

**UNIVERSIDADE FEDERAL DO RIO GRANDE  
PÓS-GRADUAÇÃO EM OCEANOGRAFIA BIOLÓGICA**

**MODELAGEM ESPACIAL E TEMPORAL DA  
ABUNDÂNCIA E RIQUEZA DE ELASMOBRÂNQUIOS  
DO SUL DO BRASIL**

**RAYD IVANOFF NUNES**

Tese apresentada ao Programa de Pós-Graduação em Oceanografia Biológica da Universidade Federal do Rio Grande – FURG, como requisito parcial à obtenção do título de Doutor.

Orientador: Prof. Dr. Paul Gerhard Kinas  
Co-orientador: Prof. Dr. Gonzalo Velasco

**RIO GRANDE  
FEVEREIRO 2019**

À conservação das espécies de tubarões e raias  
para que um dia eu possa mostrá-las ao meu filho.

## AGRADECIMENTOS

À Universidade Federal do Rio Grande onde pude concluir minha graduação e tive a oportunidade de realizar meu doutorado;

Ao Programa de Pós-graduação em Oceanografia Biológica e a CAPES pela concessão da bolsa de estudos;

Agradeço imensamente meu orientador Kinas por sempre estar presente e disponível para me ajudar;

Ao meu querido Prof. Vooren pela confiança, incentivo e carinho durante todos estes anos de convívio e amizade;

Aos membros da banca de acompanhamento Prof. Dr. Lauro Madureira, Prof. Dr. Luciano Dalla Rosa e em especial o co-orientador Prof. Dr. Gonzalo Velasco pelas contribuições ao trabalho;

Meus sinceros agradecimentos a Dra. Maria Grazia Pennino e Msc. Marie Christine Rufener, vocês foram fundamentais para que eu conseguisse concluir meu doutorado;

Aos membros do *Statistical Modeling Ecology Group* (SMEG), Prof. Dr. David V. Conesa, Prof. Dr. Antonio Lopez Quilez e Prof. Dr. Xavier Barber pelas correções e sugestões aos trabalhos;

Aos amigos Marcus Guimaraes pelo trabalho do banco de dados e Marcelo Pinho pela ajuda no Surfer;

As minhas amigas Andréia Schwingel, Fernanda Lana, Laura Miranda, Maria Tereza Veloso, Daiane Kafer, Paula Pereira, Thaynã Corrêia, Michele Pieroni, Letícia Canton que sempre me incentivaram e me apoiaram;

Ao pessoal do Laboratório de Estatística Ambiental pela excelente convivência durante estes anos;

Aos meus pais por acreditarem em mim;

Aos meus amores Leonardo Mattos e João Paulo que chegaram para fazer minha vida muito mais feliz!

Às vezes eu tropeço, caio e me quebro em pedaços. Isso me fortalece.  
Oxe, se um de mim já é forte, imagine vários!

Bráulio Bessa

## ÍNDICE

	Página
RESUMO E PALAVRAS-CHAVE .....	6
ABSTRACT E KEY WORDS.....	7
1. INTRODUÇÃO.....	8
2. OBJETIVOS.....	14
3. MATERIAL E MÉTODOS.....	15
3.1 Área de estudo.....	15
3.2 Base de dados.....	16
3.3 Análise de dados .....	17
4. RESULTADOS GERAIS.....	19
4.1 Resultados Artigo I.....	19
4.2 Resultados Artigo II.....	20
5. CONCLUSÕES.....	22
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	24
APÊNDICE I - Um banco de dados para capturas de elasmobrânquios em cruzeiros científicos.....	30
APÊNDICE II - Análise estatísticas complementares .....	38
APÊNDICE III - Long-term Rajiformes (Chondrichthyes: Elasmobranchii) spatio-temporal dynamics in the southernmost state of Brazil. ....	46
APÊNDICE IV - Modelagem espacial bayesiana para riqueza de elasmobrânquios do extremo sul do Brasil.....	47
APÊNDICE V - Códigos no R .....	48

## RESUMO

A taxa de elasmobrânquios ameaçados de extinção no Brasil excede a taxa global identificada para o grupo. A pesca é apontada como o principal fator de ameaça, muito embora a degradação do habitat, poluição e alterações climáticas também devam ser consideradas. Modelar a influência de um conjunto de variáveis bióticas e abióticas na distribuição espacial da abundância e riqueza de elasmobrânquios no litoral do Rio Grande do Sul (RS), pode gerar subsídios para a gestão visando sua conservação nesta região. Com este propósito foram utilizados métodos estatísticos robustos e inovadores visando otimizar o uso de dados coletados por meio de cruzeiros científicos ao longo de várias décadas. Um banco de dados relacional, com 10.235 registros inseridos, foi construído para sistematizar as informações e otimizar as análises. Entre 1980 a 2002 foi analisada uma série de 690 arrastos de fundo, divididos entre décadas e estações do ano. Foi verificado, através de análise multivariada, que as espécies da ordem Rajiformes formam agrupamentos consistentes, o que permitiu modelar a abundância do grupo como um todo. As variáveis ambientais mais importantes para explicar o uso do habitat pelos Rajiformes foram a batimetria e a salinidade de superfície. Dados de CPUE (kg/área varrida) observada e padronizadas indicam reduções de abundância e apontam para a necessidade de realizarem-se ações de manejo emergenciais para este grupo. Análises feitas sobre a riqueza de elasmobrânquios obtidos por meio de cruzeiro de pesquisa realizado em 2005 no litoral do RS indicam que batimetria, temperatura e salinidade de superfície são as variáveis que determinam o ambiente adequado para as espécies durante a estação do verão. A riqueza de elasmobrânquios esteve representada por 19 espécies capturadas, variando entre 1 a 9 o número de espécies por lance de pesca. Os resultados apresentados nestes estudos são oriundos de modelos limitados no tempo e espaço e por isso configuram uma imagem instantânea do relacionamento dos elasmobrânquios e o ambiente do litoral do Rio Grande do Sul. Contudo, os mapas preditivos oriundo do ajuste de modelos hierárquicos Bayesianos, cujos parâmetros foram obtidos via INLA (*Integrated Nested Laplace Approximations*), podem ser diretamente utilizados por gestores na definição de estratégias de manejo que se utilizam de áreas de proteção ou exclusão da pesca, por exemplo.

**Palavras-chave:** Conservação, Habitat essenciais, Modelos Bayesianos.

## ABSTRACT

The rate of elasmobranchs threatened with extinction in Brazil exceeds the world rate identified for the group. Fishing is pointed out as the main threat factor, although habitat degradation, pollution and climate change must also be considered. Modeling the influence of a set of biotic and abiotic variables on the spatial distribution and species richness of elasmobranchs along the coast of Rio Grande do Sul (RS), can generate subsidies to manage for their conservation in this region. We do so, robust and innovative statistical methods were used to optimize the use of all information generated in scientific cruises over several decades. A relational database with 10.235 records entered was constructed with the purpose of systematizing the information and optimizing the analyzes. Between 1980 and 2002 a series of scientific cruises were analyzed that totaled 690 bottom trawls divided between the seasons of the year. It was verified through multivariate analysis that the Rajiformes formed groupings of species, which allowed to model the abundance of the group as a whole. The most important environmental variables to explain the use of the habitat were bathymetry and surface salinity. observed and standardized CPUE's Rajiformes (kg / area swept) indicate reductions in abundance and lead to emergency management actions for this group. Analyzes made on the elasmobranch richness obtained through a research cruise carried out in 2005 on the coast of RS indicate that bathymetry, temperature and surface salinity are the variables that provide the adequate environment for the species during the summer season. The elasmobranch richness was represented to 19 captured species, which ranged from 1 to 9 the number of species per fishing haul. The results presented in these studies are derived from models limited in time and space and therefore represent an instantaneous image of the relationship of elasmobranchs and the coastal environment of Rio Grande do Sul. However, the predictive maps derived from the adjustment of Bayesian hierarchical models, parameters were obtained via INLA (Integrated Nested Laplace Approximations), can be directly used by decision makers when defining management strategies, such as areas of protection or exclusion from fishing.

**Key-words:** Conservation, Essential habitats, Bayesian models.

## 1. INTRODUÇÃO

Em números atuais, no Brasil existem 149 elasmobrânquios dos quais 87 são espécies de tubarões e 62 são de raias (ICMBIO 2019). No Rio Grande do Sul, os elasmobrânquios constituem um importante componente da ictiofauna marinha, totalizando 83 espécies, sendo 48 tubarões e 35 raias (Vooren, 1998; Menezes et al. 2003; Vooren & Oddone, 2018).

Segundo Vooren (1998) as espécies de elasmobrânquios demersais no RS apresentam um cronograma anual da sua distribuição temporal podendo ser classificadas em espécies residentes, migrantes de inverno, migrantes de verão e de ocorrência esporádica. Entende-se por residente aquelas que completam seu ciclo de vida inteiramente nas águas sul-brasileiras e que ocorrem ao longo de todo ano sem variações sazonais de abundância.

As espécies migrantes de inverno migram a partir das águas Uruguaias e Argentinas para a plataforma sul brasileira seguindo a Corrente das Malvinas. Em contrapartida as migrantes de verão utilizam as águas quentes de origem tropical (Corrente do Brasil) para avançar desde a região sudeste do Brasil para a plataforma sul no litoral do RS (Vooren 1998)

Na década de 1970, quando a frota de arrasteiros ampliou sua atuação no Brasil (Haimovici et al. 1998), as capturas de elasmobrânquios no RS costumavam ser frequentes e abundantes. Em algumas ocasiões a magnitude de captura sobre estoques de tubarões e raias chegava a exceder a obtida com os teleósteos (SUDEPE 1972, SUDEPE 1974).

Contudo, no começo da década de 1980 surgiram os primeiros indícios de que os elasmobrânquios não sustentariam uma pesca intensiva. Comparando as densidades (kg/ha) estimadas em cruzeiros de pesca científica, no período entre 1972-1978 e 1981-1987, a redução de estoques de elasmobrânquios demersais no RS foi ordem de 66% (Haimovici et al. 2007). Entre 1980 e 2000 a participação dos elasmobrânquios nas capturas demersais da plataforma externa e talude decresceu de 45% para 3% (Fischer, 2008).

Não obstante, 75% das espécies de elasmobrânquios do Rio Grande do Sul



constam como Em Perigo (EN) e 60% como Criticamente em Perigo (CR) de acordo com a atual Avaliação do Risco de Extinção da Fauna Brasileira, utilizando a metodologia da IUCN (ICMBio, 2016).

Em geral, os elasmobrânquios apresentam uma alta vulnerabilidade à exploração pesqueira em razão de suas características biológicas intrínsecas, o que lhes confere um crescimento lento, maturação tardia e baixa mortalidade natural (Carrier et al. 2004). Em decorrência disso, as populações de tubarões e raias apresentam uma pequena taxa de crescimento populacional em qualquer nível de abundância da população (Vooren & Klippel 2005; Walker 2005).

Desta forma, em pescarias mistas, os estoques de elasmobrânquios declinam rapidamente enquanto que camarões e peixes tendem a sustentar a pescaria comercial (Vooren & Klippel 2005). Por este motivo Dulvy et al. (2003) coloca a pesca como principal responsável pelo risco de extinção local e global de inúmeras espécies marinhas. Bornatowski et al. (2017) acrescentam como agravantes os efeitos das mudanças climáticas, destruição do habitat, introdução de espécies não nativas e poluição como demais ameaças à biodiversidade global de elasmobrânquios.

Tendo em vista que elasmobrânquios são, em sua grande maioria, predadores de topo, estes desempenham um importante papel de equilíbrio trófico nos oceanos (Myers et al. 2007) seja por efeitos diretos (predação), indiretos (comportamento da presa) e reciclagem de nutrientes na coluna d'água (Heithaus et al. 2008, Vooren com. pess.<sup>1</sup>). Sendo assim, a retirada descontrolada de tubarões e raias dos oceanos pode resultar em efeitos de cascata trófica, com a substituição de espécies dominantes por espécies secundárias (Velasco et al. 2007, Baum & Worm 2009).

Para minimizar os efeitos da pesca de arrasto sobre os elasmobrânquios no litoral do RS algumas medidas restritivas tem sido aplicadas como a proibição de pesca com a utilização de redes de arrasto de qualquer tipo a menos de 12 (três) milhas da costa do estado do Rio Grande do Sul (RS, 2018). Será através do monitoramento contínuo de longo prazo, em conjunto com a fiscalização desta lei, que poderemos avaliar futuramente os efeitos da recuperação dos estoques oriundos dessa regulamentação.

