidade de protozoários na amostra de

1,5 ml” (QPA). O programa aceita a entrada de dados digitados utilizando tanto vírgula quanto ponto para separar as casas decimais.

Como alternativa de entrada de dados, o usuário pode preparar, conforme o exemplo que acompanha o software, uma planilha do tipo Excel (.xls) contendo os valores das três variáveis preditoras dispostos nas três primeiras colunas, pois o software *ProtoCount* é capaz de importar e processar o arquivo .xls, facilitando bastante a entrada de dados oriundos de várias amostras.

2) Processamento dos dados e cálculo das equações

Nesta etapa os dados de entrada são temporariamente armazenados para, após o comando do usuário, serem utilizados para calcular as equações embutidas no software.

Caso a entrada dos dados tenha sido feita através da importação de uma planilha .xls, o processamento dos dados ocorre automaticamente, sendo exibida uma mensagem de confirmação (Figura iv)

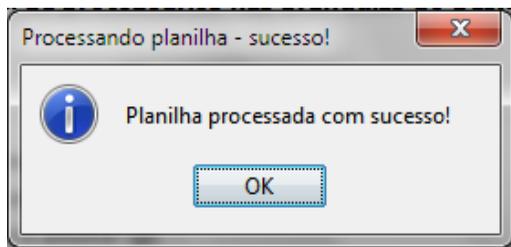


Figura xxxviii: Captura de tela com detalhe da mensagem confirmando o processamento dos dados importados com a planilha .xls.

3) Saída dos resultados

É a etapa final do processo, onde são exibidos os valores resultantes do cálculo das equações para as variáveis “Peso da Brânquia Hidratada” (PBH), “Peso seco de tecido branquial em amostra de 1,5 ml” (PSTA), “Peso seco total do tecido branquial” (PSTA) e “Quantidade total de protozoários nas brânquias do peixe” (QTP).

Caso a entrada de dados tenha sido realizada por meio da interface, os valores resultantes para as variáveis estimadas são exibidos na própria interface do software. Caso tenha sido importada uma planilha com os dados de entrada, o software gera uma nova planilha

contendo os valores das variáveis estimadas, que é salva automaticamente na mesma pasta de origem dos dados de entrada com o nome de “protocount_output”. Para evitar sobreposição de arquivos, cada planilha de resultados gerada leva junto de seu nome o *timestamp*, no formato “data-hora-minuto-segundo”, do momento de processamento.

Na Figura v está representando todo o processo envolvido na utilização do software *ProtoCount*.

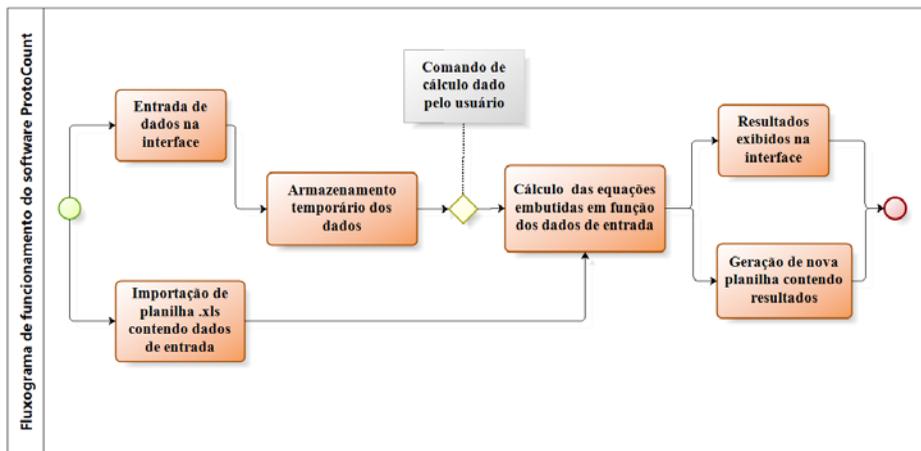


Figura xxxix: Fluxograma de funcionamento do software *ProtoCount*.

3.3. Simulação e comparação das estimativas

Foram feitas duas simulações da utilização do software *ProtoCount* para que os valores de QTP pudessem ser comparados com aqueles que seriam obtidos pelo método originalmente proposto por Tavares-Dias et al. (2001). As duas simulações utilizaram 118 pares iguais de valores de peso e comprimento do peixe, variando apenas os valores da quantidade de protozoários contados na amostra de 1,5 ml (QPA).

A primeira simulação utilizou valores de QPA entre 1 e 70 gerados aleatoriamente para verificar a dispersão das estimativas de QTP (Figura vi).

