Universidade Federal do Rio Grande - FURG

Instituto de Oceanografia

Programa de Pós-Graduação em Oceanografia Biológica

MODELAGEM ESPACIAL E TEMPORAL DOS PADRÕES DE DENSIDADE E RIQUEZA DE ESPÉCIES DE CETÁCEOS NA BACIA DE SANTOS

ANA CAROLINA BRAGA MARTINS

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 DOUTORA

Orientador: Prof. Dr. Luciano Dalla Rosa

Coorientador: Prof. Dr. Paul Gerhard Kinas

RIO GRANDE/RS, Brasil

Fevereiro/2026

M386m

Martins, Ana Carolina Braga
Modelagem espacial e temporal dos padrões de densidade e riqueza de espécies de cetáceos na Bacia de Santos / Ana Carolina Braga Martins. – 2026.
173f.

Tese (Doutorado) – Universidade Federal do Rio Grande – Programa de Pós-Graduação em Oceanografia Biológica, 2026.

Orientador: Dr. Luciano Dalla Rosa
Co-orientador: Dr. Paul Gerhard Kinas

1. Oceanografia Biológica. 2. Abundância. 3. Atlântico Sudoeste. 4. Comunidade de cetáceos. 5. Modelos de distribuição de espécies. 6. Mudanças climáticas. 7. Preditores ambientais. I.Rosa, Luciano Dalla. II.Kinas, Paul Gerhard. III.Título.

CDU 573(26)

Catalogação na fonte: Bibliotecária Brenda Sequeira CRB10/1979

UNIVERSIDADE FEDERAL DO RIO GRANDE - FURG
IO – INSTITUTO DE OCEANOGRAFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM OCEANOGRAFIA BIOLÓGICA
E-mail: ccpgob@furg.br – home-page: www.ppgocbio.furg.br

ATA DE DEFESA DE TESE DE DOUTORADO – 01/2026

Às quatorze horas do dia nove de Fevereiro do ano de dois mil e vinte e seis, Virtual Sala: https://conferenciaweb.rnp.br/sala/luciano-dalla-rosa-2/, Presencial: Sala de Reunião Prédio Recursos Renováveis – IO/FURG-C. Carreiros, reuniu-se a Comissão Examinadora de Tese de Doutorado da Acad. **Ana Carolina Martins**, composta pelos seguintes membros: Prof. Dr. Luciano Dalla Rosa – Orientador - (PPGOB – IO/FURG); Prof. Dr. Paul Gerhard Kinas – Coorientador - (IMEF/FURG); Prof. Dr. Leandro Bugoni - (PPGOB –ICB/FURG); Prof. Dra. Silvina Botta - (PPGOB –IO/FURG); Prof. Dr. Mauricio Camargo - (IO/FURG); Prof. Dr. Alexandre Azevedo (UERJ) e Prof. Dr. Rodrigo Sant'Ana (UNIVALI). Título da Tese: "**MODELAGEM ESPACIAL E TEMPORAL DOS PADRÕES DE DENSIDADE E RIQUEZA DE ESPÉCIES DE CETÁCEOS NA BACIA DE SANTOS**". Dando início à reunião, a Coordenadora do PPGOB, agradeceu a presença de todos e fez a apresentação da Comissão Examinadora. Logo após, explicou que a candidata teria um tempo de 45 a 60 min. para explanação do tema, e cada membro da Comissão, um máximo de 30 min. para perguntas. Dando prosseguimento, passou à palavra a candidata que apresentou o tema e respondeu às perguntas formuladas. Após explanação, a Comissão aprovou o que segue: As sugestões de todos os membros da Comissão Examinadora foram aceitas pela orientador/candidata e incorporadas na versão final. Foi atribuída a seguinte classificação a candidata: Prof. Dr. Leandro Bugoni – Classificação: Aprovada; Profa. Dra. Silvina Botta – Classificação: Aprovada; Prof. Dr. Mauricio Camargo – Classificação: Aprovada; Prof. Dr. Alexandre Azevedo – Classificação: Aprovada e Prof. Dr. Rodrigo Sant'Ana – Classificação: Aprovada. A candidata foi considerada APROVADA por UNANIMIDADE. Nada mais havendo a tratar, eu, Clabisnei Moura de Melo – Secretário do PPGOB, lavro a presente ata, que após lida e aprovada, será assinada pela Comissão Examinadora, pela aluna e pela Coordenadora do PPGOB.

gov.br Documento assinado digitalmente
LUCIANO DALLA ROSA
Data: 10/02/2026 12:08:51-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Luciano Dalla Rosa

gov.br Documento assinado digitalmente
PAUL GERHARD KINAS
Data: 21/02/2026 12:00:08-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Paul Gerhard Kinas

gov.br Documento assinado digitalmente
LEANDRO BUGONI
Data: 20/02/2026 14:58:00-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Leandro Bugoni

gov.br Documento assinado digitalmente
SILVINA BOTTA
Data: 30/03/2026 16:16:03-0300
Verifique em https://validar.iti.gov.br

Profa. Dra. Silvina Botta

gov.br Documento assinado digitalmente
MAURICIO GARCIA DE CAMARGO
Data: 20/02/2026 09:08:51-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Mauricio Camargo

gov.br Documento assinado digitalmente
ALEXANDRE DE FREITAS AZEVEDO
Data: 10/02/2026 13:15:26-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Alexandre Azevedo

gov.br Documento assinado digitalmente
RODRIGO SANT ANA
Data: 19/02/2026 07:35:20-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Rodrigo Sant'Ana

gov.br Documento assinado digitalmente
ANA CAROLINA BRAGA MARTINS
Data: 27/04/2026 10:24:09-0300
Verifique em https://validar.iti.gov.br

Acad. Ana Carolina Martins

Prof. Dra. Silvina Botta
Coordenadora PPGOB

gov.br Documento assinado digitalmente
SILVINA BOTTA
Data: 30/03/2026 16:13:49-0300
Verifique em https://validar.iti.gov.br

EPÍGRAFE

"*The most probable state of nature is always the one updated by the evidence*".

– Edwin Thompson Jaynes

"*Fish, amphibian, and reptile, warm-blooded bird and mammal—each of us carries in our veins a salty stream in which the elements sodium, potassium, and calcium are combined in almost the same proportions as in sea water.*"

– Rachel Carson

AGRADECIMENTOS

Meus agradecimentos à FURG e ao PPGOB pelas oportunidades concedidas e por toda a assistência durante o período. Ao estado brasileiro, por meio da CAPES, pelo financiamento da bolsa de doutorado. Durante essa jornada pude conhecer pessoas inspiradoras, adquiri experiências valiosas e pude crescer como cientista e como pessoa.

Aos meus orientadores, Luciano e Kinas, minha sincera gratidão por todo o apoio, paciência, confiança, e conhecimentos compartilhados. Kinas, meu grande amigo e incentivador, minha inspiração científica, obrigada por ter idealizado comigo esse projeto lá no início e por ter me apoiado até o final. Luciano, gratidão por todo o apoio, pela amizade, e pela ótima convivência durante todo esse período. Obrigada também por ter me inserido nesse universo GIS e fortalecido os meus laços com a modelagem.

Meu muito obrigada à minha família, desde sempre para sempre. Mariana, meu maior amor, que nasceu junto com a Tese e me ensinou que grandes responsabilidades podem trazer grandes progressos; também me lembrando sempre da importância de rir e brincar. Obrigada por entender ausências e principalmente tantas presenças em casa dedicada no computador. Te amo filha! Débora, minha madrasta, pela amizade e por ter me estendido a mão nos momentos essenciais. Meus pais, Fábio e Tatiane, pelo apoio incondicional, pelo incentivo ao estudo, e por serem meu maior exemplo de integridade. Meu padrasto João Carlos, meu irmão Rodrigo, minha tia Thaís, gratidão pela amizade e suporte fundamental de vocês, pelo carinho com a minha pequena. Ao meu companheiro Gringo por trazer tranquilidade e apoio durante a fase final deste trabalho. Minhas avós, Maria Inês e Maria da Glória, pela força e incentivo em todos os momentos da minha caminhada.

Aos amigos do LRPAME, Maria, Andrea, Deisi, Sheron, Alberto, Dayane; e à família LEA (ou NAAE, como era no ‘nosso tempo’), Flávia, Raquel, Juliano, Gustavo, Liana, Laura, Hans, Carlos, Bruno: agradeço por todos os cafés, conselhos, discussões de scripts, brincadeiras, filosofias e pelo nosso constante companheirismo.

À Vera, pela amizade, por sempre responder a todas as minhas dúvidas sobre o curso (com os devidos puxões de orelha), por acreditar em mim, e pelos conselhos únicos e originais, os quais guardo anotados ainda.

Ao Leo, pela confiança e pela excelente parceria.

Aos professores Leandro e Maurício, pelo acompanhamento, dicas e questionamentos importantes ao longo de todo o percurso.

Aos membros da banca Alexandre, Rodrigo e Silvina, pelos comentários e reflexões que contribuíram para o aprimoramento da versão final deste trabalho.

À Petrobras, pela autorização e cedência dos dados para a realização deste estudo.

Ao futuro e ao que ele reserva.