A nível de Brasil, para amenizar os impactos antrópicos sobre as espécies de

---

1. Prof. Dr. Carolus Maria Vooren, até recentemente encarregado do Lab. de Elasmobrânquios do IO-FURG.

elasmobrânquios, foi lançado em 2014 o Plano de Ação Nacional para Conservação dos Tubarões e Raias Marinhos Ameaçados de Extinção (PAN TUBARÕES) estabelecido pela Portaria ICMBio nº 125 de 2014. Dentre os objetivos deste plano encontra-se a criação e ampliação de unidades de conservação em ambientes críticos para as espécies.

Segundo o PAN, 12 espécies são consideradas prioritárias e destas 6 ocorrem no RS, a saber: *Cetorhinus maximus*, *Mustelus schmitti*, *Galeorhinus galeus* e as endêmicas Atlântico Sudoeste *Pseudobatos horkelii*, *Squatina occulta* e *Squatina guggenheim*. Isto coloca o litoral deste Estado como uma área prioritária para conservação, sendo considerado, em termos mundiais, um *hotspot* de biodiversidade que merece ser preservado dado o alto esforço de pesca (Dulvy et al. 2014).

Uma das medidas utilizadas para manejar e conservar espécies ameaçadas de extinção consiste na criação de áreas marinhas protegidas (AMPs), nas quais as práticas pesqueiras podem ser excluídas de maneira integral ou sazonal. Esta proposta é bastante atrativa do ponto de vista de conservação e comumente tem como embasamento a análise descritiva de dados de pesca e dos próprios desembarques pesqueiros.

A delimitação de AMPs, contudo, parte da premissa que existam dados biológicos e ecológicos abundantes e adequados que permitam desenvolver uma boa fundamentação estatística. Neste sentido, a presente tese teve como escopo reunir evidências robustas a respeito da distribuição espacial da abundância e riqueza dos elasmobrânquios no litoral do Rio Grande do Sul.

Modelos estatísticos aliados a Sistemas de Informações Geográficas (SIG) tem se mostrado uma ferramenta muito eficiente para compreender como se estabelece o padrão de distribuição espacial e temporal das espécies e a influência das variáveis ambientais sobre este processo. Em especial, os modelos geoestatísticos hierárquicos Bayesianos, estruturados em forma de modelos Gaussianos latentes, têm se mostrado muito adequados pois representam explicitamente todos os processos estocásticos espaciais que podem exercer algum efeito sobre o fenômeno investigado (Rue et al. 2009).

Este conjunto de modelos tem como premissa a natureza espacialmente estruturada dos dados permitindo avaliar a correlação espacial entre as coordenadas (Krainski & Lindgren, 2015). Porém modelos de distribuição de espécies necessitam ultrapassar a barreira de um domínio espacial contínuo, conhecido como Campo

aleatório Gaussiano (GRF, do inglês *Gaussian Random Field*) para a estimação de seus parâmetros. A solução mais coerente para este problema está em discretizar o GRF de modo que este assuma a forma de um Campo Aleatório Markoviano Gaussiano (GMRF, do inglês *Gaussian Markov Random Field*) (Blangiardo & Cameletti, 2015).

A metodologia de Lindgren et al. (2011) tem sido a proposta mais utilizada e se apropria de uma Equação Diferencial Estocástica Parcial (SPDE, do inglês *Stochastic Partial Differential Equations*) para realizar a aproximação de um GRF por meio de um GMRF. Desta forma a área de estimação e predição do modelo é definida dentro de uma malha (*mesh*) onde as observações são tratadas como vértices iniciais para a triangulação e vértices extras são adicionados heurísticamente para minimizar o número de triângulos necessários para cobrir toda a área de estudo (Figura 1)

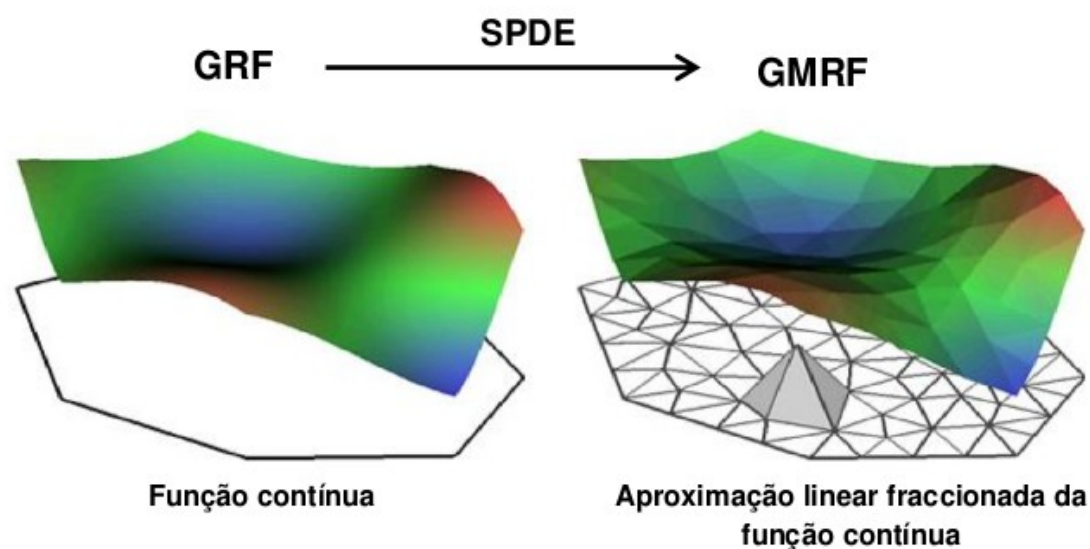


Figura 1: Ilustração da aproximação de uma função linear fracionada em uma malha triangular (Adaptado de SIMPSON et al. 2012). GRF (do inglês *Gaussian Random Field*) é o Campo aleatório Gaussiano; SPDE (do inglês *Stochastic Partial Differential Equations*) representa a Equação Diferencial Estocástica Parcial; GMRF (do inglês *Gaussian Markov Random Field*) é o Campo Aleatório Markoviano Gaussiano.

As vantagens sobre uma malha regular estão na triangulação da área que é mais densa (i.e., triângulos menores) em regiões onde há mais observações e economiza

tempo computacional pois o número de locais de previsão são menores em comparação a malha convencional (Franklin, 2010)

A inferência estatística dos parâmetros do modelo, quando obtidas num contexto Bayesiano, agregam toda a aleatoriedade das observações e demais incertezas associadas em estimativas na forma de distribuições de probabilidade (Kinas & Andrade, 2010). Neste estudo, uma solução via integração numérica denotada INLA (*Integrated Nested Laplace Approximation*) (Rue et al. 2016) consiste na adaptação da aproximação de Laplace para construir distribuições marginais posteriores de todos os parâmetros e das previsões.

Estudos com a aplicação desta metodologia para o grupo de elasmobrânquios são recentes, escassos e dirigidos a espécies do Mar Mediterrâneo (Pennino et al. 2013) e Báltico (Cosandey-Godin et al. 2015). No Brasil, trabalhos científicos aplicando análise Bayesiana através da metodologia INLA estão restritos à região Nordeste (Roos et al. 2015, Rufener et al. 2017).

Por fim, este trabalho apresentará um panorama das mudanças no tamanho das populações e da riqueza dos elasmobrânquios do sul do Brasil (Cabo de Santa Marta Grande ao Chuí) ao longo do tempo. As informações aqui apresentadas poderão subsidiar novas avaliações de risco de extinção das espécies de elasmobrânquios no Brasil e indicar pontos de referência para ações de manejo e conservação das espécies.

Para compreender melhor a sequência de desenvolvimento desta tese deve-se considerar que o primeiro capítulo inclui uma introdução geral sobre o tema abordado. Em seguida são apresentados o objetivo geral do estudo bem como os objetivos específicos. O terceiro capítulo refere-se a metodologia empregada na tese, para a qual é apresentada a área de estudo, a base de dados e a análise dos dados. Para este capítulo existem 2 apêndices relacionados, o Apêndice I que trata sobre o banco de dados que foi desenvolvido para esta tese enquanto que o Apêndice II trata das análises estatística que foram realizadas antes do ajustes dos modelos propriamente ditos e que serviram como uma forma de consolidação do banco de dados. Os capítulos subsequentes apresentam os resultados gerais dos dois trabalhos desenvolvidos nesta tese e são complementados pelos Apêndices III e IV que correspondem aos trabalhos “Long-term Rajiformes (Chondrichthyes: Elasmobranchii) spatio-temporal dynamics in the southernmost states of Brazil” e “ Modelagem espacial Bayesiana para a riqueza de elasmobrânquios do

extremo sul do Brasil”. No capítulo 5 são apresentadas as conclusões gerais do trabalho e no capítulo 6 as referências bibliográficas. A tese ainda conta com um apêndice V onde os códigos em R deste trabalho são apresentados.

## **2. OBJETIVOS**

### **2.1 Objetivo geral**

- Analisar a variabilidade espacial e temporal na composição e abundância das assembleias de elasmobrânquios e verificar a influência de fatores bióticos e abióticos do ambiente do Oceano Atlântico Sudoeste sobre o padrão de ocupação destas assembleias.

### **2.2 Objetivos específicos**

- Calcular índices de abundância apropriados para o táxon (Captura por unidade de esforço – CPUE em kg/área varrida) para avaliar a distribuição espacial de Rajiformes ao longo de uma série temporal (Apêndice III);
- Verificar a composição e possíveis alterações entre as associações de Rajiformes, em diferentes estações do ano e décadas, por meio análises multivariadas de agrupamento e matriz de correlação (Apêndice III);
- Verificar a influência das variáveis ambientais abióticas (profundidade, temperatura e salinidade) e bióticas (biomassa de elasmobrânquios e de teleósteos capturada) sobre a distribuição das espécies de elasmobrânquios (Apêndice IV);
- Predizer a distribuição das espécies a partir de modelos estatísticos e imagens de satélite (Apêndice III e IV);
- Identificar áreas e períodos prioritários para a conservação da biodiversidade de elasmobrânquios na costa do RS (Apêndice III e IV).

### 3. MATERIAL E MÉTODOS

#### 3.1 Área de estudo

Este estudo está delimitado à região da plataforma e talude continental do Sul do Brasil, uma área de aproximadamente 100.000 km<sup>2</sup> compreendida entre o Cabo de Santa Marta Grande (28° 40' S) e o Arroio Chuí (34° 40' S) (Seeliger et al. 1998).

No que diz respeito à topografia, a plataforma apresenta uma largura de 110 km em frente ao Cabo de Santa Marta Grande, alargando-se até 140 km em frente a cidade de Tramandaí, atingindo seu máximo de 170 km na barra do Rio Grande e retornando a 140 km de extensão em frente ao Chuí (Haimovici et al. 2007). As isóbatas estão dispostas aproximadamente paralelas à costa e a declividade da plataforma é suave (2 m/km), contrastando com a inclinação abrupta do talude (80 a 130 m/km) (Haimovici, 2008).

Mapas detalhados do relevo de fundo identificam nesta área importantes províncias fisiográficas na plataforma (Terraço do Rio Grande, Vale de Porto Alegre, Vale de Santa Maria) e no talude (Vale do Rio Grande e Cone do Rio Grande) (Ferreira et al. 2005). A cobertura do leito oceânico é composta principalmente por lama, areia muito fina e areia fina e predominam os sedimentos de origem litoclástica (Figueiredo & Tessler 2004).

Sobre a hidrologia da região, sabe-se que a plataforma é ocupada por duas massas de água denominadas de Água Subtropical de Plataforma - ASTP (> 18° C e 33,5 - 35,9 ups) e Água Subantártica da Plataforma – ASAP ( 10 – 14 °C no inverno e 12 – 20 °C no verão e 33,5 - 34,1 ups), ambas originadas da mistura de águas oceânicas e continentais. Na superfície oceânica do talude, são descritas duas massas de água que fluem para o sentido sul, sendo a Água Tropical da Corrente do Brasil – AT (> 20 °C e 36,4 ups) localizada sobre a Água Central do Atlântico Sul – ACAS ( limite superior 20 °C e inferior de 3 – 6 °C e 34, 6 - 36,0 ups) e fluindo em direção norte e sob a ACAS, a Água Intermediária Antártica – AIA ( 3 – 6 °C e 34,2 - 34,6 ups. (Pinet, 2009).