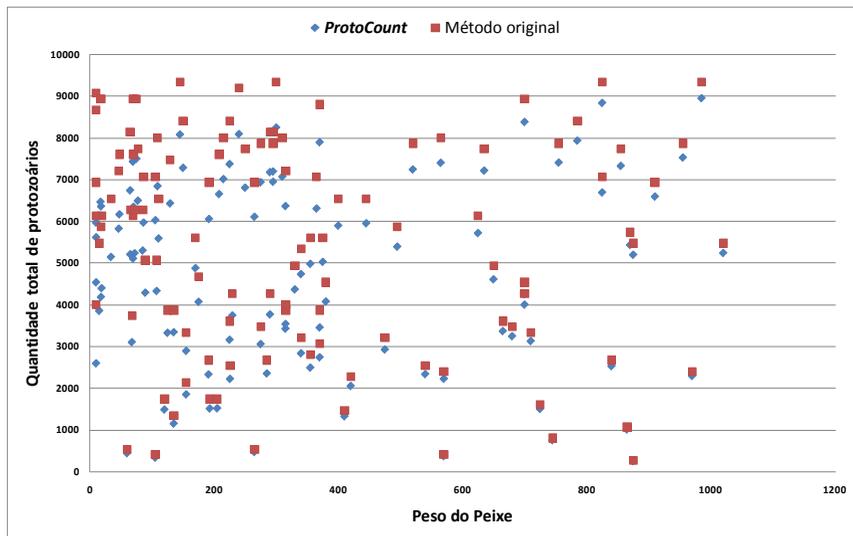


Figura xl: Gráfico de dispersão com valores de QTP estimados pelo software *ProtoCount* e pelo método original.

O resultado do teste T de Student para dados pareados (Tabela i) mostrou que as estimativas foram significativamente diferentes com $p = 9,9E^{-25}$, sendo que os valores estimados pelo *ProtoCount* apresentaram uma média mais baixa e tiveram menor variância.

Tabela xix: Resultados do Teste T pareado comparando os valores de quantidade total de protozoários (QTP) estimados pelo software *ProtoCount* e pelo método original. Florianópolis, junho de 2011.

	<i>ProtoCount</i>	<i>Método original</i>
Média	4748,016949	5489,265537
Variância	5073494,068	6755952,163
Observações	118	118
Correlação de Pearson	0,978098194	
Hipótese da diferença de média		0
gl		117
Stat t	-13,11904279	
P(T<=t) uni-caudal	4,93E-25	
t crítico uni-caudal	1,657981659	
P(T<=t) bi-caudal	9,85927E-25	
t crítico bi-caudal	1,980447532	

A segunda simulação utilizou valores fixos de QPA igual 10, com o intuito de verificar a dispersão das estimativas de QTP em função do peso dos peixes (Figura vii).

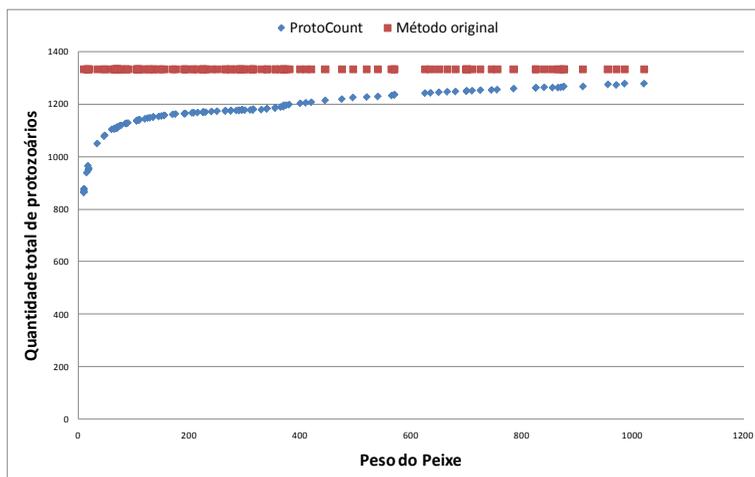


Figura xli: Gráfico de dispersão com valores de QTP estimados pelo software *ProtoCount* e pelo método original.

O resultado do teste T de Student para dados pareados (Tabela ii) também mostrou que as estimativas foram significativamente diferentes com $p = 1,9E^{-39}$, sendo que os valores estimados pelo método original praticamente não apresentaram variância, visto que os valores de QTP estimados foram os mesmos, independente do peso do peixe.

Tabela xx: Resultados do Teste T pareado comparando os valores de quantidade total de protozoários (QTP) estimados pelo software *ProtoCount* e pelo método original. Florianópolis, junho de 2011.

	<i>ProtoCount</i>	<i>Método original</i>
Média	1162,355932	1333,333333
Variância	8665,889323	5,21407E-26
Observações	118	118
Correlação de Pearson	-4,98897E-16	
Hipótese da diferença de média	0	
gl	117	
Stat t	-19,95138945	
P(T<=t) uni-caudal	9,24846E-40	
t crítico uni-caudal	1,657981659	
P(T<=t) bi-caudal	1,84969E-39	
t crítico bi-caudal	1,980447532	

DISCUSSÃO

4.1. *ProtoCount*

A escolha da plataforma Java se baseou na facilidade para o usuário, uma vez que o *Java Runtime Environment* (JRE), necessário para execução do software, é disponibilizado gratuitamente e é comumente utilizado em diversas aplicações rotineiras. Além disso, pode ser executado em qualquer sistema operacional.

Na construção do software *ProtoCount* almejou-se, também, que sua utilização fosse simples e intuitiva, por isso a interface dispõe de forma bastante direta os passos a serem tomados para obtenção dos resultados. A capacidade de importar e gerar planilhas do Excel adiciona uma praticidade extra em sua utilização, uma vez que possibilita a entrada de grandes quantidades de valores ao mesmo tempo e a disponibilização dos seus resultados de forma organizada.