ÍNDICE

RESUMO.....6

ABSTRACT.....7

PREFÁCIO.....9

CAPÍTULO I.....10

1. INTRODUÇÃO.....11

1.1. HIPÓTESES.....17

1.2. OBJETIVOS.....18

REFERÊNCIAS BIBLIOGRÁFICAS.....18

CAPÍTULO II.....24

Spatial and seasonal patterns of cetacean species richness: a Bayesian approach.….25

CAPÍTULO III.....63

Habitat-based BART models for cetaceans in the western South Atlantic: current and future distribution under climate change scenarios.....64

CAPÍTULO IV.....114

Cetacean abundance in the western South Atlantic: estimates from multi-species hierarchical distance sampling.....115

CAPÍTULO V.....145

Conclusões gerais e perspectivas futuras.....146

APÊNDICE.....151

Códigos utilizados no R.....151

RESUMO

Esta tese investigou, de forma integrada, os padrões espaciais e temporais da distribuição, riqueza de espécies e abundância de cetáceos na Bacia de Santos, sudoeste do Oceano Atlântico, com foco nos efeitos de variáveis ambientais e nas respostas potenciais às mudanças climáticas. Para isso, foram analisados dados obtidos a partir de levantamentos sistemáticos embarcados realizados entre 2015 e 2019, abrangendo diferentes estações do ano e mais de 40.000 km de esforço amostral. Utilizando uma abordagem Bayesiana baseada em modelos espaciais lineares generalizados, foi demonstrado que a riqueza de espécies de cetáceos é fortemente estruturada por gradientes ambientais, sendo mais elevada nas proximidades da isóbata de 1000 m, em áreas de topografia abrupta e no setor norte da Bacia de Santos, além de apresentar variação sazonal marcada. Foram utilizados modelos Bayesianos de árvores de regressão aditivas para avaliar padrões de distribuição e adequabilidade de habitat para diferentes grupos de cetáceos (baleias migratórias, baleias de Bryde, pequenos delfinídeos, grandes delfinídeos, cachalotes e baleias-bicudas), demonstrando elevado desempenho preditivo e evidenciando que características fisiográficas, bem como variáveis oceanográficas dinâmicas, são determinantes na estruturação dos habitats. Adicionalmente, foram projetados cenários futuros de distribuição de cetáceos sob distintos cenários de emissão de gases de efeito estufa, revelando potenciais deslocamentos espaciais, contrações e expansões de áreas ambientalmente adequadas para os diferentes grupos. Foram também produzidas as primeiras estimativas de abundância de cetáceos em escala de toda a Bacia de Santos, por meio de uma abordagem Bayesiana de amostragem por distâncias em estrutura hierárquica e multiespécies. Os resultados revelaram diferenças marcantes entre grupos de espécies, bem como variações sazonais na densidade, estabelecendo uma linha de base demográfica inédita para a região. Esta tese fornece uma visão detalhada sobre a organização espacial e sazonal da comunidade de cetáceos na Bacia de Santos, evidenciando como processos ambientais e climáticos podem influenciar sua distribuição, e ressalta a importância de continuar o monitoramento por meio de diferentes fontes de dados e variáveis bióticas e antrópicas em estudos futuros.

Palavras-chave: Abundância, Atlântico Sudoeste, Comunidade de cetáceos, Modelos de distribuição de espécies, Mudanças climáticas, Preditores ambientais.

ABSTRACT

This thesis presents an integrated investigation of the spatial and temporal patterns of distribution, species richness, and abundance of cetaceans in the Santos Basin, southwestern Atlantic Ocean, with a focus on the effects of environmental variables and potential responses to climate change. Data from systematic ship-based surveys conducted between 2015 and 2019 were analyzed, covering different seasons and over 40,000 km of survey effort. Using a Bayesian framework based on generalized linear spatial models, cetacean species richness was shown to be strongly structured by environmental gradients, being highest near the 1,000 m isobath, in areas of steep topography, and in the northern sector of the Santos Basin, while also exhibiting marked seasonal variation. Bayesian additive regression trees models were used to evaluate distribution patterns and habitat suitability for different cetacean groups (migratory whales, Bryde's whales, small delphinids, large delphinids, sperm and beaked whales), demonstrating high predictive performance and highlighting that physiographic features and dynamic oceanographic variables are key determinants of habitat structure. Future distribution scenarios under different greenhouse gas emission trajectories were projected, revealing potential spatial shifts, contractions, and expansions of environmentally suitable areas for the groups. The study also produced the first abundance estimates of cetaceans across the entire Santos Basin using a hierarchical, multispecies Bayesian distance-sampling approach. The results revealed marked differences among groups and seasonal variations in density, establishing an unprecedented demographic baseline for the region. Overall, this thesis provides a detailed view of the spatial and seasonal organization of the cetacean community in the Santos Basin, illustrating how environmental and climatic processes can influence their distribution, and emphasizes the importance of continued long-term monitoring using multiple data sources and both biotic and anthropogenic variables in future studies.

Keywords: Abundance, Cetacean community, Climate change, Environmental predictors, Southwestern Atlantic, Species distribution models.

PREFÁCIO

Os cetáceos sempre despertaram fascínio, sendo símbolos da complexidade e beleza dos oceanos. Sua presença pode indicar a saúde dos ecossistemas, e sua conservação é um desafio científico e ético global. Compreender a distribuição, abundância e riqueza de espécies em múltiplas escalas é essencial para subsidiar estratégias eficazes de conservação e manejo. Esse estudo visa contribuir para o preenchimento de lacunas de conhecimento sobre padrões ecológicos de cetáceos e fornecer subsídios científicos para o manejo ecossistêmico e a avaliação de riscos no Atlântico Sudoeste frente a mudanças ambientais, aplicando um conjunto inovador de abordagens estatísticas.

Esta tese está organizada em cinco capítulos. O Capítulo I apresenta a introdução, com a contextualização geral do tema, bem como as hipóteses e os objetivos do trabalho, redigidos em língua portuguesa. Os Capítulos II, III e IV são apresentados na forma de artigos científicos, redigidos em língua inglesa e formatados de acordo com as diretrizes das revistas escolhidas para publicação. No Capítulo II, foi avaliada a variação espacial e sazonal da riqueza de espécies de cetáceos, permitindo identificar áreas de maior diversidade e os principais fatores ambientais associados a esses padrões. O Capítulo III investigou como diferentes grupos de cetáceos se distribuem na região, bem como suas potenciais respostas a cenários futuros de mudanças climáticas. No Capítulo IV, foram geradas estimativas de abundância utilizando modelos hierárquicos multiespécies. No Capítulo V são descritas as principais conclusões desta tese. Adicionalmente, na seção de apêndices, são disponibilizados os principais códigos desenvolvidos em linguagem R para a realização das análises.

CAPÍTULO I

Introdução geral

1. INTRODUÇÃO

Os cetáceos (Infraordem Cetacea) constituem um grupo altamente especializado de mamíferos marinhos que inclui as baleias e os golfinhos. Evoluíram de ancestrais terrestres artiodáctilos há cerca de 55 milhões de anos, apresentando uma transição completa para a vida aquática, desenvolvendo adaptações como locomoção hidrodinâmica, mergulho profundo e capacidades sensoriais complexas (Fordyce 2018, Thewissen 2018). Essa radiação adaptativa permitiu a exploração eficiente de nichos ecológicos diversos, sustentando uma ampla diversidade taxonômica e funcional. Variando em tamanho desde cerca de 1 m, em golfinhos recém-nascidos, até mais de 30 m, na baleia-azul adulta (*Balaenoptera musculus*) – os maiores animais que já existiram –, os cetáceos exibem uma ampla gama de estratégias de história de vida. Ocupam praticamente todos os habitats nos oceanos do mundo, de polo a polo, da costa ao oceano aberto e da superfície a grandes profundidades, participando de múltiplos níveis tróficos, de filtradores de zooplâncton a superpredadores (Ballance 2018).

A maioria das grandes baleias-de-barbatana (Mysticeti) alimenta-se principalmente de zooplâncton, que filtram através de placas córneas especializadas chamadas barbatanas. Algumas dessas baleias, como as jubartes (*Megaptera novaeangliae*), minke (*Balaenoptera* spp.), fin (*Balaenoptera physalus*) e sei (*Balaenoptera borealis*) realizam extensas migrações sazonais, geralmente se reproduzindo no inverno em águas tropicais e deslocando-se em direção aos polos para se alimentar no verão. Cetáceos menores não realizam longas migrações, mas se movem em resposta a mudanças na disponibilidade de alimento ou a alterações físicas.

Os Odontoceti são as espécies de cetáceos que possuem dentes e um único espiráculo em vez dos dois das baleias-de-barbatana. As baleias dentadas são predadoras importantes no mar, alimentando-se principalmente de lulas ou peixes e, em alguns casos, de outros mamíferos marinhos. Estes animais não dependem de presas que vivem na superfície e podem mergulhar a profundidades de várias centenas de metros. Os odontocetos podem caçar presas por ecolocalização, na qual emitem pulsos de som e monitoram os ecos que retornam, e pelo menos algumas espécies mostram comportamento cooperativo ao reunir e capturar suas presas (Heithaus et al. 2018).

Nesse contexto, cetáceos desempenham papéis ecológicos fundamentais, influenciando a dinâmica trófica e contribuindo para a saúde dos ecossistemas marinhos, além de atuarem como indicadores de alterações ambientais em diferentes escalas espaciais e temporais (Fossi et al. 2012, Hazen et al. 2019). Apesar de sua importância ecológica e cultural, muitas populações de cetáceos sofreram declínios significativos ao longo do século XX, principalmente devido à caça comercial intensiva (Clapham & Baker 2018). Embora algumas espécies tenham apresentado sinais de recuperação, ameaças contemporâneas como captura acidental em artes de pesca; colisões com embarcações; poluição acústica, química e por resíduos plásticos; além das rápidas mudanças climáticas que podem ocasionar extremos de temperatura, continuam a afetar a distribuição, abundância e viabilidade de suas populações (van Compernolle et al. 2025). Essas pressões ambientais de origem antrópica variam no espaço e no tempo, tornando essencial o monitoramento contínuo e a compreensão dos fatores que moldam padrões populacionais.