Como um todo, a área de abrangência deste estudo é fortemente influenciada pelas condições oceanográficas e atmosféricas que atuam sobre o sudoeste do Oceano Atlântico. Sazonalmente existe um deslocamento da Convergência Subtropical do Atlântico Sul (CSAS) alcançando na margem oeste deste oceano as latitudes de 46°S no

verão e 30°S no inverno (Piola et al. 2008). Esta CSAS é formada pela confluência de águas que fluem em direções opostas, sendo uma quente, oriunda da Corrente do Brasil (CB) e a outra fria, da Corrente das Malvinas (CM). A posição da CSAS e o volume de descarga continental são os principais forçantes nesta área. Em conjunto determinam as características físicas e químicas as quais por sua vez governam a estrutura biológica e os processos ecológicos. A interação entre a CB oligotrófica e a CM com águas ricas em nutrientes e descarga continental tornam a região altamente produtiva no sentido biológico (Seeliger & Odebrecht, 1998).

A produtividade da plataforma é relativamente alta no inverno e na primavera por causa do aporte de nutrientes de origem terrígena e do deslocamento para o norte da água subantártica de plataforma. Outros mecanismos de fertilização são as ressurgências costeiras originadas pela forte ação dos ventos de nordeste atuantes na primavera e verão e que provoca a divergência de águas superficiais e causa a advecção da ACAS sobre a plataforma (Odebrecht & Castello, 2001; Haimocivi et al. 2007).

Na condição atmosférica da região, devido à presença do centro de alta pressão do anticiclone do Atlântico Sul, o padrão de ventos predominantes são os NE (velocidade média de 5m s<sup>-1</sup>) ao longo do ano seguido por ventos de SW (velocidade média de 8m s<sup>-1</sup>) durante a passagem de frentes frias, as quais são mais comuns no inverno do que no verão (Wainer & Taschetto 2006).

### **3.2 Base de dados**

Foram analisadas as capturas e esforços de pesca elasmobrânquios demersais amostrados em 27 cruzeiros científicos realizados pelo Navio Oceanográfico Atlântico Sul da Universidade Federal do Rio Grande – FURG. Além das informações sobre as espécies de elasmobrânquios amostrados dispõe-se de medidas de temperatura e salinidade da superfície e fundo bem como registros batimétricos coletados *in situ*.

Em cruzeiros pretéritos (1980 e 1990) as medidas de temperatura de superfície foram obtidas com termômetro, temperaturas de fundo por meio de termômetro de inversão e salinidade através de salinômetros. Cruzeiros científicos mais recentes (2001, 2002 e 2005) tiveram seus parâmetros ambientais registrados com instrumentos de CTDs (do inglês *Conductivity, Temperature and Depth*). Em todos os cruzeiros o



registro batimétrico dos lances de pesca foram obtidos através de Ecossonda Científica.

Um banco de dados próprio foi desenvolvido para armazenar de forma sistemática as informações georreferenciadas desta série de prospecções pesqueiras obtidas no período compreendido entre 1980 e 2005 (Apêndice I).

### 3.3 Análise de dados

Antes de ajustar os modelos de distribuição de espécies, foi realizada uma análise exploratória dos dados, na qual foram verificados as distribuições de probabilidade das variáveis resposta previamente definidas, a frequência de zero nas amostras, os *outliers* em todos conjuntos de dados e a colinearidade entre as covariáveis. No Apêndice II são apresentadas, de maneira detalhada, todas estas análises.

Modelos hierárquicos Bayesianos da abundância (Apêndice III - Long-term Rajiformes (Chondrichthyes: Elasmobranchii) spatio-temporal dynamics in the southernmost states of Brazil) e riqueza de espécies (Apêndice IV - Modelagem espacial Bayesiana para a riqueza de elasmobrânquios do extremo sul do Brasil ) serão utilizados para analisar a distribuição espacial dos elasmobrânquios na costa do Rio Grande do Sul.

A estruturação dos modelos hierárquicos apresenta a seguinte forma:

$$X|\eta \sim F(\exp(\eta))$$

$$\eta_i = \beta_{0\eta} + \sum_{m=1}^M \beta_{m\eta} X_{m_i\eta} + \sum_{l=1}^L f_l(v_{li})$$

$i = 1, \dots, n$  Onde  $\eta = \{\eta(S): s \in \mathbb{R}^d\}$  denota o preditor linear da variável resposta (CPUE e Riqueza);  $\beta_0$  é um escalar representando o intercepto;  $\beta_m$  constitui um vetor dos regressores que quantifica o efeito de covariáveis  $X_m$  sobre a variável resposta;  $f_l(.)$  são funções desconhecidas e definidas para um conjunto de covariáveis  $v_l$ . O termo  $f(.)$  pode ser usado tanto para relaxar a linearidade das covariáveis, à exemplo de funções suavizadoras, quanto para introduzir efeitos aleatórios, tais como o efeito espacial (W) e o efeito temporal (Z) (Rue et al. 2009).

Foram utilizados Modelos Gaussianos Latentes, transformados via Equação

Diferencial Estocástica Parcial (módulo SPDE) em Campos aleatórios Markovianos Gaussianos (GMRF) que assim como os GLMs permitem a inclusão de covariáveis em uma estrutura aditiva e ainda captam a estrutura de dependência espacial e temporal dos dados.

A estimativa dos parâmetros dos modelos será obtida via estatística Bayesiana utilizando o método *Integrated Nested Laplace Approximations* – INLA, o qual substitui as simulações por aproximações determinísticas, que além de serem muito mais precisas, são consideravelmente mais rápidas no quesito computacional e adequadas ao tipo de modelo proposto (Rue et al. 2016).

A seleção dos modelos foi feita pelo método manual da inserção sequencial de variáveis (*forward stepwise*). A qualidade dos modelos foi acessada mediante DIC (*Deviance Information Criterion*, Spiegelhalter et al. 2002), WAIC (*Watanabe Akaike Information Criterion*, Watanabe, 2010) enquanto que o poder preditivo por meio do LCPO (*Logarithm Conditional Predictive Ordinate*, Roos & Held, 2011). Todas as análises serão implementadas no software R versão 3.3.1 (R *Development Core Team*) a partir da biblioteca R – INLA e outras bibliotecas acessórias.

#### **4. RESULTADOS GERAIS**

Esta sessão apresenta apenas uma breve síntese dos resultados referentes às análises

de distribuição espacial de abundância e riqueza de elasmobrânquios no litoral do Estado do Rio Grande do Sul. A apresentação completa dos resultados esta inserida nos Apêndices III e IV.

#### **4.1 Resultados do manuscrito “Long-term Rajiformes (Chondrichthyes: Elasmobranchii) spatio-temporal dynamics in the southernmost states of Brazil.” (Apêndice III)**

Uma série histórica de cruzeiros científicos (1980 – 2002) com pesca de arrasto de fundo foi utilizada para avaliar a dinâmica dos Rajiformes. Para a análise dos dados, as capturas foram previamente divididas por décadas e épocas do ano. Desta maneira, foram obtidos 6 subconjuntos de dados sendo para a década de 1980 correspondente as 4 estações do ano (primavera, verão, outono e inverno); para a década de 1990 apenas para primavera e no período de 2001-2002 representam o outono-inverno em conjunto (Tabela 1)

As raias da ordem Rajiformes registradas no banco de dados são: *Atlantoraja castelnaui*, *A.*

*cyclophora*, *A. platana*, *D. leptocauda*, *Gurgesiella dorsalis*, *Psammobatis bergi*, *P. extenta*, *P. lentiginosa*, *P. rutrum*, *Rioraja agassizii*, *Sympterygia acuta* e *S.*

*bonapartii*. O conjunto da captura dessas raias foi utilizado para o cálculo de CPUE (kg/ área varrida) uma vez que a análise multivariada de agrupamento e a matriz de correlação demonstrou que as assembleias de raias são agrupada em um cluster principal com 95% de intervalo de significância (Fig. 10). Fato que explica a captura conjunta dessas raias no arrasto de fundo.

De acordo com os resultados obtidos pelos diferentes modelos da década de 80' e 90' (material suplementar – Apêndice II), observou-se que batimetria foi a variável ambiental mais relevante para descrever a sazonalidade de Rajiformes na região de estudo. Segundo os parâmetros destes modelos existe uma relação negativa com a profundidade, indicando que os Rajiformes têm preferência por águas mais próximas a costa (Tabela 3). Para os modelos de 2001 – 2002 a variável que demonstrou ser relevante foi a salinidade de superfície (material suplementar – Apêndice II), a qual

apresenta uma relação negativa com a variável resposta (Tabela 3) indicando a preferência dos Rajiformes por águas menos salinas. Os resultados do modelo GAM para CPUE ~Salinidade de superfície não encontrou nenhum ótimo de salinidade para a distribuição dos Rajiformes (Figura 8).

Mapas preditivos da CPUE média, com respectivo desvio padrão, foram elaborados para todos os modelos e permitiram identificar a localização espacial das áreas de concentração dos Rajiformes (Figuras 2 a 7). De acordo com o mapa e o cálculo do parâmetro  $\rho$  (rho), que quantifica a correlação espacial entre os dados, as áreas de ocupação espacial dos Rajiformes é maior no outono e inverno enquanto que na primavera e verão os animais apresentam uma distribuição mais agrupada.

Através do cálculo padronizado de CPUE, por meio do intercepto dos modelos, foi possível comparar as abundâncias entre os diferentes conjuntos de dados. Durante o período estudado observou-se um declínio acentuado na CPUE dos Rajiformes (material suplementar – Apêndice IV), resultados estes que corroboram com os cálculos da CPUE observada. Verificou-se que a redução da abundância inicia pelas estações de primavera, outono, verão e inverno da década de 80', seguida por primavera 90' e finalmente o período entre 2001 – 2002.

De forma geral, observou-se uma consistência na concentração de Rajiformes ao sul do RS e, assim, traz-se mais uma evidência para reforçar a proposição de criação de uma área marinha protegida nesta região para a conservação de várias espécies de elasmobrânquios.

#### **4.2 Resultados do Artigo ” Modelagem espacial Bayesiana para a riqueza de elasmobrânquios do extremo sul do Brasil” (Apêndice IV)**

Para este artigo buscou-se avaliar a relação entre a riqueza de elasmobrânquios e um conjunto de covariáveis abióticas (batimetria, temperatura e salinidade) e bióticas (biomassa de elasmobrânquios e biomassa de peixes teleósteos). Estes dados foram oriundos do cruzeiro científico realizado em Fevereiro de 2005 com 64 estações oceanográficas e lances de pesca efetuados entre 7 a 20 m de profundidade (Figura 1).

A riqueza de capturas de elasmobrânquios foi representada por um conjunto de 19 espécies, cujas mais frequentes em termos de ocorrência foram *Sympterygia acuta* (39%), *Narcine brasiliensis* (24%) e *Sympterygia bonapartii* (13%). Já em termos de

biomassa, as espécies mais abundantes foram de *S. acuta* (78 kg), *N. brasiliensis* (69 kg) e *Myliobatis ridens* (64 kg) (Tabela 1)

De acordo com o modelo geoestatístico aplicado para a riqueza destes elasmobrânquios, constatou-se que a riqueza é positivamente relacionada à batimetria e negativamente à temperatura e salinidade (Tabela 3). Portanto existe uma tendência das maiores riquezas da fauna de elasmobrânquios em se concentrar em águas mais profundas, menos salinas e mais frias. Estes resultados representados no mapa preditivo salientou um *hotspot* de concentração de elasmobrânquios no extremo sul do estado do Rio Grande do Sul (Figura 4)

Embora os resultados deste estudo foram provenientes de apenas um único cruzeiro realizado em no verão, os resultados aqui encontrados já permitem evidenciar a ocupação espacial dos elasmobrânquios na costa do RS. Além disso, esses resultados também servem como indício para comprovar a importância da região do extremo sul do Brasil como uma área marinha protegida.

## 5. CONCLUSÕES

As técnicas estatísticas utilizadas ao longo desta tese podem se enquadrar à classe de modelos conhecidas como modelos de distribuição de espécies (SDMs, do inglês Species Distribution Models), na qual a abundância (ou registro de ocorrência) de uma ou mais espécies se relaciona a um conjunto de variáveis abióticas e bióticas. Como resultado final, uma representação da distribuição espacial é construída a qual reflete também as áreas cujas condições ambientais e/ou ecológicas são mais favoráveis à sobrevivência da espécie em estudo.