A decisão de incluir a variável “Peso da brânquia hidratada” nos resultados calculados pelo software foi tomada para disponibilizar ao usuário um parâmetro intuitivo que represente a variação entre o tamanho das brânquias em função do crescimento do peixe.

4.2. Simulação e comparação das estimativas

A segunda simulação (Figura vii e Tabela ii) evidenciou com clareza a fragilidade das estimativas do método original, uma vez que é ilógico assumir que brânquias de diferentes tamanhos possuam a mesma quantidade de protozoários apenas porque o valor contado em uma amostra de mesmo volume foi igual.

Em compensação, as estimativas do software *ProtoCount* levaram em consideração as diferenças entre os tamanhos de brânquias, estimando valores menores de QTP para peixes com menores valores de peso seco de brânquia.

ProtoCount, portanto, realiza estimativas muito mais lógicas, pois tem como base uma variável (peso seco de tecido branquial) que está diretamente relacionada com o local parasitado pelos protozoários quantificados. Além disso, a relação entre o peso seco de tecido branquial na amostra (PSTA) e o peso seco total de tecido branquial (PST) representa bem a distribuição provável do parasitismo, desde que não exista uma preferência clara por *microhabitat* nas brânquias, tendência da qual não existem relatos para protozoários.

CONCLUSÃO

ProtoCount foi desenvolvido como uma ferramenta de quantificação de protozoários para os laboratórios brasileiros de parasitologia de peixes, no sentido de auxiliar e otimizar o trabalho realizado, além de torná-lo mais preciso e confiável.

Seu método prático e sistematizado dispensa o uso de equipamentos caros, podendo ser utilizado mesmo em condições de precariedade de instrumentos de medição, uma vez que o modelo preditivo embutido no software é capaz de estimar os parâmetros necessários a partir de variáveis de fácil mensuração.

Além disso, a modificação proposta no *ProtoCount*, de se utilizar o peso seco de tecido branquial como parâmetro de extrapolação ao invés do volume de solução, provou ser a escolha mais lógica e confiável, conforme mostram os resultados das simulações.

ProtoCount continuará sendo aprimorado, buscando estimativas cada vez mais precisas e abrangentes. Para o futuro é planejada a inserção de uma base de dados para o modelo preditivo, onde novos dados possam ser inseridos deixando o modelo cada vez mais robusto.

Os autores deste estudo almejam, também, desenvolver novos modelos preditivos, que futuramente possam levar em consideração diferentes espécies de peixes, diferentes volumes amostrais e diferentes locais amostrados para quantificação de protozoários parasitos.

Por fim, espera-se que a utilização do software *ProtoCount* possa contribuir para um aumento no número de trabalhos com análises quantitativas no Brasil, visando o desenvolvimento de parâmetros associados à doenças parasitárias em peixes e a diminuição da morbimortalidade em viveiros de criação.

AGRADECIMENTOS

Os autores agradecem a CAPES e CNPq pelo auxílio e pela bolsa de Mestrado e Produtividade em Pesquisa (CNPq 301072/2007-8, 302493/2010-7), e a Jivago Medeiros Ribeiro pela programação do software.

REFERÊNCIAS

- Boscardin, N. R., 2008. Produção Aqüícola, in: Ostrensky, A.; Borghetti, J.R.; Soto, D. (Eds.) Aqüicultura no Brasil: o desafio é crescer. Brasília, DF, pp. 27-72.
- Buchmann, K., Bresciani, J., 1997. Parasitic infections in pond-reared rainbow trout *Oncorhynchus mykiss* in Denmark. *Disease of Aquatic Organisms*, 28 (2), 125-138.
- Burreson, E. M.; Ragone-Calvo, L.M., 1996. Epizootiology of *Perkinsus marinus* disease of oysters in Chesapeake Bay, with emphasis on data since 1985. *Journal of Shellfish Research*, 15, 17-34.
- Cavichiolo, F.; Vargas, L.; Ribeiro, R. P.; Moreira, H. L. M.; Leonardo, J. M., 2002. Níveis de suplementação de vitamina C na ração sobre a ocorrência de ectoparasitas, sobrevivência e biomassa em alevinos de tilápia do Nilo (*Oreochromis niloticus* L.). *Acta Scientiarum*, 24 (4), 957-964.
- COSTA, N. C. A. da., 2008. Lógica Indutiva e Probabilidade , third ed. Hucitec, São Paulo.
- Crosby, M.P.; Roberts C.F., 1990. Seasonal infection intensity cycle of the parasite *Perkinsus marinus* (and the absence of *Haplosporidium* spp.) in oysters from a South Carolina salt marsh. *Diseases of Aquatic Organisms*, 9, 149-155.
- Daims, H.; Wagner, M., 2007. Quantification of uncultured microorganisms by fluorescence microscopy and digital image analysis. *Applied microbiology and biotechnology*, 75 (2), 237-48
- Eiras, J C; Chemes, S. B.; Casabianca, O. A.; Takemoto, R. M. ; Rossi, L. M. ; Pavanelli, G. C., 2008. An unusual infection of *Henneguya* sp. (Myxozoa , Myxosporea , Myxobolidae) in the gills of *Hoplosternum littorale* Hancock , 1928 (Osteichthyes , Callichthyidae) from Middle Paraná River , Santa Fe , Argentina. *Canadian Journal of Zoology*, 28, 22-26.