A compreensão da distribuição, abundância e riqueza de espécies em múltiplas escalas é um dos pilares da ecologia e da conservação. O estado de conservação dos cetáceos e seus papéis ecológicos só podem ser plenamente entendidos quando analisados em relação aos ambientes que ocupam. As relações entre cetáceos e seu habitat são moldadas por variáveis oceanográficas estáticas e dinâmicas, essas influenciadas por processos sazonais, interanuais e pelo clima (Ballance 2018). Na maioria dos casos, as presas são os prováveis determinantes das relações entre espécies individuais de cetáceos e seus habitats. No entanto, como variáveis e interações bióticas são mais difíceis de mensurar e representar (Soberón & Nakamura, 2009), a maioria das abordagens tem se restringido a representações das condições abióticas nas quais as espécies podem se estabelecer, sobreviver e se reproduzir (Melo-Merino et al., 2020). Entre os fatores ambientais, destaca-se a importância da temperatura da superfície do mar, batimetria e declividade, que afetam direta ou indiretamente a disponibilidade de presas (Whitehead et al. 2008, Cañadas et al. 2002, Di Tullio et al. 2016). Entretanto, os efeitos das variáveis derivadas (por exemplo, declividade, rugosidade, e frentes de temperatura) e suas interações com outros fatores permanece menos compreendido, e há

escassez histórica de modelos aplicados à riqueza e à estrutura comunitária de cetáceos - especialmente no Brasil (Melo-Merino et al. 2020, Rodrigues et al. 2023).

Globalmente, são reconhecidas aproximadamente 94 espécies de cetáceos distribuídas em 14 famílias (Committee on Taxonomy 2025), das quais pelo menos 46 ocorrem no Brasil (ICMBio 2020). A Bacia de Santos (BS), no Atlântico Sudoeste, é uma das áreas de maior diversidade de cetáceos no mundo (Tittensor et al. 2010, Kaschner et al. 2011), abrigando cerca de 83% das espécies registradas no país (Lodi e Borobia 2013, ICMBio 2020). A BS estende-se entre aproximadamente 23° e 28°S ao longo da margem continental sudeste brasileira e integra a região oceanográfica conhecida como South Brazil Bight. Trata-se de uma bacia sedimentar extensa, que se estende por cerca de 1100 km ao longo da costa, com isóbatas aproximadamente paralelas à linha de costa. Sua produtividade é moderada a alta, com forte variabilidade temporal e espacial, influenciada principalmente pela Corrente do Brasil e pela intrusão sazonal da Água Central do Atlântico Sul (ACAS), além da atuação da Corrente das Malvinas durante o inverno. Esses processos modulam a estrutura térmica e a disponibilidade de nutrientes na plataforma e no talude, favorecendo a formação de frentes oceânicas, ressurgências e vórtices, associados a elevada diversidade biológica (Ekau & Knoppers 2003, Pereira e Gomes 2009). Em função de sua complexidade físico-oceanográfica e elevada diversidade biológica, a BS é reconhecida como uma área prioritária para conservação no Brasil (MMA 2018). Adicionalmente, a Bacia de Santos insere-se em um contexto de intensificação das mudanças climáticas no Atlântico Sudoeste. Como o aquecimento dos oceanos ocorre de forma heterogênea, a região está associada a padrões relevantes de aquecimento em escala regional e global, frequentemente descritos como *hotspots* de mudança climática (Hobday & Pecl 2013). Projeções climáticas para o Atlântico Sudoeste indicam a continuidade do aumento da temperatura da superfície do mar, da estratificação da coluna d'água e de possíveis alterações na circulação oceânica e na produtividade primária (Franco et al. 2020), com potenciais efeitos sobre a distribuição das presas (Costa et al. 2021), a adequabilidade de habitat (Guisan & Thuiller 2005) e, consequentemente, os padrões espaciais de ocorrência dos cetáceos.

No escopo dos estudos com cetáceos, a região é uma das mais estudadas no país, com trabalhos abordando uso de habitat, acústica, hábitos alimentares, bioacumulação de contaminantes e parâmetros de história de vida (Moreno et al. 2005, Siciliano et al. 2007, Melo et al. 2010, Dorneles et al. 2013, Santos et al. 2019, Dalpaz et al. 2021, Pires et al. 2021, Correa et al. 2023, Lima et al. 2024). Tanto espécies tropicais quanto de águas temperadas compõem a comunidade de cetáceos da área, incluindo espécies frequentemente associadas ao uso local do habitat, bem como grandes baleias migratórias observadas sazonalmente ao longo do talude continental (Di Tullio et al. 2016). Ao comparar a diversidade e a distribuição de cetáceos no talude continental do sul e sudeste do Brasil durante a primavera e o outono, Di Tullio et al. (2016) registraram 17 espécies na região da Bacia de Santos e não observaram variações aparentes na diversidade total nem na composição de espécies entre as estações amostradas. Os autores também destacaram a relevância do talude continental como área de ocorrência para uma comunidade diversa de cetáceos. Apesar dos importantes avanços no conhecimento sobre a comunidade de cetáceos da região, ainda persistem lacunas relevantes, especialmente quanto aos padrões de distribuição em nível de comunidade ao longo de toda a área de estudo e à abundância da maioria das espécies que utilizam a região. A maior parte das análises ainda se concentra em espécies isoladas ou em poucos grupos taxonômicos, frequentemente restritas a recortes espaciais e temporais específicos, incluindo abordagens comunitárias em escalas delimitadas no espaço e no tempo (Di Tullio et al. 2016, Maricato et al. 2022, Dalpaz et al. 2023, Tardin et al. 2025). Além disso, embora a região abrigue uma elevada diversidade de cetáceos (Lodi & Borobia 2013, ICMBio 2020), ainda há forte limitação de informações quantitativas sobre abundância e tendências populacionais. Apenas algumas populações, como a de baleias-jubarte, possuem séries históricas de monitoramento (Andriolo et al. 2010, Zerbini et al. 2019), enquanto delfinídeos e outros misticetos seguem incertos quanto ao tamanho populacional (IUCN 2023).

Apesar da elevada diversidade taxonômica e comportamental, cetáceos apresentam convergências em padrões macroecológicos relacionados ao uso do habitat, distribuição e resposta a gradientes ambientais, particularmente quando analisados em grupos baseados em sua filogenia e características ecológicas. A macroecologia oferece

uma estrutura conceitual adequada para investigar essas respostas, ao examinar padrões amplos de abundância, distribuição e diversidade de espécies em grandes escalas espaciais e temporais. Para isso, integra princípios de biogeografia, ecologia de comunidades e geociências na busca por regularidades ecológicas e áreas onde a riqueza de espécies é excepcionalmente elevada (Ballance 2018).

No contexto de cetáceos, analisar as espécies em nível de comunidade, e não de maneira isolada, permite compreender como diferentes grupos ocupam o ambiente marinho. Ao agrupar espécies de acordo com sua filogenia e aspectos ecológico-comportamentais, como baleias migratórias, baleias de Bryde, pequenos e grandes golfinhos, além de cachalotes e baleias-bicudas, é possível representar variações importantes nas estratégias ecológicas e de uso do habitat, bem como identificar padrões estruturais mais robustos da comunidade (Di Tullio et al. 2016, Roberts et al. 2016). Essa abordagem permite explorar, de forma integrada, a riqueza da comunidade como um todo, além de padrões espaciais de distribuição para todos os grupos ecológicos e estimativas de densidade para aqueles com maior detectabilidade (Royle e Dorazio 2008, LeDuc 2018). A perspectiva comunitária é particularmente relevante em ambientes marinhos, onde limitações logísticas e baixa detectabilidade podem levar à subestimação de espécies raras ou mais elusivas. A partir dessas relações ecológicas e da escala espacial envolvida, a caracterização de padrões populacionais exige a aplicação de modelos estatísticos adequados à complexidade dos sistemas ecológicos.

Um modelo estatístico é uma construção matemática baseada na teoria da probabilidade que visa reconstruir o sistema ou o processo em estudo; os dados são observações desse sistema ou processo. De forma geral, um modelo estatístico pode ser decomposto em duas partes: uma componente determinística, que descreve a estrutura média do processo por meio de uma relação matemática, e uma componente estocástica (aleatória), que representa a variabilidade inerente ao processo e às observações (Kéry 2010). Fazemos uso de modelos para sintetizar um conhecimento e especular sobre um determinado sistema, e a diferença entre o modelo e as observações aponta para aspectos que merecem investigação mais profunda e eventuais correções no modelo. Portanto, analisar dados usando modelos estatísticos raramente se resume a aplicar um

único modelo aos dados; trata-se de um processo iterativo que envolve comparar estruturas alternativas, avaliar suposições e quantificar as incertezas que emergem tanto dos dados quanto da estrutura dos modelos. Analisar dados usando modelos estatísticos é um processo permanente de aprendizado. Assim, nenhum modelo é perfeito, mas um bom modelo é útil para nos aproximar da verdade (Box 1979).

Em ecologia, a modelagem estatística é necessária pois permite inferir a estrutura e o funcionamento dos sistemas naturais e prever os efeitos de processos ecológicos. Grande parte da teoria ecológica contemporânea, bem como aplicações em conservação e gestão de recursos naturais, concentra-se em entender a variação na abundância ou ocorrência das espécies. Essa variação ocorre no espaço e no tempo e está frequentemente associada a diferenças mensuráveis nas características ambientais. Com a crescente quantidade e complexidade dos dados, precisamos de modelos mais elaborados que resumam os dados de forma coerente. O avanço da capacidade computacional expandiu o uso de modelos estatísticos mais flexíveis e realistas, com destaque para métodos Bayesianos que incorporam explicitamente diferentes fontes de incerteza (Gelman et al. 2020, Kinas & Andrade 2020). A abordagem Bayesiana é particularmente adequada para estudos ecológicos, pois permite combinar dados observacionais com conhecimento prévio de forma coerente e quantificar as incertezas de maneira explícita em todos os níveis da modelagem. Além disso, modelos bayesianos oferecem vantagens distintas em contextos de volume limitado de dados, devido ao uso de simulações MCMC e de prioris regularizadores, o que aumenta a sua estabilidade computacional e a precisão inferencial (Gelman et al. 2020). A obtenção das distribuições posteriores, por meio de algoritmos de amostragem MCMC, fornece inferência probabilística completa sobre parâmetros e previsões. Métodos recentes, como Bayesian Additive Regression Trees (BART) (Chipman et al. 2010, Carlson 2020), expandem o poder preditivo de modelos de distribuição de espécies, enquanto o Hierarchical Distance Sampling (HDS) permite estimar densidade e abundância ao corrigir a detectabilidade e incorporar variáveis ambientais em múltiplas espécies ou grupos (Sollmann et al. 2016).