O presente trabalho utilizou modelos hierárquicos Bayesianos, com efeito aleatório espacialmente estruturado, para estimar e prever a abundância de Rajiformes e a riqueza de elasmobrânquios em função de covariáveis temporais e ambientais com base em dados coletados em pesca científica durante o período de 1980-2005.

Numa série temporal de dados de captura de Rajiformes foram caracterizadas as áreas de ocorrência preferencial das espécies de acordo com as estações do ano. Também foi verificado que diferentes associações e relacionamentos entre as espécies foram estabelecidas com passar dos anos e podemos suspeitar que essas mudanças sejam pela diminuição de alguns estoques resultando num rearranjo de assembleias de Rajiformes. As informações de CPUE observadas, bem como as previsões efetuadas a partir dos modelos complementam-se no sentido de demonstrarem a redução dos estoques e atentam para a necessidade de ações imediatas para conservação das raias que estão expostas ao risco de colapso como evidenciado em vários estudos populacionais deste grupo desde a década de 1990 (Vooren e Klippel, 2005)

Foram reconhecidas áreas preferenciais de concentração de elasmobrânquios durante a estação do verão, cujas condições ambientais estão relacionadas às variáveis temperatura, salinidade e profundidade. Uma vez que a etapa preditiva dos modelos foi elaborada com base em dados climatológicos para períodos mais quentes, infere-se que este seja o padrão preferencial de ocupação da região durante o verão. Considerando que dezembro, janeiro e fevereiro correspondem ao período de reprodução de muitas espécies de elasmobrânquios na costa do RS, este estudo contribuiu com novas informações a respeito da distribuição das espécies em períodos críticos do ciclo de vida.

Este trabalho foi realizado numa abordagem Bayesiana usando modelos espaciais versáteis e técnicas computacionais muito eficientes. Os resultados gerados pelos modelos, combinados com elementos da geoestatística apresentam uma interpretação gráfica bastante intuitiva. A partir dos mapas é possível identificar espacialmente regiões de maior e menor índice de abundância de Rajiformes assim como para riqueza de elasmobrânquios.

Vale ressaltar que os resultados obtidos dos modelos refletem a distribuição espacial dos elasmobrânquios considerando as condições ambientais presentes no momento da sua captura. Para uma análise temporal contínua do uso do habitat, a incorporação de variáveis que representem anomalias térmicas pode contribuir para que tenhamos modelos mais realistas a respeito do fenômeno investigado.

Os modelos apresentados nesta tese também podem ser utilizados fora do contexto de cruzeiros científicos. No entanto, para caracterizar o cenário atual das pescarias de elasmobrânquios é necessário uma estatística pesqueira de qualidade, com informações de geolocalização das capturas, as quais podem ser obtidas de mapas de bordo ou de observadores de bordo. A correta identificação das espécies nos desembarques contribui para que se possa desenvolver modelos refinados. Considerando que desembarques de tubarões e raias ocorrem como carcaças (decapitados e sem nadadeiras), identificação a nível de espécie é possível por meio de caracteres visuais e medidas corporais.

Este trabalho evidencia a importância da região costeira e adjacente a costa do Rio Grande do Sul como área relevante para a conservação de um grande número de famílias de elasmobrânquios, as quais devem e merecem ser preservadas a fim de evitar sua extinção antes mesmo que tenhamos o conhecimento da sua importância no ecossistema marinho.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

BAUM, J. K.& WORM, B. 2009. Cascading top-down effects of changing oceanic predator abundance. *Journal of Animal Ecology*, (78): 699 – 714.

BLANGIARDO, M. & CAMELETTI, B. G. 2015.

Spatio and Spatio-temporal Bayesian Models  
with R -INLA. United Kingdom, John Wiley &  
Sons Ltd. 322p.

BORNATOWSKI, H.; ANGELINI, R.; COLL, M. ; BARRETO, R. P. ; AMORIM, A. F.  
Ecological role and historical trends of large pelagic predators in a subtropical marine  
ecosystem of the South Atlantic. *REVIEWS IN FISH BIOLOGY AND FISHERIES* , p.  
1-19, 2017.

CARRIER, JC.; PRATT, HL. & CASTRO, JL. 2004 Reproductive biology of  
Elasmobranchs. In: CARRIER, JC; JA, MUSICK JA & MR, HEITHAUS (ed). *Biology  
of Sharks and Their Relatives*. CRC Press. pp.269-286.

COSANDEY-GODIN, A.; KRAINSKI, ET.; WORM, B. & FLEMMING, JM. 2015.  
Applying Bayesian spatiotemporal modelst o fisheries bycatch in the Canadian Arctic.  
*Can. J. Fish. Aquat. Sci.*, 72: 1-12.

DULVY, NK.; SADOVY , Y & REYNOLDS, JD. . 2003. Extinction vulnerability in  
marine populations. *Fish and Fisheries*., 4: 25 - 64.

DULVY, N. K., FOWLER, S. L., MUSICK, J. A.,

CAVANAGH, R. D., KYNE, P. M., HARRISON, L. R., CARLSON, J. K.,  
DAVIDSON, L.

N. K., FORDHAM, S. V., FRANCIS, M. P.,

POLLOCK, C. M., SIMPFENDORFER, C. A.,

BURGESS, G. H.,CARPENTER, K. E., COMPAGNO, L. J. V., EBERT, D. A.,



GIBSON, C.,

HEUPEL, M. R., LIVINGSTONE, S. R., SANCIANGCO, J. C., STEVENS, J. D.,  
VALENTI,

S. & WHITE, W. T. 2014. Extinction risk and  
conservation of the world's sharks and rays.

<https://doi.org/10.7554/eLife.00590.001>

FERREIRA, CS.; MADUREIRA, LSP.; KLIPPEL, S.; WEIGERT, S.; HABIAGA,  
RGP. & DUVOISIN, AC. 2005. Mapas do relevo marinho das regiões sudeste, sul e  
central do Brasil: acústica e altimetria por satélite. São Paulo, Instituto Oceanográfico –  
USP. 41p.

FIGUEIREDO, AG & TESSLER, MG. 2004. Topografia e composição do substrato  
marinho da Região Sudeste-Sul do Brasil. São Paulo, IO – USP. 65 p.

FISCHER, LG. Caracterização da estrutura da comunidade nêctônica demersal da  
plataforma externa e talude superior da região sul do Brasil (28°S – 34°S). 2008. 123f.  
Dissertação (Mestrado em Oceanografia Biológica) – Instituto de Oceanografia,  
Universidade Federal do Rio Grande, Rio Grande. 2008.

FRANKLIN, J., 2010. Mapping species distributions: spatial inference and prediction,  
first edition.

Cambridge University Press, New York.

GOVERNO DO RIO GRANDE DO SUL, 2018.

Lei Nº 15.223. Institui a Política Estadual de Desenvolvimento Sustentável da Pesca no  
Estado do Rio Grande do Sul e cria o Fundo Estadual da Pesca. DOE nº 172, de 6 de  
setembro de 2018)

HAIMOVICI, M.; CASTELLO, J.P.; VOOREN, C.M. 1998. Pescarias. In: SEELIGER,  
U; C, ODEBRECHT & JP CASTELLO (eds). Os Ecossistemas Costeiro e Marinho do  
Extremo Sul do Brasil. Rio Grande: Ecoscienza, Chap. 9 p. 205 – 218.

HAIMOVICI, M; ÁVILA-DA-SILVA, AO; MIRANDA, LV. & KLIPPEL, S. 2007. Prospecções na região sudeste-sul. In: HAIMOVICI, M. (Org.). A prospecção pesqueira e abundância de estoques marinhos no Brasil nas décadas de 1960 a 1990. Brasília: MMA/SMCQ, p. 35- 73.

HAIMOVICI, M.; ROSSI-WONGTSCHOWSKI, CLDB.; BERNARDES, RA.; FISCHER, LG .; VOOREN, CM.; SANTOS, RA.; RODRIGUES, AR., & SANTOS., S 2008. Prospecção pesqueira de espécies demersais com rede de arrasto-de-fundo na Região Sudeste-Sul do Brasil. São Paulo, Instituto Oceanográfico – 184p.

HEITHAUS, MR.; FRID, A.; WIRSING, AJ. & WORM, B. 2008. Predicting ecological consequences of marine top predator declines. 2008. Trends in Ecology & Evolution., (23): 202 – 210.

INSTITUTO CHICO MENDES. 2016. Avaliação do risco de extinção dos elasmobrânquios e quimeras no Brasil: 2010-2012. Disponível em: [http://www.icmbio.gov.br/cepsul/images/stories/biblioteca/download/trabalhos\\_tecnicos/pub\\_2016\\_avaliacao\\_elasmo\\_2010\\_2012](http://www.icmbio.gov.br/cepsul/images/stories/biblioteca/download/trabalhos_tecnicos/pub_2016_avaliacao_elasmo_2010_2012).

INSTITUTO CHICO MENDES. 2019. Sistema de Avaliação do Estado de Conservação da Biodiversidade – SALVE. Disponível em <https://salve.icmbio.gov.br/salve-consulta/>

KINAS, P. G. & ANDRADE, H. A. 2010. Introdução à Análise Bayesiana (com R). Porto Alegre, maisQnada. 240P

KRAINSKI, ET & LINDGREN, F. 2014. The R-INLA tutorial: SPDE models. < <http://www.math.ntnu.no/inla/r-inla.org/tutorials/spde/spde-tutorial.pdf> > acesso em março de 2015.

LINDGREN, F.; RUE, H. & LINDSTRÔM, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach (with discussion). J. R.

Stat. Soc. Ser. B., 73: 423-498.

MENEZES, N. A., BUCKUP, P. A., FIGUEIREDO, J. D. & MOURA, R. D. 2003. Catálogo das espécies de peixes marinhos do Brasil. São Paulo, Museu de Zoologia da Universidade de São Paulo. 159 p.

MYERS, RA.; BAUM, JK.; SHEPHERD, TD.; POWERS, SP. & PETERSON, CH. 2007. Cascading effects of the Loss of Apex predatory Sharks from the Coastal Ocean. Science., (315): 1846

ODEBRECHT, C. & CASTELLO, J.P. 2001. The Convergence Ecosystem in the Southwest Atlantic. In: Seeliger, U. & B. Kerjve (eds.). Coastal Marine Ecosystems of Latin America, Springer: 147-165.

PENNINO, MG.; MUÑOZ, F.; COESA, D.; LÓPEZ-QUÍLEZ, A.& BELLIDO, JM. 2013. Modeling sensitive elasmobranch habitats. Journal of sea research 83: 209 - 218

PINET, PR. 2009. Invitation to oceanography 5th edition. Canada. Jones and Bartlett Publishers

PIOLA, AR.; MöLLER, OO.; GUERREIRO, RA. & CAMPOS, E. 2008. Variability of the subtropical shelf front off eastern South America: Winter 2003 and summer 2004. Continental Shelf Research 28: 1639– 1648

R Development Core Team, 2018. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

R – INLA. <http://www.r-inla.org/>

ROOS, M. & HELD, L., 2011. Sensitivity analysis in Bayesian generalized linear mixed models for

binary data. *Bayesian Analysis* 6, 259-278.

ROOS, N. C.; CARVALHO, A. R.; LOPES, F. M. & PENNINO, M. G. 2015. Modeling sensitive parrotfish (Labridae: Scarini) habitats along the Brazilian coast. *Mar. Env. Res.*, 110 : 92-100.

RUE, H.; MARTINO, S. & CHOPIN, N. 2009. Approximate bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Statist. Soc. B.*, 71: 319-392.

RUE, H.; RIEBLER, A.; SØRBYE, SH.; ILLIAN, JB.; SIMPSON, DP & LINDGREN, F. 2016. Bayesian Computing with INLA: A Review . Available at HYPERLINK "http://arxiv.org/abs/1604.00860"<http://arxiv.org/abs/1604.00860>

RUFENER M. C., KINAS P. G., NÓBREGA M. F.

& OLIVEIRA J. E. L. 2017 Bayesian spatial predictive models for data-poor fisheries. *Ecological Modelling* 348: 125–134

SIMPSON, D.; LINDGREN, F. & RUE, H. 2012. In order to make spatial statistics computationally feasible, we need to forget about the covariance function. *Environmetrics*, 23: 65-74.

SEELIGER, U.; ODEBRECHT , C & CASTELLO, JP (eds) 1998. Os Ecossistemas Costeiro e Marinho do Extremo Sul do Brasil. Rio Grande: Ecoscientia – 326p

SPIEGELHALTER, D.J., BEST, N., CARLIN, B.P. & VAN DER LINDE, A., 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society Series B (Statistical Methodology)* 64, 4., 1-34. doi: 10.1111/1467-9868.02022.