- Eiras, J.C.; Abreu, P. C.; Robaldo, R.; Pereira Júnior, J., 2007. *Myxobolus platanus* n. sp. (Myxosporea, Myxobolidae), a parasite of *Mugil platanus* Günther, 1880 (Osteichthyes, Mugilidae) from Lagoa dos Patos, RS, Brazil. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 59 (4), 895-898
- EPAGRI, Empresa de pesquisa Agropecuária e Extensão Rural de Santa Catarina, 2008. Desempenho da pesca e da Aqüicultura – Ano 2005-2006. Instituto Cepa - Centro de Socioeconômica e Planejamento Agrícola, Florianópolis.
- FAO, Food and Agriculture Organization Fisheries Department, 2010. The State of World Fisheries and Aquaculture. FAO, Rome.
- FRIEDMAN, J., 1991. Multivariate adaptive regression splines. Annals of Statistics, 19, 1-141.
- Higuchi, R.; Dollinger, G.; Walsh, P.S.; Griffith R., 1992. Simultaneous amplification and detection of specific DNA sequences. Biotechnology, 10, 413-417.
- Leonardo, J. M. L. O.; Pereira, J. V.; Krajevieski, M. E., 2006. Ocorrência de ectoparasitas e estacionalidade em alevinos de tilápia-donilo (*Oreochromis niloticus*) após a reversão sexual, na região noroeste do Paraná. Iniciação Científica CESUMAR, 8, 185-191.
- Martins, M. L.; Moraes, J. R. E.; Andrade, P. M.; Schalch, S. H. C.; Moraes, F. R., 2001. *Piscinoodinium pillulare* (Schäperclaus 1954) Lom, 1981 (Dinoflagellida) infection in cultivated freshwater fish from Northeast region of São Paulo State, Brazil. Parasitological and pathological aspects. Revista Brasileira de Biologia, 61 (4), 639-644.
- Martins, M. L.; Onaka E. M.; Moraes, F. R. De; Bozzo, F. R.; Paiva, A. De M. E F. C.; Gonçalves, A., 2002. Recent studies on parasitic infections of freshwater cultivated fish in the state of São Paulo, Brazil. Acta Scientiarum, 24, 981-985.
- Mayhew, T. M.; Gundersen, H. J., 1996. If you assume, you can make an ass out of u and me': a decade of the disector for stereological counting of particles in 3D space. Journal of anatomy, 188, 1-15.

- Monis, P. T.; Giglio, S.; Keegan, A. R.; Andrew Thompson, R. C., 2005. Emerging technologies for the detection and genetic characterization of protozoan parasites. *Trends in parasitology*, 21 (7), 340-6.
- Ranzani-Paiva, M. J. T.; Felizardo, N. N.; Luque, J. L., 2005. Parasitological and hematological analysis of Nile tilapia *Oreochromis niloticus* Linnaeus, 1757 from Guarapiranga reservoir, São Paulo State, Brazil. *Acta Scientiarum. Biological Sciences*, 29, 231-237
- Remacha-Triviño, A.; Borsay-Horowitz, D.; Dungan, C.; Gual-Arnau, X.; Gómez-Leon, J.; Villamil, L.; Gómez-Chiarri, M., 2008. Numerical Quantification of *Perkinsus marinus* in the American Oyster *Crassostrea virginica* (Gmelin, 1791) (Mollusca: Bivalvia) by Modern Stereology. *Journal of Parasitology*, 94 (1), 125-136.
- Tavares-Dias, M.; Martins, M. L.; Moraes, F. R., 2001. Fauna parasitária de peixes oriundos de “pesque-pague” do município de Franca, São Paulo, Brasil. I. Protozoários. *Revista Brasileira de Zoologia*, 18, 67 – 79.
- Tavares-Dias, M.; Brito, M. L. S.; Lemos, J. R. G., 2009. Protozoários e metazoários parasitos do cardinal *Paracheirodon axelrodi* Schultz, 1956 (Characidae), peixe ornamental proveniente de exportador de Manaus, Estado do Amazonas, Brasil. *Acta Scientiarum. Biological Sciences*, 31 (1), 23-28.

CONCLUSÕES GERAIS

Apliquei um grande esforço na tentativa de desenvolver um trabalho metodologicamente criterioso, que analisasse a questão sob diferentes aspectos, e que fosse capaz de oferecer uma alternativa confiável, precisa e de fácil aplicação na quantificação de protozoários parasitos no Brasil.

Inicialmente, almejava-se que o método desenvolvido pudesse auxiliar na quantificação de protozoários parasitos presentes nas brânquias e no muco de tilápias do Nilo. Entretanto, devido ao curto tempo disponível para o volume de análises e ao fato da maioria das análises e aferições terem sido feitas pessoalmente, o modelo preditivo e o software que o utiliza ficaram restritos à análise das brânquias de tilápias do Nilo. Já existe, porém, o planejamento de futuros trabalhos tratando sobre a quantificação de protozoários presentes no muco de peixes.