Diante disso, este trabalho integra avanços analíticos para explorar padrões de comunidade, grupos ecológicos e variáveis ambientais em um hotspot de diversidade de cetáceos do Atlântico Sudoeste. O estudo preenche lacunas críticas sobre os padrões ecológicos na Bacia de Santos e fornece base científica para o manejo ecossistêmico e a avaliação de riscos frente às mudanças ambientais, à luz de novas abordagens quantitativas.

1.1. HIPÓTESES

Considerando as informações apresentadas nesta introdução, o presente estudo partiu das seguintes hipóteses, formuladas no projeto de Tese, sobre a riqueza de espécies e da densidade de cetáceos na Bacia de Santos:

(i) a riqueza e a densidade de cetáceos aumentam em resposta a condições ambientais associadas à maior disponibilidade de presas, sendo positivamente relacionadas à produtividade primária e apresentando relação não linear com a temperatura da superfície do mar; e

(ii) a distribuição espacial da riqueza e densidade de cetáceos é estruturada por características geomorfológicas, com maiores valores associados à plataforma externa e ao talude devido à maior heterogeneidade ambiental e às condições que favorecem a disponibilidade de recursos tróficos.

Ao longo do desenvolvimento da pesquisa, a incorporação de abordagens preditivas baseadas em modelos Bayesianos permitiu expandir o escopo das hipóteses originais. Assim, passou-se a testar adicionalmente que:

(iii) mudanças projetadas na temperatura da superfície do mar e na produtividade oceânica alteram a adequabilidade de habitat para diferentes grupos ecológicos de cetáceos, resultando em mudanças na distribuição potencial e em respostas diferenciais entre grupos com distintas afinidades ecológicas e ambientais.

1.2. OBJETIVOS

1.2.1. Geral

Avaliar os padrões espaciais e temporais de distribuição, densidade e riqueza de espécies de cetáceos na Bacia de Santos, considerando a influência de variáveis ambientais e processos oceanográficos.

1.2.2. Específicos

(i) Relacionar distribuição, densidade e riqueza de espécies de cetáceos com covariáveis ambientais e processos oceanográficos, utilizando modelagem estatística;

(ii) Identificar padrões espaciais e sazonais de densidade e riqueza de espécies de cetáceos na Bacia de Santos;

(iii) Avaliar padrões de distribuição atuais e potenciais futuros entre grupos de cetáceos, com base em modelos preditivos.

REFERÊNCIAS BIBLIOGRÁFICAS

Andriolo A, Kinas PG, Engel MH, Martins CCA, Rufino AM (2010) Humpback whales within the Brazilian breeding ground: distribution and population size estimate. Endang Species Res 11:233–243.

Ballance LT (2018) Cetacean ecology. In: Würsig B, Thewissen JGM, Kovacs K (eds) Encyclopedia of marine mammals. Academic Press, London.

Box GP (1979) Robustness in the strategy of scientific model building. In: Launer RL, Wilkinson GN (eds) Robustness in statistics. Academic Press, New York.

Cañadas A, Sagarminaga R, García-Tiscar S (2002) Cetacean distribution related with depth and slope in the Mediterranean waters off southern Spain. Deep-Sea Res I 49:2053–2073.

Carlson CJ (2020) embarcadero: species distribution modelling with Bayesian additive regression trees in R. Methods Ecol Evol 11:850–858.

Chipman HA, George EI, McCulloch RE (2010) BART: Bayesian additive regression trees. Ann Appl Stat 4:266–298.

Clapham PJ, Baker CS (2018) Whaling, modern. In: Würsig B, Thewissen JGM, Kovacs K (eds) Encyclopedia of marine mammals. Academic Press, London.

Committee on Taxonomy (2025) List of marine mammal species and subspecies. Society for Marine Mammalogy. www.marinemammalscience.org.

Corrêa TP, Wedekin LL, Cremer MJ (2023) Occurrence and habitat use of pantropical spotted dolphins (Stenella attenuata) at the edge of their distribution in the southwestern Atlantic Ocean. Mar Mamm Sci 40:e13094.

Costa MD, Wilson KA, Dyer PJ, Pitcher R, Muelbert JH, Richardson AJ (2021) Potential future climate-induced shifts in marine fish larvae and harvested fish communities in the subtropical Southwestern Atlantic Ocean. Clim Change 165:66.

Dalpaz L, Paro AD, Daura-Jorge FG, Rossi-Santos M, Norris TF, Ingram SN, Wedekin LL (2021) Better together: Analysis of integrated acoustic and visual methods when surveying a cetacean community. Mar Ecol Prog Ser 678:197-209.

Dalpaz L, Cypriano-Souza AL, Lodi L, Wedekin L, Daura-Jorge F (2023) Bryde's whales in South Brazil Bight: Evidence of low genetic diversity and seasonal habitat use. Mar Biol 170:94.

Di Tullio JC, Gandra TB, Zerbini AN, Secchi ER (2016) Diversity and distribution patterns of cetaceans in the subtropical southwestern Atlantic outer continental shelf and slope. PLoS ONE 11:e0155841.

Dorneles PR, Sanz P, Eppe G, Azevedo A.F, Bertozzi C.P, Martínez MA, Secchi ER, Barbosa LA, Cremer M, Alonso MB, Torres JPM, Lailson-Brito J, Malm O, Eljarrat E, Barceló D, Das K (2013). High accumulation of PCDD, PCDF, and PCB congeners in marine mammals from Brazil: a serious PCB problem. Sci Total Environ 463–464.

Ekau W, Knoppers B (2003) A review and redefinition of the large marine ecosystems of Brazil. In: Sherman K, Hempel G. (eds) Large Marine Ecosystems of the World: Trends in Exploitation, Protection and Research. Elsevier Science, Amsterdam.

Fordyce RE (2018) Cetacean evolution. In: Würsig B, Thewissen JGM, Kovacs K (eds) Encyclopedia of marine mammals. Academic Press, London.

Fossi MC, Casini S, Caliani I, Panti C, Marsili L, Viarengo A, Giangreco R, di Sciara GN, Serena F, Ouerghi A, Depledge MH (2012) The role of large marine vertebrates in the assessment of the quality of pelagic marine ecosystems. Mar Environ Res 77:156–158.

Franco BC, Defeo O, Piola AR, Barreiro M, Yang H, Ortega L, Gianelli I, Castello JP, Vera C, Buratti C, Pájaro M, Pezzi LP, Möller OO (2020) Climate change impacts on the atmospheric circulation, ocean, and fisheries in the southwest South Atlantic Ocean: a review. Clim Change 162:2359–2377.

Gelman A, Vehtari A, Simpson D, Margossian CC, Carpenter B, Yao Y, Kennedy L, Gabry J, Bürkner PC, Modrák M (2020) Bayesian workflow. arXiv preprint arXiv:2011.01808.

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993–1009.

Hazen EL, Abrahms B, Brodie S, Carroll G, Jacox MG, Savoca MS, Scales KL, Sydeman WJ, Bograd SJ (2019) Marine top predators as climate and ecosystem sentinels. Front Ecol Environ 17:565–574.

Heithaus MR, Dill, LM, Kiszka JJ (2018) Feeding strategies and tactics. In: Würsig B, Thewissen JGM, Kovacs K (eds) Encyclopedia of marine mammals. Academic Press, London.

Hobday AJ, Pecl GT (2014) Identification of global marine hotspots: sentinels for change and vanguards for adaptation action. Rev Fish Biol Fish 24:415–425.

ICMBio (Instituto Chico Mendes de Conservação da Biodiversidade) (2020) Guia ilustrado de identificação de cetáceos e sirênios do Brasil. ICMBio/CMA, Brasília.

IUCN SSC Cetacean Specialist Group (2023) Red-list status and extinction risk of the world's whales, dolphins and porpoises. Conserv Biol 37:e14090.

Kaschner K, Tittensor DP, Ready J, Gerrodette T, Worm B (2011) Current and future patterns of global marine mammal biodiversity. PLoS ONE 6:e19653.

Kéry M (2010) Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses. Academic Press, Burlington.

Kinas PG, Andrade HA (2020) Introdução à análise Bayesiana (com R). Consultor Editorial, Porto Alegre.

LeDuc R (2018) Biogeography. In: Würsig B, Thewissen JGM, Kovacs K (eds) Encyclopedia of marine mammals. Academic Press, London.

Lima GS, Pedrobom JH, Suarez CA, Torres-Florez JP, Vidal LG, Domit C, Menegario AA (2024) Bioaccumulation of trace elements in marine mammals: new data and transplacental transfer on threatened species. Sci Total Environ 948:174854.

Lodi L, Borobia M (2013) Baleias, botos e golfinhos do Brasil: guia de identificação. Technical Books Editora, Rio de Janeiro.

Maricato G, Tardin R, Lodi L, Wedekin LL, Daura-Jorge FG, Maciel I, Maria TF, Alves MAS (2022) Identifying suitable areas for common bottlenose dolphin in anthropized waters. Mar Biol 169:110.