SUDEPE. 1972. “Mestre Jeronimo” Mapas de Bordo No° 1 1-4/72. M. A. - SUDEPE.

SUDEPE. 1974. “Zeus” Mapas de Bordo No° 1 1-8/74. M. A. - SUDEPE.

VELASCO, G.; ARAÚJO, J. N.; CASTELLO, J. P.; ODDONE, M. C. 2007. Exploring MSY strategies for elasmobranch fishes in an ecosystem perspective. *Pan. J. Aquat. Sci.*, 2(2): 63-178.

VOOREN, CM. 1998. Ambientes Costeiros e Marinhos e sua Biota, Elasmobrânquios Demersais. In: SEELIGER, U; C, ODEBRECHT & JP CASTELLO (eds). *Os Ecossistemas Costeiro e Marinho do Extremo Sul do Brasil*. Rio Grande: Ecoscientia, Chap. 6 p. 157 – 162.

VOOREN, CM. & KLIPPEL, S. (eds) 2005. *Ações para conservação de tubarões e raias no sul do Brasil*. Porto Alegre: Igaré - 261 p

VOOREN, C.M & ODDONE, M.C. 2018. La diversidad de los condriictios del extremo Sur de Brasil: las especies, sus origenes, y sus modos de reproducción. In: Muniz, P., Venturini, N., Brugnoli, E. & Conde, D.. (Org.). *Ciencias marino-costeras en el umbral del Siglo XXI*. 1ed.Ciudad de México: AGT Editor, S.A., 2018, v. , p. 171-214.

WAINER, I & TASCHETTO, AS. 2006. Climatologia da região entre o Cabo de São Tomé (RJ) e o Chuí (RS). Diagnóstico para os períodos relativos aos levantamentos pesqueiros do Programa REVIZEE. In: ROSSI – WONGTSCHOWSKI, CLDB & MADUREIRA, LSP. (ed.). *O ambiente oceanográfico da plataforma continental e do Talude na região sudeste-sul do Brasil*. Edusp, São Paulo, Chap. 2: 121-160

WALKER, T. 2005. Management measures. In: MUSICK, JA & BONFIL, R (ed). *Management techniques for elasmobranch fisheries*. Rome: Food and Agriculture Organization of the United Nations. Chap. 13: 216 – 243.

WATANABE, S. 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11, 3571-3594

## APÊNDICE I

### *Uma abordagem baseada em banco de dados*

#### **1. INTRODUÇÃO**

Inicialmente, alguns aspectos técnicos são relevantes para o entendimento de uma abordagem sistemática baseada em coleta, armazenamento e processamento de dados. Bancos de dados salvaguardam coleções de dados que apresentam características distintas e que estão dispostas em locais específicos denominados de entidade. As entidades nada mais são do que tabelas que detêm dados que se assemelham entre si e para as quais existe uma coleção de atributos definidos. Nestes atributos os dados são identificados como registros (Silberschatz et al. 2016).

Para formar o diagrama Entidade – Relacionamento e permitir que associações sejam estabelecidas entre entidades, são utilizados atributos específicos denominados de chaves primárias ou id. A maneira pela qual as ligações entre as Entidades podem ser efetuadas, a partir da troca de chaves primárias, ocorre pela cardinalidade do banco.

Quando chaves primárias são trocadas entre entidades, esta passa a ser chamada de chave estrangeira.

Como o volume de dados é algo em aberto, é importante deixar claro que essas relações são únicas e não se sobrepõem. Com isso, é assegurado que pesquisas futuras relacionando informações de diversas tabelas resultem em informações consistentes e válidas sob o ponto de vista de análise de dados.

Os principais tipos de cardinalidade são as nulas, simples e múltiplas. No caso das cardinalidades nulas, um registro de uma entidade não apresenta relacionamento com o de outra entidade. Em termos gerais, representam conjunto de dados que estão apenas armazenando valores, sem qualquer associação ou referência com outras tabelas (Schwartz et al. 2008).

Na cardinalidade simples existe uma ligação um – para – um, ou seja, valores de uma tabela podem ser referência para outra no processo de execução de análise dos dados. No caso da cardinalidade múltipla um registro na entidade tem ligação como múltiplos registros de outra entidade ou ainda ligação múltiplos – para – múltiplos (Russel, 2004).

Uma vez que as entidades com seus atributos e relacionamentos estejam devidamente estruturadas, pode-se estabelecer o modelo de banco de dados. Levando em consideração que estas características são estruturais e após sua implementação são imutáveis.

Intuitivamente o modelo de organização de banco de dados mais amplo é o tipo relacional (armazenamento em tabelas). Este modelo diferentemente dos tipos hierárquicos e em rede permite que a partir de um registro no banco seja possível acessar diferentes níveis de informação (Crockett 2008). Assim, o modelo relacional se torna mais eficiente que os demais em razão desta capacidade em estabelecer não apenas ligações entre os dados como também relações entre si, ou seja, suas funcionalidades permitem integrar gerenciamento de dados dos tipos tradicionais aos objetos complexos como séries temporais e dados geoespaciais (Harrington 2002).

Contudo, independente do modelo de organização do banco de dados, para assegurar sua qualidade operacional o mesmo deve ter como princípios a durabilidade, o isolamento, a consistência e a atomicidade (representados pelo acrônimo DICA).

Por durabilidade se entende que todo dado que é colocado, alterado, manipulado

no banco tem que permanecer durável dentro deste. Já o isolamento refere-se às transações simultâneas que ao serem executadas não exerçam interferência uma na outra. A consistência está relacionada à capacidade do banco de levar os dados de um estado consistente para outro na mesma categoria. Por fim a atomicidade diz respeito à transação, a qual deve ser íntegra no sentido de que ela (a transação) acontece completa ou então nada acontece (Date 2016).

Esses princípios são assumidos como inerentes a qualquer aplicação que vise uma estabilidade e escalabilidade. Dessa forma, sistemas mais robustos e complexos podem ser implementados visando facilitar o acesso e consequente processamento da informação, sabendo que suas estruturas são coerentes e condizentes com os relacionamentos propostos no modelo relacional.

Assim, uma abordagem preliminar mais teórica serve para esclarecer alguns aspectos técnicos com os quais um serviço baseado em coleta e processamento de dados deve preocupar-se em fase de planejamento para posterior implementação computacional, visando garantir a integridade, qualidade, acesso e processamento adequado das informações obtidas.

## **2. MATERIAL E MÉTODOS**

### **2.1 A interface de acesso aos dados**

O sistema para cadastro e controle foi desenvolvido utilizando linguagem PHP e um banco de dados MySQL. A estrutura e interface de acesso ao sistema seguiu os padrões encontrados nos cadastros físicos, em pastas, para assim preservar toda informação previamente obtida de maneira manual. Os resguardos referentes aos tipos de variáveis, sendo reais, inteiras, data e hora e textos literais, foram mantidos para que futuras análises pudessem ser conduzidas a fim de explorar o conjunto de dados e transformar a informação contida nas pastas em conhecimento útil para o agente utilizador do ambiente computacional proposto.

Consequentemente, configurações personalizadas dessa ferramenta foram implementadas para permitir a conexão direta entre o sistema estatístico R e o servidor MySQL. Dessa forma, toda análise exploratória e processamento estatístico podem também ser realizados remotamente e em tempo real independente do acesso à



plataforma. Devido à estruturação realizado pelo sistema proposto e a escolha adequada do modelo relacional do banco de dados, foi possível manter a hierarquia e relacionamento das tabelas e informações coletadas visando agilizar o processo de análise e tomada de decisão.

Dessa forma, entrega-se um ambiente computacional completo, capaz não apenas de armazenar dados, mas também de fornecer acesso e processamento externo com aplicações baseadas em estatística computacional já consolidadas em ambiente tanto acadêmico quanto profissional.

### **3. RESULTADOS**

#### **3.1 Estudo de caso: cruzeiros científicos**

Um banco de dados próprio para consulta particular foi estruturado para cadastrar as capturas de elasmobrânquios em cruzeiros científicos realizados pela FURG. Este sistema proposto tem por objetivo otimizar a busca dos dados tornando as análises mais eficientes no que diz respeito à manipulação de informações e todo processo de transformação desses em conhecimento aplicado.

Porém, antes de realizar este processo foi necessário um levantamento sobre os dados disponíveis no extinto laboratório de Elasmobrânquios do Instituto de Oceanografia. Com orientação e participação do Prof. Dr. Carolus Maria Vooren, responsável pela coleta destes dados, foi realizado um inventário (Tabela I) deste material e solicitado oficialmente ao Instituto de Oceanografia a incorporação deste acervo como patrimônio da FURG.

Para realizar este trabalho foram reunidos relatórios de cruzeiros, cadernos de bordo, fichas de capturas das espécies e planilhas de dados registradas em papel ou arquivo digital. Foi constatado que os dados coletados apresentavam falta de padronização nos registros destas capturas e por isso foi iniciado o projeto de estruturação e desenvolvimento de um sistema para cadastro e organização dessas informações na forma digital.

Primeiramente foram idealizadas as entidades do banco e suas respectivas variáveis de interesse, cuja descrição e número de itens cadastrados encontram-se na Tabela 2. Particularmente na entidade *Grupos*, o banco vai além de sistematizar dados e

busca resgatar informações de indivíduos coletados durante os cruzeiros e que se encontram preservados na coleção científica do Laboratório de Elasmobrânquios. Esse acervo tem valor imensurável para o ensino e pesquisa de condrictes, pois guarda material biológico para estudos de idade, crescimento, reprodução, alimentação e morfologia de espécies.

Além disso, no banco de dados, cada indivíduo cadastrado apresenta um id único, o que possibilita que a ficha de captura original da espécie seja identificada e resgatada a qualquer instante para qualquer tipo de confirmação de informação.

De acordo com os procedimentos adotados, a integridade das informações incluídas no banco de dados é assegurada, pois todos os dados referentes aos cruzeiros foram manualmente inseridos no banco de dados. Com a estrutura proposta é possível garantir que todos os dados, antes informados em tabelas manuscritas, agora possam ser facilmente acessados e processados buscando respostas e identificando os padrões necessários para o andamento da pesquisa.

Para garantir que eventuais incidentes não afetem o bom andamento do trabalho realizado por todas as partes, e a integridade dos dados não seja violada por qualquer circunstância imprevista, um protocolo de segurança foi estabelecido e implementado visando garantir a segurança dos dados. Para isso, uma rotina específica de backup, ou seja, um procedimento que resguarde uma cópia dos dados como segurança foi vinculado remotamente a uma conta do serviço de armazenamento em nuvem *Dropbox*.

A realização dessa cópia de segurança foi programada para automaticamente salvar as alterações no banco de dados desenvolvido a cada 30 minutos de operação. A escolha por um armazenamento remoto dessas informações reforça a segurança para todo o processo implementado até o momento. Dessa forma, em casos extremos, onde nenhum dado físico ou computacional seja acessível, o presente sistema de dados pode ser resgatado rapidamente da nuvem e posto em operação em apenas alguns minutos.

O grande diferencial deste sistema é que essa ferramenta permite a extração de informações relacionadas e hierarquicamente armazenadas diretamente do sistema desenvolvido tanto local, diretamente no servidor de dados, quanto remotamente através de sistemas estatísticos consolidados, tais como o R, para que esse, em conjunto com seus pacotes acessórios, possa analisar os dados de uma maneira mais robusta e eficiente. Desta forma, os questionamentos do pesquisador se tornam o único fator

limitante para o processamento desses dados.

Com isso, entrega-se uma plataforma robusta, única e desenvolvida exclusivamente com esse propósito de armazenar e processar as informações coletadas referente às capturas de elasmobrânquios em cruzeiros científicos realizados pela FURG.

#### **4. REFERÊNCIAS BIBLIOGRÁFICAS**

CROCKETT, K.; ROB, P. & CORONEL, C. 2008 Database Systems: Design, Implementation & Management, Cengage Learning EMEA.

DATE, C. J. 2016. Introdução a sistemas de banco de dados, Elsevier Brasil,

HARRINGTON, J.L. 2002. Projetos de banco de dados relacionais, Elsevier Editora.

RUSSEL, C., & STEPHENS, J. 2004. Beginning MySQL Database Design and Optimization: From Novice to Professional, Apress, Apresspod Series.

SCHWARTZ, B., ZAITSEV, P. & TKACHENKO, V. 2008. High Performance MySQL: Optimization, Backups, Replication, and More, O'Reilly Media Inc, 2ed.