Além disso, existe o interesse em aprimorar o modelo preditivo desenvolvido, de modo que novas variáveis, como diferentes volumes amostrais e diferentes volumes de solução, possam ser modeladas, garantindo estimativas padronizadas e confiáveis, a despeito da técnica adotada na contagem dos protozoários.

A elaboração desta Dissertação de Mestrado foi o começo de um longo caminho a ser percorrido. No entanto, um importante passo foi dado, de forma que futuros trabalhos possam ampliar os estudos já realizados ao desenvolver modelos preditivos para outras espécies de peixes. Acredito que estes novos modelos possam ser acrescidos futuramente ao software *ProtoCount*, contribuindo para o desenvolvimento e aprimoramento desta ferramenta de quantificação que, além de ser disponibilizada gratuitamente, não exige a utilização de nenhum equipamento de alto custo.

Espero que os frutos deste estudo, apresentados sob a forma de um software e de um modelo preditivo, possam contribuir de forma relevante com a parasitologia de peixes nacional, possibilitando um aumento não só no número de trabalhos com análises quantitativas, mas, também, na sistematização, confiabilidade e precisão dessas análises.

REFERÊNCIAS DA INTRODUÇÃO

- AL-RASHEID, K.A.S. Trichodinid ectoparasites (Ciliophora: Peritrichida) of some River Nile fish, Egypt. **Parasitology International**, Tokyo, v. 49, p. 131-137, 2000.
- AMANN, R. I.; LUDWIG, W.; SCHLEIFER, K. H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. **Microbiological Reviews**, v. 59, n. 1, p. 143-69, 1995.
- ANDRADE, S.M.S.; MALTA, J.C.O.; FERRAZ, E. Fauna Parasitológica de Alevinos de Matrinhã *Brycon Cephalus* (Gunther, 1869) coletados nos rios Negro e Solimões, na Amazônia Central. **Acta Amazonica**, Manaus, v. 31, n. 2, p. 263-273, 2001.
- ARTHUR, J. R.; LOM, J. Trichodinid protozoa (Ciliophora: Peritrichida) from freshwater fishes of Rybinsk Reservoir, USSR. **Journal of Protozoology**, Lawrence, v. 31, p. 82-91, 1984.
- AZEVEDO, T. M. P. D.; MARTINS, MAURÍCIO LATERÇA; BOZZO, F. R.; MORAES, F. R. D. Haematological and gill responses in parasitized tilapia from valley of Tijucas River, SC, Brazil. **Scientia Agricola**, v. 63, n. 2, p. 115-120, 2006.
- BARKER, D. E.; CONE, D. K.; BURT, M. D. B. Trichodina murmanica (Ciliophora) and Gyrodactylus pleuronecti (Monogenea) parasitizing hatchery-reared winter flounder, *Pseudopleuronectes americanus* (Walbaum): effects on host growth and assessment of parasite interaction. **Journal of Fish Diseases**, v. 25, n. 2, p. 81-89, 2002.
- BOSCARDIN, N.R. Produção Aqüícola. *In*: OSTRENSKY, A.; BORGHETTI, J.R.; SOTO, D. (organizadores). **Aqüicultura no Brasil: o desafio é crescer**. Brasília, DF: 2008, p.27-72.
- BUCHMANN K., BRESCIANI J. Parasitic infections in pond-reared rainbow trout *Oncorhynchus mykiss* in Denmark. **Disease of Aquatic Organisms**, v. 28, n.2, p. 125-138, 1997.

BUCHMANN, K.; SIGH, J.; NIELSEN, C. V.; DALGAARD, M. Host responses against the fish parasitizing ciliate *Ichthyophthirius multifiliis*. **Veterinary parasitology**, v. 100, n. 1-2, p. 105-16, 2001.

BURRESON, E.M.; RAGONE-CALVO, L.M., Epizootiology of *Perkinsus marinus* disease of oysters in Chesapeake Bay, with emphasis on data since 1985. **Journal of Shellfish Research**, v.15, p. 17-34, 1996.

CAVICHIOLO, F.; VARGAS, L.; RIBEIRO, R. P.; MOREIRA, H. L. M.; LEONARDO, J. M. Níveis de suplementação de vitamina C na ração sobre a ocorrência de ectoparasitas, sobrevivência e biomassa em alevinos de tilápia do Nilo (*Oreochromis niloticus* L.). **Acta Scientiarum**, v. 24, n. 4, p. 957-964, 2002.

CECCARELLI, P. S.; FIGUEIRA, L. B.; FERRAZ DE LIMA, C. L. B.; OLIVEIRA, C. A.. Observações sobre a ocorrência de parasitos no CEPTA entre 1983 e 1990. **Boletim Técnico CEPTA**, Pirassununga, v. 3, p. 43-54, 1990.