MMA (2018) Zona Costeira e Marinha: Áreas e ações prioritárias para conservação da Biodiversidade Brasileira. Ministério do Meio Ambiente, Brasília.

Melo CLC, Santos RA, Bassoi M, Araújo AC (2010) Feeding habitats of delphinids (Mammalia:Cetacea) from Rio de Janeiro State, Brazil. J Mar Biolog Assoc U.K. 90: 1509-1515.

Melo-Merino SM, Reyes-Bonilla H, Lira-Noriega A (2020) Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. Ecol Model 415:108837.

Moreno IB, Zerbini AN, Danilewicz D, de Oliveira Santos MC, Simões-Lopes PC, Lailson-Brito J, Azevedo AF (2005) Distribution and habitat characteristics of dolphins of the genus Stenella (Cetacea: Delphinidae) in the southwest Atlantic Ocean. Mar Ecol Prog Ser 300:229-240.

Pereira RC, Gomes AS (2009) Biologia Marinha. Editora Interciência, Rio de Janeiro.

Pires CR, Rossi-Santos MR, Paro AD, Wedekin LL (2021) Whistles of the pantropical spotted dolphin (Stenella attenuata) in Santos Basin, western South Atlantic Ocean. J Acoust Soc Am 149:3241–3249.

Rodrigues LS, Daudt NW, Cardoso LG, Kinas PG, Conesa D, Pennino MG (2023) Species distribution modelling in the Southwestern Atlantic Ocean: a systematic review and trends. Ecol Model 486:110029.

Royle JA, Dorazio RM (2008) Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Academic Press, London.

Santos, MC de O, Laílson-Brito J, Flach L, Oshima JEF, Figueiredo GC, Carvalho RR, Ventura ES, Molina JMB, Azevedo AF (2019) Cetacean movements in coastal waters of the southwestern Atlantic ocean. Biota Neotrop 19:e20180670.

Siciliano S, Ramos RMA, di Beneditto APM, Santos MCO, Fragoso AB, Brito JL, Azevedo AF, Vicente AFC, Zampirolli E, Alvarenga FS, Barbosa L, Lima NRW (2007) Age and growth of some delphinids in south-eastern Brazil. J Mar Biol Assoc U. K. 87: 293–303.

Soberón J, Nakamura M (2009) Niches and distributional areas: concepts, methods, and assumptions. Proceedings of the National Academy of Sciences 106 (Supplement 2): 19644–19650.

Sollmann R, Gardner B, Williams KA, Gilbert AT, Veit RR (2016) A hierarchical distance sampling model to estimate abundance and covariate associations of species and communities. Methods Ecol Evol 7:529–537.

Tardin R, Maricato G, Kiszka JJ, Cantor M, Maciel I, Melo-Santos G, Vale MM (2025) Optimistic climate mitigation scenario halves projected range loss in a neotropical dolphin. Ocean Coast Manag 269:107800.

Thewissen JGM (2018) Archaeocetes, Archaic. In: Würsig B, Thewissen JGM, Kovacs K (eds) Encyclopedia of marine mammals. Academic Press, London.

Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, Berghe EV, Worm B (2010) Global patterns and predictors of marine biodiversity across taxa. Nature 466:1098–1101.

van Compernolle M, Morris J, Calich HJ, et al. (2025) Vulnerability of marine megafauna to global at-sea anthropogenic threats. Conserv Biol 40:e70147.

Whitehead H, McGill B, Worm B (2008) Diversity of deep-water cetaceans in relation to temperature: implications for ocean warming. Ecol Lett 11:1198–1207.

Zerbini AN, Adams G, Best J, Clapham PJ, Jackson JA, Punt AE (2019) Assessing the recovery of an Antarctic predator from historical exploitation. R Soc Open Sci 6:190368.

CAPÍTULO II

Spatial and seasonal patterns of cetacean species richness: A Bayesian approach

(artigo publicado na revista Deep-Sea Research Part I – IF 2.3, CAPES Qualis Biodiversidade A1)

DOI: https://doi.org/10.1016/j.dsr.2023.104046

Spatial and seasonal patterns of cetacean species richness: a Bayesian approach

Ana Carolina Martins[a,b*], Paul Gerhard Kinas[a,b], Leonardo Liberali Wedekin[c] and Luciano Dalla Rosa[a,d]

[a] Programa de Pós-Graduação em Oceanografia Biológica, Universidade Federal do Rio Grande – FURG, 96203-900, Rio Grande, RS, Brazil

[b]Laboratório de Estatística Ambiental, Instituto de Matemática, Estatística e Física, Universidade Federal do Rio Grande – FURG, 96203-900, Rio Grande, RS, Brazil

[c] Socioambiental Consultores Associados, 88015-200, Florianópolis, SC, Brazil

[d] Instituto de Oceanografia, Universidade Federal do Rio Grande – FURG, 96203-900, Rio Grande, RS, Brazil

* *Corresponding author: tel: +55(53)984193893*; *e-mail*: ana.carolbm@hotmail.com

Abstract

Research about community-wide patterns in cetaceans is challenging, so seasonal cycles and patterns of distribution are still poorly understood for many locations. In the present study, we proposed Bayesian Spatial General Linear Models (GLMs) to quantify and map cetacean species richness in southeastern Brazil's Santos Basin, considering environmental effects, effort, and seasonality. Systematic surveys were conducted in Santos Basin along zigzag-shaped transects over four years (November 2015 to April 2019), allowing consistent data collection over 40,000 km of effort through all seasons. A set of nine static physiographic data and eight dynamic environmental covariates were included in the models. A total of 371 sightings were recorded comprising 24 species. The best-fit model contained sea surface temperature, distance to 1000 m isobath, latitude, slope, and effort as covariates. We report that the cetacean species richness at Santos Basin is higher near the 1000 m isobath, in abrupt topographies, and off the northern portion; varies seasonally; is not spatially correlated; and is mainly affected by environmental predictors. These models and maps established a current characterization

of cetacean species richness in these waters and have the potential to be applied in planning marine conservation, mainly in spatial prioritization programs.

Keywords: Atlantic Ocean; Community ecology; Distribution; Environmental factors; GLMs; Habitat.

1. Introduction

Quantifying species richness and identifying the variables that exert the most influence on this process are critical components of current studies in ecology and conservation biology. The most recent assessment of cetacean species considers 92 species distributed among 14 families (Committee on Taxonomy, 2021). Thus far, at least 46 species have been registered off the Brazilian coastline (ICMBio, 2020).

The conservation status of cetaceans and their roles in the ecosystems can be truly understood in the context of the environments they inhabit. And considering the multitude of threats they face in these environments, understanding the relationships cetaceans have with their habitat is urgent; this is especially true regarding climate change, which impacts the distribution and geographic scope of habitats on which cetaceans depend (Ballance, 2018; Stachowitsch *et al*., 2019). The study of cetacean species-habitat relationships involves the integration of these two different types of data (species and habitat) with sophisticated analytical and computational tools (Redfern *et al*., 2006). Within the context of spatially structured ecological data, two of the main objectives are to describe patterns of distribution and understand ecological processes. Among the many methodological tools that have been proposed for modelling species distributions are the correlative models (Guisan and Thuiller, 2005; Franklin, 2010). In most cases, prey are the likely drivers of individual cetacean species-habitat relationships. However, since biotic variables and interactions are harder to measure and depict in a raster grid format (Soberón and Nakamura, 2009), most correlative models have been restricted to representations of the abiotic conditions under which the species can become established, survive and breed (Melo-Merino *et al*., 2020). Environmental

covariates in correlative species distribution models provide information about the underlying processes that drive species distribution. It is important to point out that marine habitats are not static in space and time; rather, they change as currents alter the properties of local water masses and with seasonal, interannual, and climatic cycles (Ballance, 2018).

Although our knowledge of individual cetacean species has advanced rapidly, community-wide patterns of abundance and diversity are less understood. Regarding environmental variables, sea surface temperature has been considered the most influential for cetacean diversity (Whitehead *et al*., 2008; Tittensor *et al*., 2010; Kaschner *et al*., 2011). Depth and slope have been studied in association with preferential cetacean habitat and community structure (Cañadas *et al*., 2002; Tittensor *et al*., 2010; Di Tullio *et al*., 2016). A growing number of studies have identified other environmental variables, such as primary productivity, salinity, and currents, as well as oceanographic processes, such as ocean fronts, as influencing the general distribution of cetaceans (Davis *et al*., 2002; Ballance *et al*., 2006; Whitehead *et al*., 2010; Dalla Rosa *et al.*, 2012; Dey *et al*., 2021). However, the application of this modelling approach to the study of cetacean richness remains scarce in the literature and, to our knowledge, has not be employed off the coast of Brazil.

The Santos Basin (SB) is a sedimentary basin in the southeastern region of the Brazilian continental margin between 23° and 28° S. It extends approximately 1100 km along the coast, and the isobaths run approximately parallel to the coastline. In the oceanographic literature, this region of the Brazilian coast is referred to as the South Brazil Bight. The region's productivity is moderate to high and characterized by great temporal and spatial variability (Ekau and Knoppers 2003). The SB is mostly influenced by the tropical waters of the Brazil Current, which transports warm and oligotrophic waters. The most prominent oceanographic characteristic of the area is the seasonal presence of the South Atlantic Central Water (SACW) in the interior region of the continental shelf. SACW upwelling regions, with an increase in the primary productivity, can be observed mainly in the summer and on the Rio de Janeiro coast. During winter months, the SACW retracts towards the edge of the continental shelf, and

the Malvinas Current exerts greater influence over the area (Pereira and Gomes, 2009). The encounter of this cold current with the warm waters of the Brazil Current characterizes the Subtropical Convergence (Campos *et al*., 1995). However, the Malvinas Current is more influential further to the south in the Santos Basin. These dynamics drive upwelling and increase productivity. Therefore, a considerable number of Brazilian fisheries target this area, mainly for the local sardine *Sardinella brasiliensis* (Brandini *et al*., 2018). With respect to cetacean diversity, the waters of southeastern Brazilian are among the richest worldwide (Tittensor *et al*., 2010, Kaschner *et al*., 2011). Due to its high biological diversity in general, this region is a high conservation priority in Brazil (MMA, 2018). In the SB, based on the sighting history and stranding statistics, about 38 species are potentially available to sighting (Lodi and Borobia, 2013); this is about 83% of all cetacean species registered in Brazilian waters. While comparing cetacean diversity and distribution at the continental slope in southern and southeastern Brazil during spring and autumn, Di Tullio *et al*. (2016) encountered 17 species in this region of SB and found no apparent variation in overall diversity and species composition from one covered season to another.