SILBERSCHATZ, A., KORTH, .H F; & SUDARSHAN, S. 2016. Sistema de banco de dados, ed.6, Elsevier Brasil.

**Tabela 1:** Inventário dos dados de elasmobrânquios obtidos pelo Laboratório de Elasmobrânquios e Aves Marinhas do Instituto de Oceanografia da FURG.

ID	TIPO	Cruzeiro de pesquisa	Período
EV1	Pasta	3/80	05 - 18/06/1980
EV2	Pasta	3/80	05 - 18/06/1980
EV3	Pasta	4/80	10 - 13/07/1980
EV4	Pasta	6/80 - AREPE	20 - 22/08/1980
EV5	Pasta	7/80	23/09/1980 - 08/10/1980
EV6	Pasta	10/80	20 - 27/11/1980
EV7	Pasta	1/81	18 - 27/02/1981
EV8	Pasta	2/81	22/04/1981 - 12/05/1981
EV9	Pasta	7/81	17 - 25/08/1981
EV10	Pasta	8/81 e 9/81	09 - 13/09/1981 e 22/09/1981 - 02/10/1981
EV11	Pasta	1/82	14 - 23/01/1982
EV12	Pasta	3/82	22/07/1982 - 02/08/1982
EV13	Pasta	4/83	15 - 26/04/1983
EV14	Pasta	9/83 e 10/83	09 - 31/08/1983
EV15	Pasta	13/83	08 - 19/11/1983
EV16	Pasta	BELAP	Janeiro 1984
EV17	Pasta	4/84 e 5/84	14 - 24/06/1984 e 07 - 19/11/1984
EV18	Pasta	2/85 e 3/85	17 - 30/06/1985 e 12/07/1985
EV19	Caixa	86/87 - TALUDE	19/07/1986 - 25/05/1987
EV20	Pasta	1989 - Santa Luzia	18/08/1989 - 03/09/1989
EV21	Pasta	DEZ/89 - CORVINA	09 - 22/12/1989

EV22	Pasta	OUT/91	07 - 11/10/1991
EV23	Pasta	2/92 - DIADEMA	29/10/1992 - 02/11/1992
EV24	Pasta	4/95 - DARWIN INITIATIVE	19 - 22/12/1995
EV25	Pasta	96/97 - MARGUS	28/08/1996 - 15/06/1997
EV26	Caixa	96/97/98 - ARGO	13/11/1996 - 05/04/1998
EV27	Caixa	2001/2002 - REVIZEE	09/08/2001 - 30/04/2002
EV28	Caixa	2005 - SALVAR	13 - 28/02/2005
EV29	Pasta	82; 85 - PRAIA	
EV30	Caixa	87/88; 94/95/97 - FÁBRICA	
EV31	Caixa	72; 78 - SUDEPE	Dados pretéritos da SUDEPE 1972 - 1978
EV32	Pasta	MUSTELUS	
EV33	Caixa	MYLIOBATIS	
EV34	Caixa	ZAPTERYX	
EV35	Caixa	SQUATINA	
EV36	Caixa	MUSTELUS SCHMITTI/ 80-81	
EV37	Caixa	MUSTELUS SCHMITTI/ 82-84	
EV38	Caixa	SQUALUS	
EV39	Caixa	GALEORHINUS	
EV40	Caixa	RAJIDAE	

Tabela 2. Estrutura do banco de dados, com suas respectivas entidades e variáveis, para cadastro de capturas de elasmobrânquios em cruzeiros científicos da FURG

Entidades	Campos	Número de registros no banco de dados
Embarcação	Nome da embarcação	25
	Comprimento do casco	
	Potência do motor	
	Ano do cruzeiro	
	Id do projeto de origem	
Estação de Lance	Data	694
	Hora inicial e final	
	Número do lance de pesca	
Local de Lance	Latitude e Longitude	714
	Profundidade	
	Direção e velocidade do vento	
Ambiental	Temperatura	714
	Salinidade de superfície e de fundo	
	Tipo de fundo	
Modalidade de pesca	Arrasto de fundo, meia-água e superfície	714
	Espinhel de superfície e fundo	
Operação de pesca	Pesca de arrasto: velocidade e tempo de arrasto	714
	Espinhel: tempo de imersão e número de anzóis	
Petrecho de pesca	Comprimento e tamanho do anzol	714
	Comprimento e tamanho da malha	

Grupo	Comprimento e tamanho do saco	
	Abertura horizontal e abertura vertical da rede	
	Classe taxonômica	5.080
	Gênero	
	Espécie	
	Número de indivíduos capturados	
	Captura em quilogramas	
Indivíduos coletados	Comprimento total e furcal	865
	Largura do disco	
	Peso total	
	Peso eviscerado	
	Sexo e estágio de desenvolvimento reprodutivo	
	Número espécime	
	Peso da gônada	
Total	Peso do fígado	
		10.235

## APÊNDICE II

### *Análises estatísticas complementares*

#### 1. INTRODUÇÃO

É muito importante reconhecer a diferença entre análise exploratória e avaliação dos pressupostos dos modelos. Durante a análise exploratória dos dados avalia-se a melhor distribuição de probabilidade, a frequência de zeros (zero -inflação), os outliers e a colinearidade. Enquanto que na verificação dos pressupostos dos modelos verifica-se a normalidade, a independência entre as observações e a homocedasticidade. Desta maneira, a análise exploratória é realizada sobre os dados brutos enquanto que a verificação dos pressupostos é sobre os resíduos dos modelos.

Este apêndice destina-se a apresentar um conjunto de análises exploratórias que foram conduzidas sobre os dados utilizados nesta tese. Uma vez que estes foram diretamente importados da base de dados particular, esta análise exploratória serviu também para que uma verificação sobre a digitação das informações fosse efetuada

numa espécie de consolidação do banco de dados.

A título de exemplo os resultados apresentados referem-se a um sub-conjunto de dados do manuscrito I “Long-term Rajiformes (Chondrichthyes: Elasmobranchii) spatio-temporal dynamics in the southernmost states of Brazil.” e a base de dados do artigo II “Modelagem espacial bayesiana para riqueza de elasmobrânquios do extremo sul do Brasil”.

Todas as análises foram conduzidas no software estatístico R versão 3.4.1 (R Core Team, 2018) utilizando os pacotes Rmysql (Ooms et al. 2019) e dplyr (Wickham et al. 2018).

## **2. ANÁLISE EXPLORATÓRIA DOS DADOS**

O conhecimento do pesquisador sobre seus dados é parte fundamental no processo de elaboração de modelos estatísticos. Portanto, a exploração dos dados consiste em, preliminarmente, entender as suas características por meio de múltiplos recursos técnicos, predominantemente de exploração visual (Idreos et al. 2015).

De acordo com Zuur et al. (2010) esta etapa do estudo pode representar 50% do tempo gasto em uma análise. Contudo uma vez que se dedique a isto o analista diminui o risco de propor modelos pouco realistas e, desta forma, eximir-se de obter estimativas nos parâmetros dos modelos erroneamente ajustados.

### **2.1 Distribuição de probabilidades**

As distribuições de frequências para uma variável aleatória (v.a.) podem apresentar diferentes padrões de distribuição. Muito embora a distribuição normal seja a mais conhecida, outras tantas distribuições existem para melhor descrever a variável de interesse. Entende-se por v.a. como sendo uma variável que possui um único valor para cada resultado de um experimento, e a distribuição de probabilidades nada mais é que o conjunto de possibilidades que uma v.a. pode apresentar (Kinas e Andrade, 2010).

São reconhecidos dois grupos de distribuições de probabilidade, um para variáveis discretas e outro para variáveis contínuas. As variáveis quantitativas discretas apresentam valores finitos e enumeráveis, enquanto que as variáveis quantitativas contínuas assumem valores em intervalos dos números reais. Através de funções

matemáticas é possível adequar uma variável aleatória de interesse a um tipo de distribuição de probabilidade específica.

## **2.2 Dados zero – inflacionados**

Quando uma v.a. é zero-inflacionada, a frequência de zeros na amostra é superior aos valores diferentes de zero. A avaliação é feita diretamente por comparação da frequência absoluta dos valores. Considerar essa característica da v.a. é imprescindível para que os parâmetros estimados pelos modelos não sejam obtidos a partir de estimativas errôneas. Para o caso de amostras zero-inflacionadas existem duas soluções, ou aplicamos modelos com suas versões zero-inflacionadas ou podemos ajustar modelos truncados em zero.

## **2.3 Outliers**

São os valores extremamente raros, no sentido de que estão muito afastados da maioria dos dados mas que podem revelar informações importantes (Triola, 1999). Dentre uma das formas de identificar um *outlier* está o gráfico Cleveland dotplot (Cleveland 1993). Neste gráfico os pontos que se destacam dos valores observados ao lado direito ou ao lado esquerdo são consideravelmente maiores ou menores que a maioria das observações e requerem investigação adicional (Zuur et al. 2010). Desconsiderar a verificação de *outliers* pode enviesar negativamente o resultado de uma análise, as quais tendem a ser influenciadas por valores extremos.

## **2.4 Colinearidade**

A existência de forte correlação entre covariáveis pode ser identificada através de gráficos de dispersão pareados e seus respectivos coeficientes de correlação de Pearson. Estes gráficos também são úteis em verificar se as relações entre as variáveis são lineares ou não-lineares.

Para solucionar problemas de colinearidade basta desconsiderar uma das covariáveis, preferencialmente a que tiver menor sentido ecológico no estudo. Ignorar a colinearidade pode resultar em uma análise estatística confusa em que nada seja significativo no momento do ajuste dos modelos (Zuur et al. 2010).



### *Estudo de caso I*

A variável aleatória é a CPUE (kg/área varrida), cujos valores assumidos são inteiros e contínuos, e segue uma distribuição Gama (Fig. 1a). Para esta amostra pode-se observar os valores máximos e mínimos de CPUE e analisar que a frequência de zeros não foi superior aos demais valores da amostra (Fig. 1b)

A presença de *outliers* foi verificada para a variável CPUE e covariáveis batimetria, temperatura e salinidade de superfície e fundo, sendo que nenhum caso de *outlier* foi identificado. Nos gráficos os valores nulos representam ausência de informações (Fig. 1c). Para as covariáveis latitude e longitude também analisou a presença de *outliers* mas nenhum valor extremo foi verificado, o que indica o correto registro das informações de local de captura.

A colinearidade foi identificada para as variáveis temperatura e salinidade de superfície ( $r = 0,85$ ); temperatura de fundo e salinidade de fundo ( $r = 0,96$ ); profundidade e temperatura de fundo ( $r = -0,94$ ) e profundidade e salinidade de fundo ( $r = -0,91$ ). Os valores positivos e negativos indicam o modo de correlação (Fig. 1d).

### *Estudo de caso II*

A variável aleatória é a riqueza de espécies, que apresenta distribuição de probabilidade discreta, uma vez que se origina do número de ocorrências de um evento (neste caso a captura de uma espécie) em uma determinada área. Entre as opções de distribuição discreta a de Poisson se adequa melhor aos dados pois parte do pressuposto de que a média dos valores seja igual à variância (Fig. 2a).

Os valores de riqueza iguais a zero denotam captura nula e são inexistentes neste gráfico dada a maneira como foram conduzidas as análises para esta variável. Os arrastos com captura de apenas uma espécie foram os mais frequentes que os demais (Fig. 2b).

A análise de *outliers* para a variável riqueza identificou valores distantes da média das observações, porém constatou-se que são valores reais. Para as covariáveis abióticas (profundidade, temperatura e salinidade de superfície e fundo) e bióticas (biomassa de teleósteos e de elasmobrânquios), nenhuma discordância dos valores médios foi observada (Fig. 2c).

Os *outliers* também foram analisados para as variáveis latitude e longitude para verificar possíveis erros de registro de informação georeferenciada, mas todos os

valores encontrados foram como o esperado.

Foram consideradas colineares as variáveis temperatura de superfície e de fundo ( $r = 0,83$ ), salinidade de superfície e de fundo ( $r = 0,98$ ), ambas com correlação positiva. (Fig. 2d).

### 3. AVALIAÇÃO DOS PRESSUPOSTOS

A qualidade do ajuste dos modelos é avaliada por meio de gráficos diagnósticos que verificam os pressupostos de normalidade e homocedasticidade nos resíduos. Uma vez que estas análises estão contempladas nos resultados dos artigos desta tese, apenas uma breve descrição sobre este assunto será apresentada.