CROSBY, M.P.; C.F. ROBERTS. Seasonal infection intensity cycle of the parasite *Perkinsus marinus* (and the absence of *Haplosporidium* spp.) in oysters from a South Carolina salt marsh. **Disease of Aquatic Organisms**, v. 9, p. 149-155, 1990.

DAIMS, H.; WAGNER, M. Quantification of uncultured microorganisms by fluorescence microscopy and digital image analysis. **Applied microbiology and biotechnology**, v. 75, n. 2, p. 237-48, 2007.

EDWARDS, K., LOGAN, J. AND SAUNDERS, N. **Real-Time PCR, An Essential Guide**. Wymondham: Horizon Bioscience, 2004.

EFRON, B; TIBSHIRANI, R. J. **An Introduction to the Bootstrap**. New York: Chapman and Hall, 1994.

EPAGRI, Empresa de pesquisa Agropecuária e Extensão Rural de Santa Catarina. **Desempenho da pesca e da Aqüicultura – Ano 2005-2006**. Instituto Cepa - Centro de Socioeconômica e Planejamento Agrícola, 2008. Disponível em:
<http://cepa.epagri.sc.gov.br/Publicacoes/sintese_2006/pesca_2006.pdf>
. Acesso em: 20 abr 2008.

FAO, Food and Agriculture Organization Fisheries Department. **The State of World Fisheries and Aquaculture**. Rome, FAO, 2010.

FERRAZ DE LIMA, C.L.B.; REIS, N.S.; CECCARELLI, P.S. Caracterização histológica da ictiofitiríase em pacu, *Piaractus mesopotamicus* Holmberg, 1887(Teleostei, Serrasalminae). **Boletim do CEPTA**, v.4, n.2, p. 39-46, 1991.

FRIEDMAN, J. Multivariate adaptive regression splines. **Annals of Statistics**, v. 19, p. 1-141, 1991.

GALLI, L. F.; TORLONI, C. E. Criação de Tilápias. In: **Criação de Peixes**. 3.ed., São Paulo: Nobel, 1986. p. 74-85

GHIRALDELLI, L.; MARTINS, M.L.; JERÔNIMO, G.T.; YAMASHITA, M.M.; ADAMANTE, W.B. Ectoparasites communities from *Oreochromis niloticus* cultivated in the State of Santa Catarina, Brazil. **Journal of Fisheries and Aquatic Science**, New York, v. 1, n. 2, p:181-190, 2006.

GUIDO, M. C. **Relação Parasita–Hospedeiro**. Aulas: Epidemiologia, 2009. Disponível em: <<http://www.mcguido.vet.br/parasita-hospedeiro.htm>>. Acesso em: 15 março 2011.

GUNDERSEN, H. J. G. Stereology of arbitrary particles. A review of unbiased number and size estimators and the presentation of some new ones, in memory of Willian R. Thompson. **Journal of Microscopy**, v. 143, p. 3–45, 1986.

HIGUCHI, R.; DOLLINGER, G.; WALSH, P.S.; GRIFFITH R. Simultaneous amplification and detection of specific DNA sequences. **Biotechnology**, New York, v. 10, p. 413–417, 1992.

KLESZIUS, P.; ROGERS, W. Parasitisms of catfish and other farm-raised food fish. **Journal of the American Veterinary Medical Association**, v.207, n.11, p.1473–1478, Dec, 1995.

LITTLE, D. C., D. SIKAWA AND J. JUNTANA. Commercial production and marketing of Nile tilapia (*Oreochromis niloticus*) fry in Chonburi and Chachoengsao Provinces, Thailand. **The NAGA ICLARM Quarterly**, Abril, p. 14-17, 1994.

LUDWIG W., SCHLEIFER K.H. How quantitative is quantitative PCR with respect to cell counts. **Systematic and Applied Microbiology**, v. 23, p. 556-562, 2000.

MANCINI, M.; LARRIESTRA, A.; SANCHEZ, J. Estudio ictiopatólogico en poblaciones silvestres de la región centro-sur de la provincia de Córdoba - Argentina. **Revista de medicina veterinaria**, Buenos Aires, v. 81, n. 2, p. 104-108, 2000.

MARTINS, M. L. Doenças Infecciosas e Parasitárias de Peixes. **Boletim Técnico CAUNESP**, São Paulo, Brasil, 66 p., 1998.

MARTINS, M. L.; MORAES, J. R. E.; ANDRADE, P. M.; SCHALCH, S. H. C.; MORAES, F. R.. *Piscinoodinium pillulare* (Schäperclaus 1954) Lom, 1981 (Dinoflagellida) infection in cultivated freshwater fish from Northeast region of São Paulo State, Brazil. Parasitological and pathological aspects. **Revista Brasileira de Biologia**, Rio de Janeiro, v. 61, n. 4, p. 639-644, 2001.

MARTINS, M. L.; ROMERO, N.G. Efectos del parasitismo sobre el tejido branquial en peces cultivados: estudio parasitologico e histopatologico. **Revista Brasileira de Zoologia**, Curitiba, v. 13, n. 2, p. 489-500, 1996.