The expansion of computer power over the last decade has allowed for the use of more complex and realistic statistical models. Among such developments, Bayesian analyses are at the forefront (Gelman *et al*., 2020). The advantage of incorporating Bayesian inference into predictive models is that Bayesian analyses allow for the integration of a variety of uncertainties and exclusively using probability as its metric. Combining the uncertainty in the data (expressed by the likelihood) with additional data information (expressed by prior distributions), posterior probability distributions are developed for all unknown quantities of interest using Bayes' theorem (Gelman *et al*., 2013; Kinas and Andrade, 2020). Studies modelling species richness, abundance, and density of birds (Dorazio *et al*., 2006; Royle and Kéry, 2007; Kéry, 2010; Nunes and Bugoni, 2017), insects (Royle and Kéry, 2007; Szewczyk and McCain, 2019) and marine fishing resources (Kinas and Andrade, 2007; Rufener *et al*., 2017, Rodrigues *et al.*, 2022) have increasingly utilized Bayesian tools due to their flexibility and potential to handle more complex and realistic models. However, the use of this approach to examine cetacean ecology is still scarce. Within this framework, our objective was (i) to

evaluate and model the effect of covariates on cetacean species richness and (ii) to predict species richness spatially and temporally at two periods of the year, taking into account environmental effects.

2. Material and Methods

2.1. Data acquisition and processing

2.1.1. Cetacean surveys

The Santos Basin was sampled from the shoreline to the boundary of the Brazilian Exclusive Economic Zone during eight systematic ship surveys conducted as part of the "Projeto de Monitoramento de Cetáceos da Bacia de Santos" (PMC-BS) project. These seasonal surveys covered approximately 44,000 km of zigzag-shaped transect lines between 2015 and 2019. Survey periods and routes are detailed in Table 1 and Figure 1. The transect lines were intended to cover the Santos Basin in a systematic and homogeneous way, so that the cetacean counts could be extrapolated to the entire study area. The transects were oriented perpendicular to the coast, crossing the depth contours of the study area, and were shifted between 60 and 90 km between surveys in order to avoid spatial biases in the sampling. The vessel travelled along the transect lines at an average speed of 7 knots.

2.1.2. Effort and sighting data

Searches were conducted during daylight period, with three observers positioned at the vessel's highest point, approximately 7 metre above sea level. The three observers rotated periodically with two other observers at rest and searched for groups of cetaceans with naked eye and with the aid of binoculars (7x50 and 15x50). Observers in rest always tried to obtain photographs to help identification. The sampling effort was strictly regimented, and the time and location of all observations were recorded. For each detection of a cetacean group, the horizontal angle in relation to the trackline and the vertical angle to the horizon were recorded with an angle board and reticulated binoculars, respectively. This allowed for the geographic position (registered with a

GPS) of the cetacean group to be corrected using the perpendicular distance from the transect while following the established distance sampling methodology (Buckland *et al*., 2004). Sighting data included time, location, species identification, group size and composition, the presence of offspring and behaviour (Dalpaz *et al.,* 2021).

Table 1. Survey periods and summaries for cetacean sightings and search effort by cruise.

cruise	**year**	**period**	**N sightings**	**N species**	**N species (cumulative)**	**effort (h)**	**effort (km)**
1	2015 - 2016	20 Nov - 5 Jan	31	10	10	389.9	4,937.1
2	2016	10 May - 20 Jun	57	12	16	410.9	5,462.5
3	2016	27 Sep - 6 Nov	74	14	18	420.1	5,500.2
4	2017	8 Mar - 18 Apr	38	8	18	434.9	5,541.9
5	2017	25 Aug - 7 Oct	94	11	18	435.7	5,583.8
6	2018	12 Jan - 20 Feb	44	11	21	444.3	5,863.4
7	2018	1 Sep - 20 Oct	110	14	21	477.2	5,930.5
8	2019	13 Mar - 26 Apr	58	17	24	425.3	5,502.3
Total	-	-	506	-	24	3,438.3	44,321.7

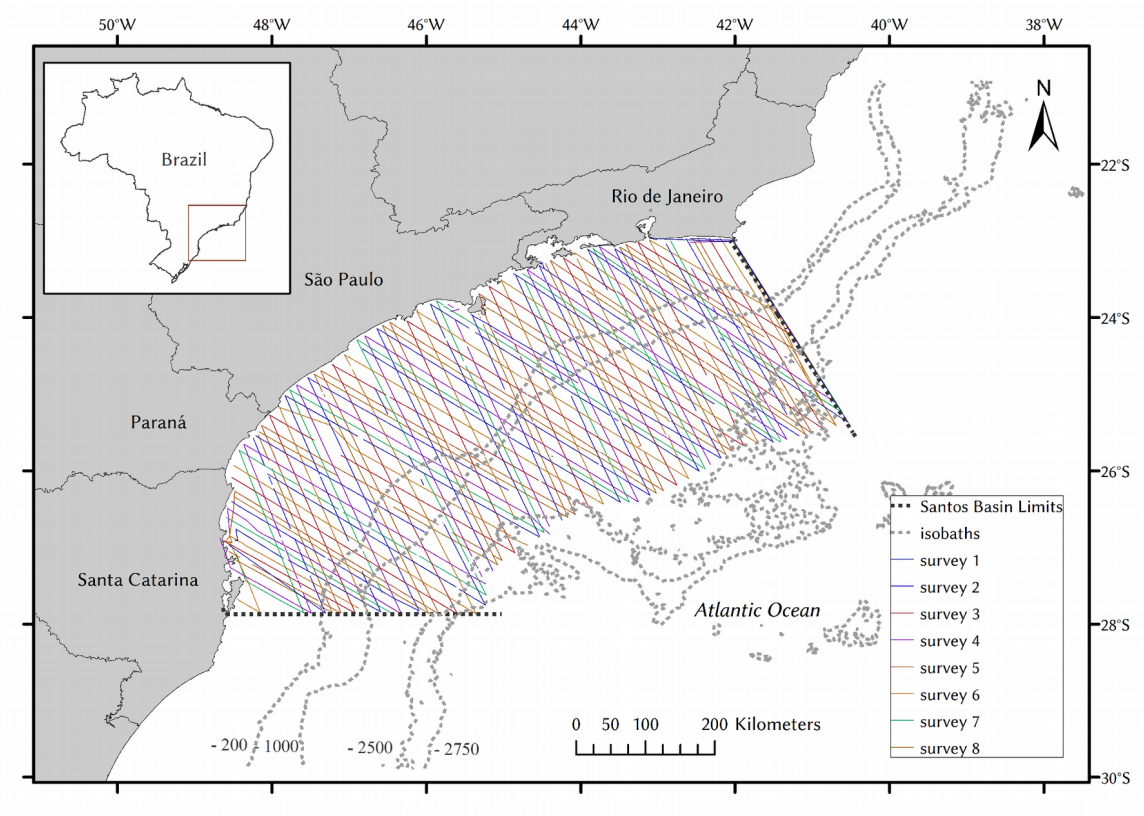


Figure 1. Study area highlighting the transect lines of the cetacean surveys conducted in the Santos Basin.

2.1.3. Environmental data

Habitat predictors included a suite of static physiographic data and dynamic environmental covariates and were obtained from gridded remote sensing and ocean model products (Table 2). These were chosen based on their perceived ecological relevance for cetacean distribution. Monthly values of environmental variables were associated with the latitude and longitude at the centre of each grid cell. To match the data with the scale of the grid cells, covariates were averaged by spatial resolution to obtain a single monthly value for each cell. All geoprocessing and sampling of covariates were performed using ArcGIS 10.6.1 (ESRI, 2018) and the Marine Geospatial Ecology Tools toolbox (Roberts *et al*., 2010). More details are provided in Supplementary Appendix A.

Table 2. Candidate environmental covariates for the spatial-seasonal models.

environmental variable	**unit**	**transfomation**	**temporal resolution**	**acronym**	**source**
Latitude	degree	none	static	lat	--
Longitude	degree	none	static	long	--
Depth	m	none	static	elev	GEBCO
Slope	degree	log	static	slope	GEBCO
Rugosity	index	none	static	rug	GEBCO
Distance to land	m	square root	static	d_coast	IBGE
Distance to 50 m isobath	m	square root	static	d_50	IBGE
Distance to 200 m isobath	m	square root	static	d_200	IBGE
Distance to 1000 m isobath	m	square root	static	d_1000	IBGE
Temperature	°C	none	monthly	sst	NASA Ocean Color
Chlorophyll a	mg/m3	log	monthly	chl	NASA Ocean Color
Salinity	psu	none	monthly	sal	HYCOM
Sea Surface Height	m	none	monthly	ssh	HYCOM
Sea Level Anomaly	m	none	monthly	sla	AVISO
Zonal current velocity	m/s	none	monthly	vel_u	HYCOM
Meriodional current velocity	m/s	none	monthly	vel_v	HYCOM
Distance to SST fronts	m	none	monthly	dist_fronts	NASA Ocean Color

2.1.4. Data organization

The study area was spatially gridded into 0.5° x 0.5° cells. Effort was defined as the total transect length (km) within a given grid cell multiplied by the cell area (km^2). On effort sightings and tracklines were imported into geodatabases in ArcGIS. The grid cells were therefore the statistical unit, and the number of cetacean species identified per cell represented species richness during the two periods: warm (October – March) and cold (April – September). The periods were defined based on known hydrological characteristics of SB and visual inspection of environmental images. The same months were grouped through the years to form one dataset for each period.