#### 3.1 Normalidade

Considerando que todo teste de normalidade supõe que os dados têm média zero e variância constante, apenas os resíduos (em relação à média) atendem essa premissa, uma vez que os dados apresentam médias dependentes do efeito das variáveis explicativas. A normalidade nos resíduos pode ser avaliada pelo gráfico do Quantil-Quantil, onde valores observados da variável são relacionados aos valores esperados. Se há uma boa aderência dos dados à distribuição Normal os pontos estarão próximos a reta de referência apresentada no gráfico. A presença dos *outliers* na amostra bem como a natureza dos dados (p. ex. dados de contagem) podem resultar em fuga da normalidade dos resíduos.

#### 3.2 Homocedasticidade

A homocedasticidade é o termo para designar variância constante dos erros experimentais para observações distintas. Para verificar a variância dos resíduos pode-se relacionar os valores dos resíduos versus valores ajustados (valores preditos). Se no gráfico de dispersão os pontos estão aleatoriamente distribuídos em torno do 0, sem nenhum comportamento ou tendência, temos indícios de que a variância dos resíduos é homoscedástica.

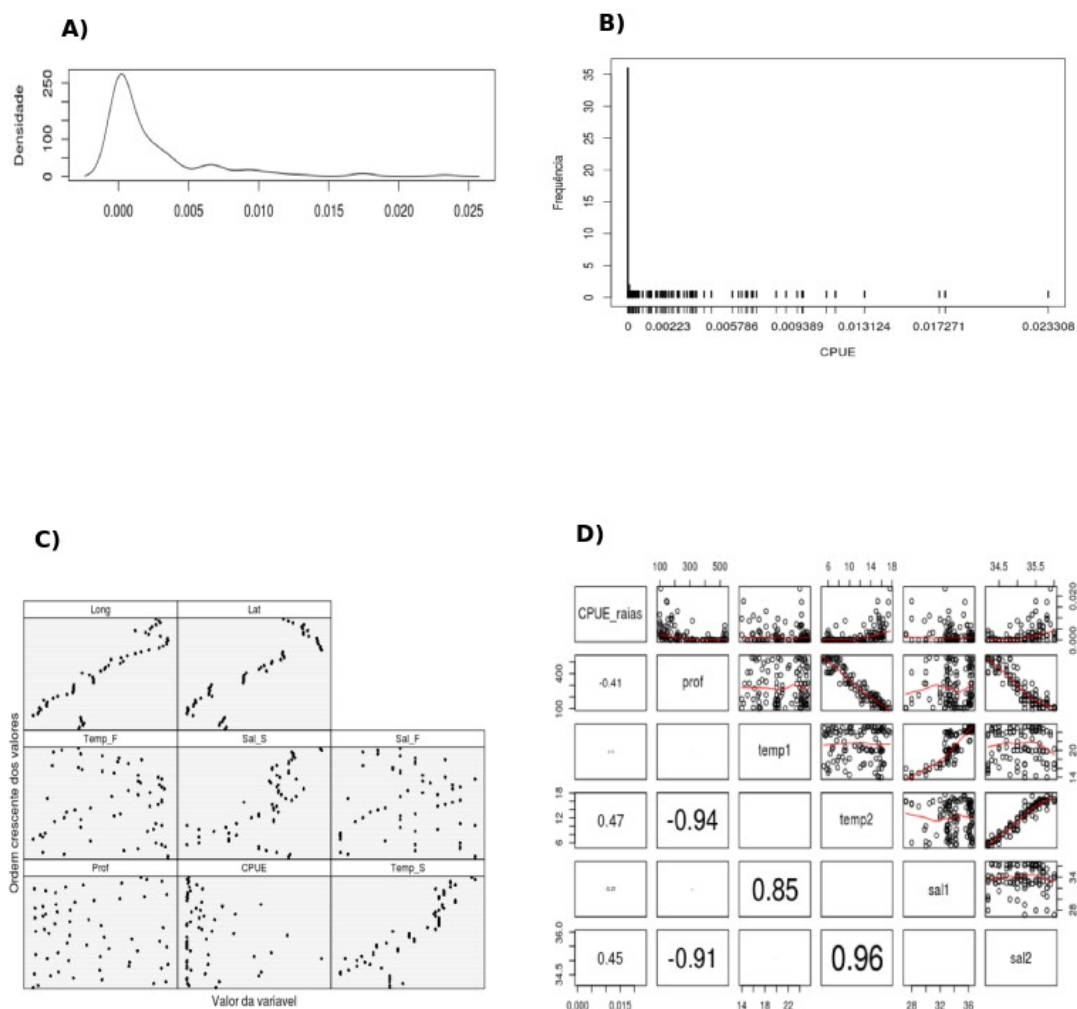


Figura 1. Análise exploratória: a) Distribuição de probabilidade da variável resposta

CPUE; b) Frequência de valores zeros na amostra de CPUE; c) Análise de *outliers* para CPUE e demais covariáveis; d) Análise de colinearidade entre variáveis aleatórias.

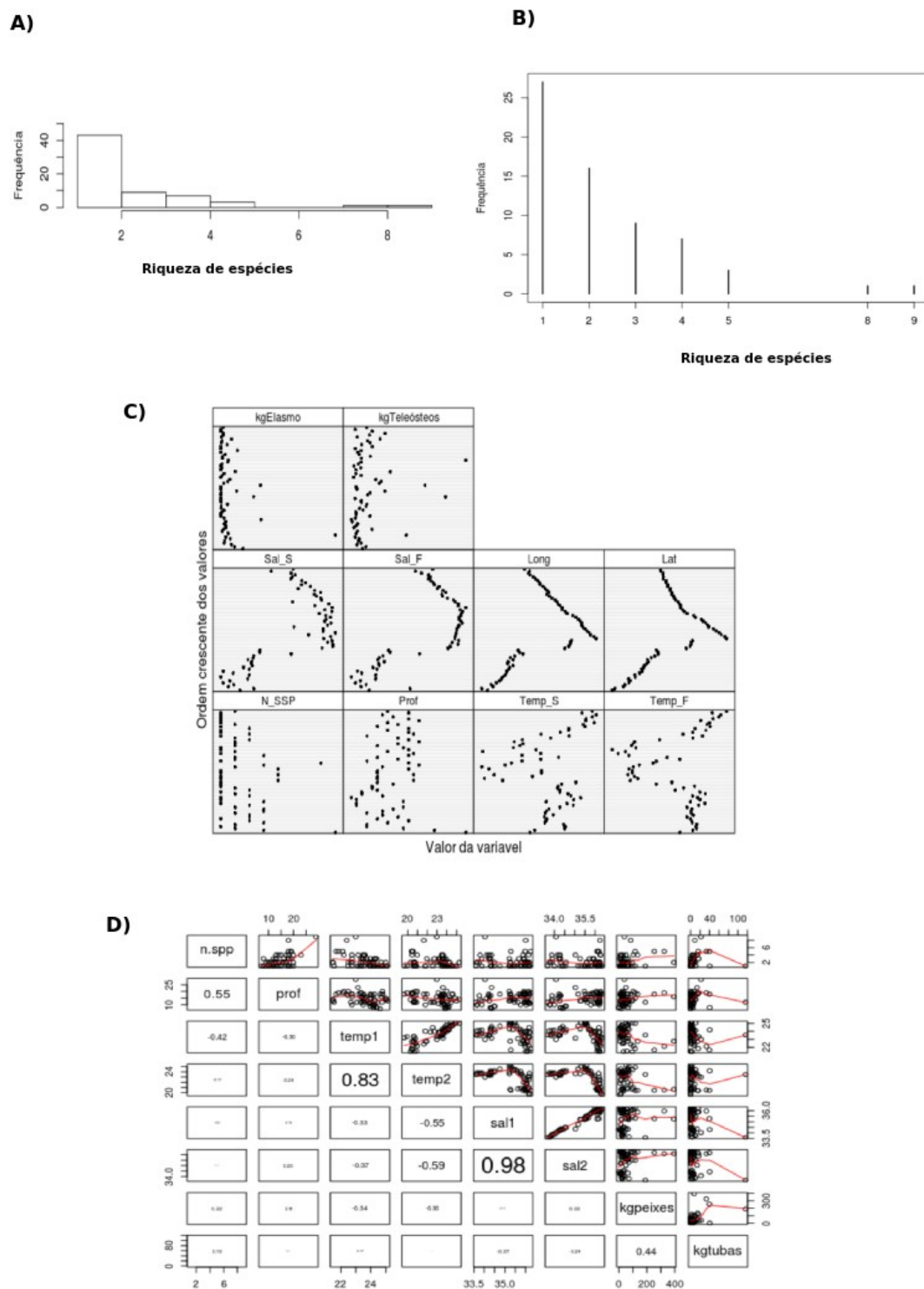


Figura 2. Análise exploratória: a) Histograma da distribuição de probabilidade para a variável resposta riqueza; b) Frequência de valores zeros na amostra; c) Análise de *outliers* para variável resposta e demais covariáveis; d) Análise de colinearidade entre variáveis aleatórias.

#### 4. REFERÊNCIAS BIBLIOGRÁFICAS

IDREOS, S.; PAPAEMONOUIL, O. & CHAUDHURI, S. 2015. "Overview of data exploration techniques", *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data SIGMOD '15*, pp. 277-281

KINAS, P. G. & ANDRADE, H. A. 2010. Introdução à Análise Bayesiana (com R). Porto Alegre, maisQnada. 240p.

OOMS, J., JAMES, D., DEBROY, S., WICKHAM, H. & HORNER, J. (2019). Database Interface and 'MySQL' Driver for R. R package Version 0.10.16

R Development Core Team, 2018. R: A language and environment for statistical computing.

Vienna, Austria: R Foundation for Statistical Computing.

RAJU, P. S. N.; RAJU, J. V . N & SWAMI, S. V. 2017. A Novell Approach for Comparative Study of Classification Algorithms Using R-Programming. International Journal of Advanced Engineering and Global Technology, ISSN No: 2309 – 4893

TRIOLA, M. F. 2005. Introdução a estatística. Rio de Janeiro. LTC \_ Livros técnicos e Científicos Editora S.A. 410p.

ZUUR, A. F., IENO, E. N. & ELPHICK, C. S. 2010. A protocol for data exploration to avoid common statistical problems. *Meth. Ecol. Evol.*, 1: 3-14.

WICKHAM, H., FRANÇOIS, R., HENRY, L., MÜLLER, K.(2018). A Grammar of Data Manipulation . R package Version 0.7.8

### APÊNDICE III

“Long-term Rajiformes (Chondrichthyes: Elasmobranchii) spatio-temporal dynamics in the southernmost states of Brazil”

**IVANOFF<sup>1,2</sup>, R.; PENNINO, M.G.<sup>3,4,5</sup>; RUFENER, M. C.<sup>4,6</sup>; VOOREN, C.M.<sup>1,2</sup> & KINAS, P. G<sup>2</sup>**

<sup>1</sup>Programa de Pós-graduação Oceanografia Biológica, Instituto de Oceanografia, Universidade Federal do Rio Grande – FURG, Avenida Itália km 8, Carreiros, CEP - 96201-900, Rio Grande, RS, Brasil, raydnunes@gmail.com.

<sup>2</sup>Laboratório de Estatística Ambiental - Instituto de Matemática, Estatística e Física, FURG.

<sup>3</sup>Fishing Ecology Management and Economics (FEME), Universidade Federal do Rio Grande do Norte – UFRN. Depto. de Ecologia, Natal (RN), Brazil.

<sup>4</sup>Statistical Modeling Ecology Group (SMEG). Departament d’Estadística i Investigació Operativa, Universitat de València, C/Dr. Moliner 50, Burjassot, 46100, Valencia, Spain.

<sup>5</sup>Instituto Español de Oceanografía, Centro Oceanográfico de Murcia, C/ Varadero 1, San Pedro del Pinatar, 30740, Murcia, Spain.

<sup>5</sup>Technical University of Denmark, Institute for Aquatic Resources, Kemitortet, Building 201, 2800, Kgs. Lyngby, Denmark.

**Submetido para:**

Journal of Marine Systems

## APÊNDICE IV

“Modelagem espacial bayesiana para riqueza de elasmobrânquios do extremo sul do Brasil”

**IVANOFF<sup>1,2</sup>, R.; PENNINO, M.G<sup>3,4,5</sup>; RUFENER, M. C.<sup>4,6</sup>; VOOREN, C.M.<sup>1,2</sup> &  
KINAS, P. G<sup>2</sup>**

<sup>1</sup>Programa de Pós-graduação Oceanografia Biológica, Instituto de Oceanografia, Universidade Federal do Rio Grande – FURG, Avenida Itália km 8, Carreiros, CEP - 96201-900, Rio Grande, RS, Brasil, raydnunes@gmail.com.