MARTINS, M. L.; ONAKA E. M.; MORAES, F. R. de; BOZZO, F. R.; PAIVA, A. de M. e F. C.; GONÇALVES, A. Recent studies on parasitic infections of freshwater cultivated fish in the state of São Paulo , Brazil. **Acta Scientiarum**, p. 981-985, 2002.

MAYHEW, T. M.; GUNDERSEN, H. J. If you assume, you can make an ass out of u and me': a decade of the disector for stereological counting of particles in 3D space. **Journal of anatomy**, v. 188 (Pt 1, p. 1-15, 1996.

MOLNÁR, K. Effect of decreased water oxygen content on common carp fry with *Dactylogyrus vastator* (Monogenea) infection of varying severity. **Diseases of Aquatic Organisms (DAO)**, Oldendorf, v.20, p.153-157, 1994.

MONIS, P. T.; GIGLIO, S.; KEEGAN, A. R.; ANDREW THOMPSON, R. C. Emerging technologies for the detection and genetic characterization of protozoan parasites. **Trends in parasitology**, v. 21, n. 7, p. 340-6, 2005.

MORAES, F.R.; MARTINS, M.L. Condições predisponentes e principais enfermidades de teleósteos em piscicultura intensiva. In: Cyrino, J.E.P., Urbinati, E.C., Fracalossi, D.M., Castagnolli, N. **Tópicos especiais em piscicultura de água doce tropical intensiva**. São Paulo: TecArt, 2004. p. 343-383.

NOGA, EDWARD J. **Fish disease: diagnosis and treatment**. Wiley-Blackwell, 2000.

PAVANELLI, G.C.; J.C. EIRAS; R.M. TAKEMOTO. **Doenças de peixes: profilaxia, diagnóstico e tratamento**. Maringá: Universidade Estadual de Maringá, 2002.

POULIN, R. The Diversity of Parasites. **The Quarterly Review of Biology**, v. 75, p. 277–293, 2000.

RAMSING, N. B.; FOSSING, H.; FERDELMAN, T. G.; ANDERSEN, F.; THAMDRUP, B. Distribution of bacterial populations in a stratified fjord (mariager fjord, denmark) quantified by in situ hybridization and related to chemical gradients in the water column. **Applied and environmental microbiology**, v. 62, n. 10, p. 3915, 1996.

RANZANI-PAIVA, M. J. T.; SILVA-SOUZA, A T. Co-infestation of gills by different parasite groups in the mullet, *Mugil platanus* Günther, 1880 (Osteichthyes, Mugilidae): effects on relative condition factor. **Brazilian journal of biology**, v. 64, n. 3B, p. 677-82, 2004.

RAY, S. M. A culture technique for the diagnosis of infections with *Dermocystidium marinum*. Mackin, Owen, and Collier, in oysters. **Science**, v. 166, p. 360–361, 1952.

_____. A review of the culture method for detecting *Dermocystidium marinum*, with suggested modifications and precautions. **Proceedings of the National Shellfisheries Association**, v. 54, p. 55-69, 1966.

REMACHA-TRIVIÑO, A.; BORSAY-HOROWITZ, D.; DUNGAN, C.; GUAL-ARNAU, X.; GÓMEZ-LEON, J.; VILLAMIL, L.; GÓMEZ-CHIARRI, M. Numerical Quantification of *Perkinsus marinus* in the American Oyster *Crassostrea virginica* (Gmelin, 1791) (Mollusca: Bivalvia) by Modern Stereology. **Journal of Parasitology**, v. 94, n. 1, p. 125-136, 2008.

ROBERTS, R.J. **Patología de Los Peces**. Madrid: Mundi-Prensa, 1981.

SCHMID, M.; TWACHTMANN, U.; KLEIN, M.; STROUS, M.; JURETSCHKO, S.; JETTEN, M.; METZGER, J.W.; SCHLEIFER, K.H.; WAGNER, M. Molecular evidence for genus level diversity of bacteria capable of catalyzing anaerobic ammonium oxidation. **Systematic and applied microbiology**, v. 23, p. 93-106, 2000.

SIKKA, G., KAUR, A. & UDDIN, M. Estimating Function points: Using Machine Learning and Regression Models. In: INTERNATIONAL CONFERENCE ON EDUCATION TECHNOLOGY AND COMPUTER, 2, 2010, Shanghai, p. 52-56.

SNIESZKO, S.F. The effects of environmental stress on outbreaks of infectious diseases of fishes. **Journal of Fish Biology**, v. 6, p. 197-208, 1974.

STATSOFT, Inc. **Electronic Statistics Textbook**. Tulsa, OK: StatSoft, 2011. Disponível em: <<http://www.statsoft.com/textbook/>>

TAVARES-DIAS, M.; MARTINS, MAURÍCIO L.; MORAES, F. R. Fauna parasitária de peixes oriundos de “pesque-pague” do município de Franca, São Paulo, Brasil. I. Protozoários. **Revista Brasileira de Zoologia**, v. 18, p. 67 - 79, 2001.