2.2. Analytical methods

2.2.1. Modelling species richness

Exploratory data analysis was conducted to identify potential problems in the data that could bias the model results. Pairplots of all explanatory variables were used to identify variables that were correlated. Variables in highly correlated pairs (Pearson correlation coefficient, $|r| > 0.7$) were not used together in the same model. Most explanatory variables were transformed according to the data characteristics to improve the fit (Zuur *et al*., 2007).

General linear models (GLMs) were used to describe species richness within a spatial lattice data structure (Blangiardo and Cameletti, 2015). The response variable $Y(s)$ was defined as the random aggregate of the counted number of cetacean species over areal units s with well-defined boundaries in the region A. This global region A represents a countable collection of 2-dimensional spatial units $\{s_1, s_2, \ldots, s_G\}$. To simplify notation, we replaced each area (or cell) unit s_i by the cell index i, such that $Y(s_i) = Y_i$ for $i = 1, 2, \ldots, G$, where G is the total number of spatial cells in A.

We start by assuming that Y_i, the counted number of species in spatial cell i, follows a Poisson distribution with parameter λ_i.

$$Y_i \sim Poisson\ (\lambda_i)$$

Using a log-link function, we define η_i in terms of a linear predictor,

$$\log(\lambda_i) = \eta_i = \log\ (f_i) + \sum_{j=1}^{M} \beta_j \cdot x_{i,j}$$

where f_i is the effort in km and was included in each spatial cell i as an offset, while covariates were entered as a vector $X_i = \{x_{i,1}, x_{i,2}, \ldots, x_{i,M}\}$ of length M, which is the total number of model parameters to be estimated. The linear and quadratic orthogonal effects for *SST* and *chlorophyll* were also included. Log-effort was included as an offset (as in the linear predictor given above), but also tested as a covariate (with a coefficient to be estimated from the data).

2.2.2. Model selection and Bayesian parameter estimation

Modelling proceeded in two steps. First, due to the high number of environmental variables available, a stepwise procedure (both forward and backwards) was applied to select the model that best fit the data, following the structure described in 2.2.1 and in conjunction with AIC (Akaike's Information Criterion; Akaike, 1973) and BIC (Bayesian Information Criterion; Schwarz, 1978). Goodness of fit was evaluated according to standard graphical checks, such as by residual scatter and quantile-quantile plots, and by conducting chi-square tests on model deviances. Spatial autocorrelation in the residuals was investigated by inspecting the spatial distribution of residuals and through a variogram analysis. The independence among residuals was assessed by a Monte Carlo test using the empirical variogram, where the envelope reflects the premise of no spatial correlation and is given by permutations of the data across the sampling locations (Diggle and Ribeiro, 2007).

In the second modelling stage, once model selection and evaluation had been completed, the results from the best fit model were used for inference through posterior distributions, which were calculated assuming flat prior distributions (Gelman and Hill, 2006). We drew 2000 random values for each parameter while taking the correlation between the parameters into account; that is, from the joint posterior distribution of the

model parameters (Gelman and Hill, 2006; Korner-Nievergelt *et al*., 2015). Posterior predictive model checking was investigated using empirical cumulative distribution function (ECDF) plots, by comparing the ECDFs of simulated predictions from the models to the ECDF of the observed data. The fit of the posterior distribution was also evaluated by calculating the Bayes p-value as a goodness-of-fit statistic (Gelman *et al*., 2013; Kinas and Andrade, 2020), that is, the expected posterior probability that the replicated data could be more extreme than the observed data. All modelling procedures were conducted using the R statistical software (R Core Team, 2021), by means of the following packages: *arm* (Gelman *et al*., 2021), *geoR* (Ribeiro Jr and Diggle, 2020), *sp* (Pebesma and Bivand, 2005), and *stats* (R Core Team, 2021).

2.2.3. Predicting species richness

Output maps of predicted cetacean species richness were produced to display spatial and seasonal variation and map sensitive habitats. For mapping purposes, the effort was fixed at its average. The two periods were represented by their *SST* means, and we calculated 2000 draws from the posterior distribution of the estimated species richness using the combination of the predictor levels (Korner-Nievergelt *et al*., 2015). Finally, each cell displayed represents the average of 2000 posterior predictive simulations.

3. Results

3.1. Data overview

A total of 44,321.7 km of tracklines were divided into 244 spatial-temporal cells (122 for each period) that covered offshore and coastal areas. A total of 371 cetacean group sightings were available for 144 of these spatial-temporal cells (75 cold season cells and 69 in the warm season). Baleen whales were represented by six species of the family Balaenopteridae. Toothed cetaceans were represented by fifteen species in the Delphinidae family and one species each in the Kogiidae, Physeteridae and Ziphiidae families. Most (88%) of the 24 identified species were recorded during the first six

surveys (Supplementary Figure 4). The cetacean sightings per species and period (cold/warm) are summarized in Supplementary Table 1, and visual information related to the sightings and effort by grid cell is reported in Supplementary Figures 1-3.

The observed species richness per cell ranged from 0 to 6, and the mean posterior predicted species richness per cell ranged from 0 to 3. The general species richness predicted average for the cold period was higher than the general average for the warm period (Table 3).

Table 3. Posterior summary of species richness estimates for 122 cells in warm and cold periods (Sd = standard deviation; CI95% = 95% credible interval).

Period	Mean	Sd	CI95%	
			Q0.025	Q0.975
warm	1.056	359	544	1.740
cold	1.437	489	739	2.370

3.2. Model selection and Bayesian parameter estimation

Among the environmental predictors, high correlations were found between (a) salinity and sea surface height, (b) salinity and chlorophyll-a concentrations, (c) 200 m and 1000 m isobaths and (d) distance to shore and 50 m isobath. Therefore, these covariates were not used together during the modelling procedures.

Table B1 (see supplementary Appendix B) summarizes the most relevant models selected by stepwise procedures regarding all combinations of environmental predictors for cetacean species richness. The most parsimonious model for species richness (Model C1 in Table B1) included (i) sea surface temperature and its quadratic term, (ii) rugosity, (iii) distance to the 1000 m isobath, (iv) latitude, (v) slope and (vi) effort as relevant predictors. Square root transformation was used to best fit variable (iii), while logarithm transformation was used with variables (v) and (vi). We also tested models that included smoothing effects; however, they did not show relevant fit improvement. Goodness-of-fit measures of the GLMs were determined to be adequate as evaluated by residuals and

autocorrelation measures. More details of stepwise procedures, criteria and fit visualizations are provided in Supplementary Appendix B.

Posterior summary statistics for all parameters of the selected model are denoted in Table 4. Figures 2 and 3 show posterior credibility intervals and distributions for the model coefficients. According to the best model, *SST* had a quadratic effect on species richness, indicating maximum expected richness at temperatures around 23°C (Figure 4). In this figure, we can also see that there is a marked division between observations in the cold and warm periods defined in this study.

Table 4. Summary of parameters from the Bayesian GLM and their associated statistics (Sd = standard deviation; CI95% = 95% credible interval).

Parameter	Mean	Sd	CI95%	
			Q0.025	Q0.975
(Intercept)	-19.136	4.736	-28.275	-10.094
log (esf)	567	140	301	838
rug	144	62	24	265
sqrt (d_1000)	-354	93	-527	-172
sst	-3.167	1.203	-5.569	-845
sst2	-2.878	1.234	-5.398	-583
lat	167	68	40	301
log (slope)	175	93	8	358

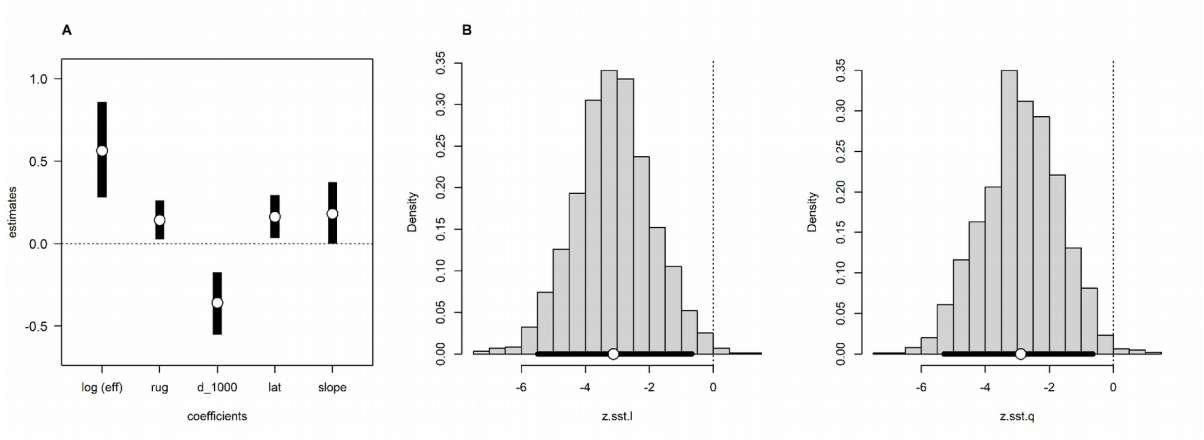


Figure 2. Graphical visualization of parameters from the Bayesian Poisson GLM, where (A) represents posterior 95% credibility intervals of *eff*, *rug*, *d_1000*, *lat* and *slope*, and (B) posterior histograms and 95% credibility intervals for linear and orthogonal quadratic *SST*.