<sup>2</sup>Laboratório de Estatística Ambiental - Instituto de Matemática, Estatística e Física, FURG.

<sup>3</sup>Fishing Ecology Management and Economics (FEME), Universidade Federal do Rio Grande do Norte – UFRN. Depto. de Ecologia, Natal (RN), Brazil.

<sup>4</sup>Statistical Modeling Ecology Group (SMEG). Departament d’Estadística i Investigació Operativa, Universitat de València, C/Dr. Moliner 50, Burjassot, 46100, Valencia, Spain.

<sup>5</sup>Instituto Español de Oceanografía, Centro Oceanográfico de Murcia, C/ Varadero 1, San Pedro del Pinatar, 30740, Murcia, Spain.

<sup>6</sup>Technical University of Denmark, Institute for Aquatic Resources, Kemitortet, Building 201, 2800, Kgs. Lyngby, Denmark.

**Publicado em:**

Revista CEPsul - Biodiversidade e Conservação Marinha



## APÊNDICE V

### *Códigos no R*

#### **1. Banco de dados**

```
dbGetQuery(con, "Show Tables")

# faz uma busca pelo id da embarcação de acordo com o que está escrito na observação
do cruzeiro

as.numeric(dbGetQuery(con, "SELECT id FROM embarcacao WHERE obs='*'))

# faz uma busca pelo id do lance de pesca inserindo o número do lance e o id da
embarcação

as.numeric(dbGetQuery(con, "SELECT id from estacao_lance WHERE nr_lance= *
AND id_emb='*'))

# seleciona as informações ambientais do id do lance de pesca

dbGetQuery(con, "SELECT * from ambiental WHERE id_est='*')

# seleciona pelo id da embarcação as informações georeferenciadas

id_elance <- as.list(dbGetQuery(con, "SELECT id from estacao_lance WHERE
id_emb='*'))

for (i in 1: length(id_elance$id)) {

  print(dbGetQuery(con, paste('SELECT id_emb, lat_dec_i, lon_dec_i FROM
local_lance WHERE id_est = ', id_elance$id[i])))}
```

## 2. Análises exploratórias

```
# análise da massa de probabilidade
```

```
par(mfrow = c(2, 3), mar = c(5, 4, 3, 2))
```

```
plot(density(dados$resposta), ylab = "Frequência")
```

```
hist(dados$resposta, xlab = "", ylab = "Frequência")
```

```
# teste de normalidade
```

```
shapiro.test(dados$resposta)
```

```
# frequência de zeros
```

```
par(mfrow = c(1, 1), mar = c(5, 4, 3, 2))
```

```
plot(table(dados$resposta), xlab = "", ylab = "Frequência")
```

```
table(dados$resposta)
```

```
# outliers variáveis explicativas
```

```
W <- cbind(dados$resposta, dados$v1, dados$v2)
```

```
colnames(W) <- c("resposta", "v1", "v2")
```

```
dotplot (as.matrix(W), groups = FALSE, strip = strip.custom(bg = 'white', par.strip.text  
= list(cex = 0.8)), scales = list(x = list(relation = "free"), y = list(relation = "free"), draw  
= FALSE), col = 1, cex = 0.5, pch = 16, xlab = "Valor da variavel", ylab = "Ordem  
crescente dos valores")
```

```

# Colinearidade

pairs(resposta ~ v1 + v2 , pch = 19)

panel.cor <- function(x, y, digits=2, prefix="", cex.cor)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(0, 1, 0, 1))
  r = (cor(x, y))
  txt <- format(c(r, 0.123456789), digits=digits)[1]
  txt <- paste(prefix, txt, sep="")
  if(missing(cex.cor)) cex <- 0.8/strwidth(txt)
  text(0.5, 0.5, txt, cex = cex * abs(r))}

pairs(resposta ~ v1 + v2 , pch = 19) , upper.panel=panel.smooth, lower.panel=panel.cor)

cor(resposta)

```

### 3. Modelos hierárquicos Bayesianos com INLA

```

resposta ~ v1 + efeito espacial
Distribuição gamma

# Para criar a matriz de coordenadas**
str(dados)
head(coords<-as.matrix(dados[,3:2]))
coords<-as.matrix(dados[,3:2])

# Importando o shapefile da costa
RS<- readShapePoly("RS")

```

```
plot(RS)
```

```
# Fazendo o plot do mesh
```

```
bound.pts <- inla.nonconvex.hull(coords, 0.3, 0.3)
```

```
mesh <- inla.mesh.2d(boundary=bound.pts, max.edge=c(0.20, 0.40), offset=c(0.2, 0.5),  
cutoff=0.05)
```

```
plot(mesh, asp=1) ; points(coords,col=2,pch=20) ; plot(RS,col="grey50",add=T)
```

```
# Identificando as coordenadas dos vértices dos triângulos da malha (definição dos  
pontos de predição)**
```

```
mesh.points<- SpatialPoints(  
  coords = cbind(mesh$loc[,1], mesh$loc[,2]),  
  proj4string = CRS("+proj=longlat +datum=WGS84 +no_defs"))  
points(mesh.points)
```

```
##### ESTIMAÇÃO #####
```

```
# Criando a matriz de estimação**
```

```
A.est<- inla.spde.make.A(mesh, loc=coords)
```

```
dim(A.est)
```

```
# Modelo SPDE com a função de correlação de Matérn**
```

```
spde <- inla.spde2.matern(mesh, alpha=2)
```

```
#Padronizando (escla Z) os dados coletados in situ
```

```
vp<-deconstand(dados$v1,"standardize", na.rm=T )
```

```
#Gerando o data frame de estimação
```

```
N<-NROW(dados$resposta)
```

```
E<-data.frame(Intercept=rep(1,N), v=vp)
```

```
# Stack - estimação**
```

```
stk.est<-inla.stack(data=list(y=dados$resposta), A = list(A.est,1), effects =
list(spatial=1:spde$n.spde, E), tag = "est")
```

```
# Combinação de variáveis
```

```
FI<- y ~ 0 + Intercept + f(spatial, model=spde)
```

```
# Modelo com distribuição Gamma
```

```
modI<- inla(FI, family = "Gamma", data = inla.stack.data(stk.est),control.predictor=list
(A= inla.stack.A(stk.est),compute=TRUE),verbose=TRUE,
control.compute=list(return.marginals=TRUE, waic=T,dic=TRUE, cpo =TRUE))
```

```
# Resultado do modelo
```

```
summary(modI)
```

```
plot(modI)
```

```
# Análise preditiva do modelo
```

```
getCPO <- function(res) { if(is.list(res$cpo)) return(res$cpo$cpo)
else return(`res`-$cpo)}
```

```
lcpo<--mean(log(getCPO(modI)[which(getCPO(modI)>0)]))
```

```
lcpo
```

```
##### PREDIÇÃO #####
```

```
ext<-extent(-54.0, -47.00, -35.0, -28.00)
```

```
#Extrair os valores dos preditores (raster) em cada ponto do mesh (mesh.points)
```

```
raster=extract(mean, mesh.points)
```

```
#Extrair os valores dos preditores (raster) em cada ponto de coleta (coords)
```

```
e_raster<-extract(mean, coords)
```

```
#Padronizar os valores do raster para "DATA FRAME PREDIÇÃO"
```

```
s_raster<-decostand(s_raster, "standardize", na.rm=T)
```

```
summary(s_raster)
```

```

# Criando a matriz de predição**
A.pred = inla.spde.make.A(mesh,loc=matrix(c(mesh$loc[,1],mesh$loc[,2]),ncol=2))
head(mesh$loc)
dim(A.pred)

# Gerando o data frame da predição (informações de raster)**

P<-data.frame(Intercept=1,
              v.a=s_raster)

##Stack - predição**
stk.pred<-inla.stack(data=list(y=rep(NA,spde$n.spde)),
A=list(A.pred,1),effects=list(spatial=1:spde$n.spde, P), tag="pred")

## Stack estimação + predição**
stk=inla.stack(stk.est, stk.pred)

FI<- y ~ 0 + Intercept + f(spatial, model=spde) # MODELO COM MENOR DIC

# Modelo com distribuição Gamma
modI<- inla(FI,
            family = "Gamma",
            data = inla.stack.data(stk),
            control.predictor=list (A= inla.stack.A(stk), compute=TRUE),
            control.compute=list(return.marginals=TRUE, waic=T,dic=TRUE, cpo
=TRUE))

# Projeção do efeito espacial/preditor linear em um grid
**atenção ao nome do modelo!**

pgrid<-inla.mesh.projector(mesh,dims=c(600,450))

```

```

prd.m<-inla.mesh.project(pgrid,modI$summary.ran$spat$mean) # média do GMRF
prd.s<-inla.mesh.project(pgrid,modI$summary.ran$spat$s) # desvio padrão do GMRF

```

```

# Extrair o índice de predição**

```

```

igr <- inla.stack.index(stk, "pred")$data # média do preditor linear
ms.m<-inla.mesh.project(pgrid,(exp(modI$summary.fitted.values[igr,1])))
sd.s<-inla.mesh.project(pgrid,(modI$summary.fitted.values[igr,2]))

```

```

# Projetar média/desvio do GMRF/preditor linear**

```

```

grid.arrange(levelplot(prd.m, col.regions=topo.colors(99),
                        main="latent field mean", xlab=" ", ylab=" ",
                        scales=list(draw=F)),
              levelplot(prd.s,col.regions=topo.colors(99),main="latent field SD",
                        xlab=" ", ylab=" ", scales=list(draw=F)),
              levelplot(ms.m,
                        xlab="", ylab="",main="response mean",
                        col.regions=tim.colors(99),scales=list(draw=F)),
              levelplot(sd.s,
                        xlab=" ", ylab=" ", main="response SD",
                        col.regions=topo.colors(99),scales=list(draw=F)),
              nrow=2)

```

```

# Distribuição posterior do campo markoviano nos locais específicos**

```

```

inddp<-inla.stack.index(stk, tag="pred")$data
str(inddp)

```

```

# Extract estimation index

```

```

(x<-inla.stack.index(stk, tag='est')$data) #

```

```

# Extract model's residuals

```

```

residuals<- dados$respostamodI$summary.linear.predictor[x,1]

```

```

# valores observados vs preditos
plot(dados$resposta~modI$summary.fitted.values[x'mean'],
      ylab="Observados",xlab="Preditos") ; abline(0,1)
cor.test((dados$resposta,modI$summary.fitted.values[x,'mean']))

# resíduos vs preditos
plot(residuals~modI$summary.fitted.valuesx,'mean'],
      ylab="resíduos", xlab="preditos")
cor(residuals, modI$summary.fitted.values[x,'mean'])

# quantil vs quantil
qqnorm(residuals,main=NULL,
        ylab = "quantil observado", xlab = "quantil teórico")
qqline(residuals,col="grey30")

# histograma dos resíduos
hist(residuals)

# Sumário da Distribuição posterior do campo markoviano**
s<-round(modI$summary.linear.predictor[inddp,],4)

#Extrair os parâmetros do GMRF**
result=inla.spde2.result(modI, "spatial", spde)

#Valor médio do kappa**
inla.emarginal(function(x)x,result$marginals.kappa[[1]])

# Valor médio do tau**
inla.emarginal(function(x)x,result$marginals.variance.nominal[[1]])

# rho(km) e intervalo de credibilidade a 95% (distancia em que rho = 0,1)**
(rangecor<-c(mean=inla.emarginal(function(x) x, result$marginals.range[[1]]),

```



```

      q=inla.hpdmarginal(0.95, result$marginals.range[[1]])(c(1,2,3))
# média e desvio padrão do campo latente**
inla.emarginal(function(x) x, result$marginals.range.nominal[[1]])

range.nom.quantiles = inla.qmarginal(c(0.025, 0.5, 0.975), range.nom.marg)
range.nom.quantiles
# 0.1422716 0.4442432 1.8469577

# Extrair a distribuição marginal e HPD a 95% da variável resposta**
**atenção ao nome do modelo!**

ind2<-inla.stack.index(stk,"est")$data
s2<-round(modI$summary.fitted.values[ind2,])
mr<-modI$marginals.fitted.values[ind2]

inla.hpdmarginal(0.95,modI$marginals.fixed$salsup)

```