THATCHER, V. E.; BRITES-NETO, J. B. Diagnóstico, Prevenção e Tratamento de Enfermidades de peixes neo tropicais de água doce. **Revista Brasileira de Medicina Veterinária**, Rio de Janeiro, v. 16(3), p. 111-128. 1994

THILAKARATNE, I. D. S. I. P.; RAJAPAKSHA, G.; HEWAKOPARA, A.; RAJAPAKSE, R. P. V. J.; FAIZAL, A. C. M. Parasitic infections in freshwater ornamental fish in Sri Lanka. **Diseases of Aquatic Organisms**, Oldendorf, v. 54, p. 157- 162, 2003.

VARGAS, L.; POVH, J.A.; RIBEIRO, R.P.; MOREIRA, H.L.M. Prevalência de ectoparasitos en tilapia del Nilo (*Oreochromis niloticus*) de origen tailandesa de Maringá, Paraná. **Arquivos de Ciências Veterinárias e Zoologia**, UNIPAR, v. 3, p. 32-37, 2000.

WEST, M. J.; SLOMIANKA, L.; GUNDERSEN, H. J. G. Unbiased stereological estimation of the total number of neurons in the subdivisions of the rat hippocampus using the optical fractionator. **The Anatomical Record**, v. 231, p. 428–497, 1991.

WEST, M. J. Stereological methods for estimating the total number of neurons and synapses: issues of precision and bias. **Trends in neurosciences**, v. 22, n. 2, p. 51-61, 1999.

XU, K. *et al.* Trichodinid ectoparasites (Ciliophora: Peritrichida) of some marine fishes from coastal regions of the Yellow Sea and Bohai Sea. **Systematic parasitology**, Dordrecht, v. 50, p. 69-79, 2001.

ZANIBONI FILHO, E. *Piscicultura das espécies exóticas de água doce*. In: Poli, C.R.; Poli, A.T.B.; ANDREATTA, E.; BELTRAME, E. **Aqüicultura: experiência brasileira**. Florianópolis: Multitarefa, 2004. p. 309-336.

GLOSSÁRIO

Application Programming Interface (API) = conjunto de bibliotecas que fornecem recursos e funcionalidades específicas para o desenvolvimento de software.

Árvores de decisão = é uma representação de uma tabela de decisão sob a forma de uma árvore. Tem a mesma utilidade da tabela de decisão. Trata-se de uma maneira alternativa de expressar as mesmas regras que são obtidas quando se constrói a tabela. (Wikipédia. Disponível em: <<http://pt.wikipedia.org>>)

Diagramas de Venn = Ilustrações de conjuntos, relações matemáticas ou relações lógicas. Os diagramas de Venn foram introduzidos em 1881 pelo filósofo e matemático britânico John Venn. (Wikipédia. Disponível em: <<http://pt.wikipedia.org>>)

Distância Euclidiana = Em matemática, distância euclidiana (ou distância métrica) é a distância entre dois pontos, que pode ser provada pela aplicação repetida do teorema de Pitágoras. (Wikipédia. Disponível em: <<http://pt.wikipedia.org>>)

Integrated Development Environment (IDE) = programa de computador que reúne características e ferramentas de apoio ao desenvolvimento de software.

Java Runtime Environment (JRE) = ambiente que é utilizado para executar as aplicações da plataforma Java. É composto por bibliotecas (APIs) e pela Máquina virtual Java (JVM).

Microhabitat = *Microhabitat* é a vizinhança imediata do local onde vive um determinado espécime animal ou vegetal. No caso de parasitos de brânquias, representa uma área específica da brânquia, como determinado arco ou posição.

Polimerase = enzima que catalisa a ligação de nucleotídeos e leva à formação de cadeia de ácido nucléico. (Dicionário Houaiss. Disponível em: <<http://search.4shared.com>>)

Princípio da parcimônia = Preferência pela explicação mais simples para uma observação, geralmente considerada a melhor maneira de julgar hipóteses.

APÊNDICE
Fotografias das etapas de desenvolvimento do modelo preditivo.
Florianópolis, Junho, 2011.



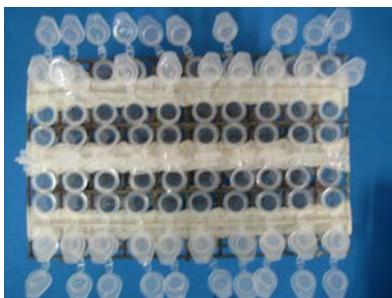
Balança de precisão modelo
Bel Mark série M



Pesagem na balança eletrônica
modelo Marte



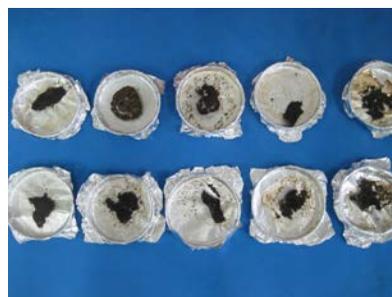
Fotografia do peixe contra fundo azul,
utilizada para análise de imagem



Estrutura utilizada para secagem dos tubos Eppendorf contendo tecido branquial em uma amostra de 1,5 ml



Detalhe do tecido branquial seco dentro do tubo Eppendorf



Placas de Petri de vidro forradas com papel alumínio, utilizadas para secagem do tecido branquial total



Detalhe do tecido branquial hidratado



Detalhe do tecido branquial após secagem em estufa