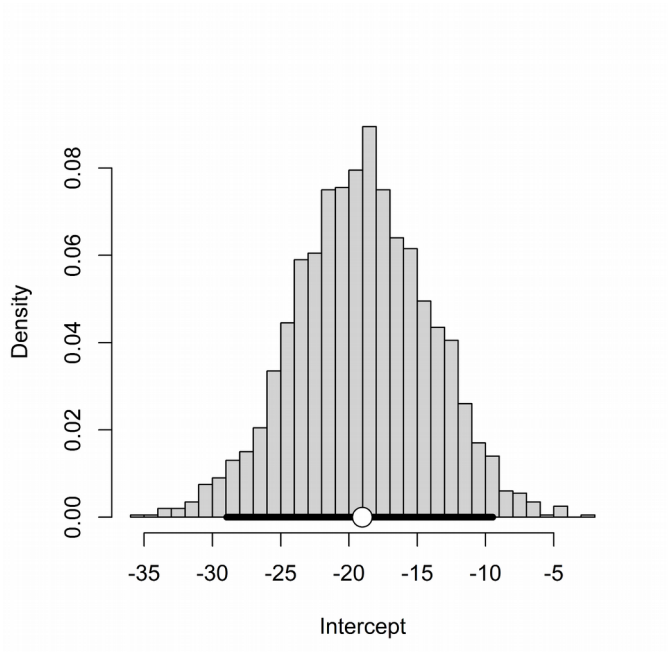


Figure 3. Posterior distribution histogram for the marginal distribution of parameters from the Bayesian Poisson GLM.

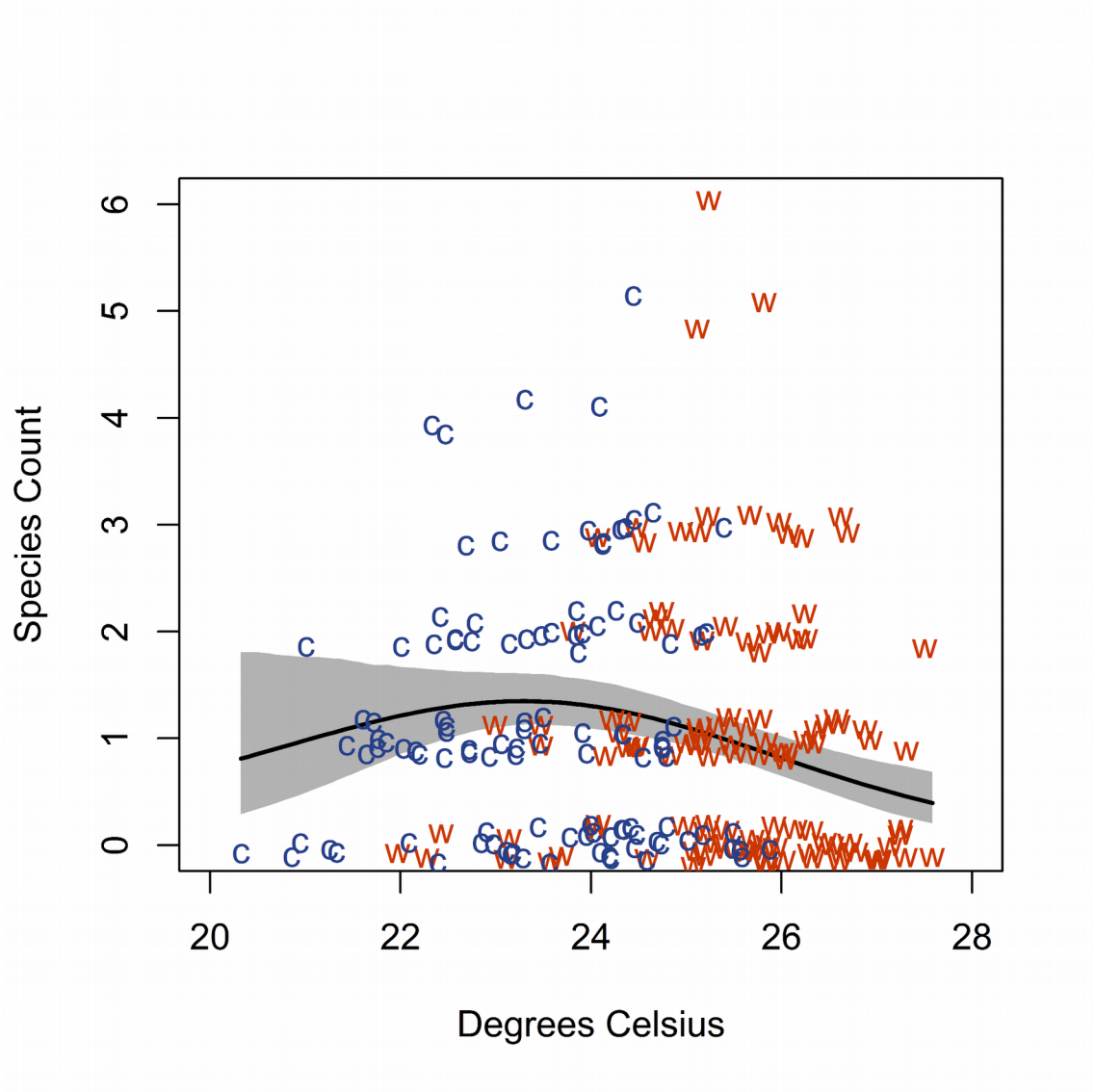


Figure 4. Covariate effect of orthogonal quadratic *SST* on cetacean species count in the cold (blue points) and warm (red points) periods. The continuous black line indicates the posterior means, and the shaded area indicates the 95% credibility intervals.

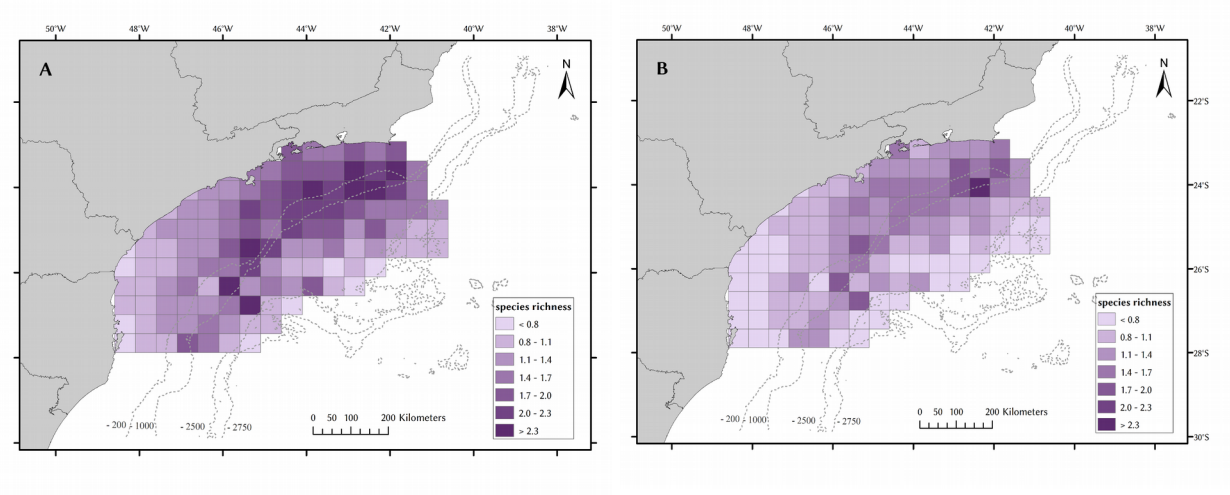


Figure 5. Posterior mean of the predictive distribution of cetacean species richness by cell in A) cold period and B) warm period, obtained by the Bayesian Poisson GLM.

3.3. Species richness prediction

Based on the mean prediction maps, the highest species richness cells were found in the northernmost part of the survey region within the Santos Basin, regardless of the survey period (Figure 5). In general, lower species richness was estimated for cells located far from the 1000 m isobath in both periods. Rugosity and slope, which are used as indicators of seabed morphology, are accentuated in the region around this isobath. Cetacean richness increased in the southeast region during the cold period. According to the diagnostic tools, the adjustment of the predictions was satisfactory (Supplementary Table B2 and Figure B3).

4. Discussion

4.1. Modelling issues

GLM and GAM are useful modelling tools for ecological data due to their flexibility in handling a variety of probability distributions for the response variables and their nonlinear relationships with covariates. In this application, the inclusion of

quadratic polynomials was sufficient to capture such nonlinearities in GLMs, and there was no need to advance further into GAMs.

Variable selection as part of the broader goal of finding a "best model" in some (relative) sense is a challenge. Being aware of all steps described by Burnham and Anderson (2002), we proceeded carefully by combining scientific knowledge (to define a subset of plausible models), exploratory data analysis (to explore correlation among variables) and automatic stepwise procedures (forward and backwards) to reach our goal. We used AIC as the criterion to compare models, taking into account the discussions in Burnham and Anderson (2004) and Aho and Peterson (2014).

The log-effort is usually included as an offset into the linear component of the Poisson family in GLMs. This is equivalent to defining its coefficient equal to 1 a priori. A more flexible approach includes this coefficient among the parameters to be estimated. That is, log-effort merely becomes another covariate, and the model has one more parameter to estimate. We explored both possibilities and ultimately included log-effort as a covariate since a lower AIC favoured its inclusion.

The sampling units in our model describe a spatial lattice structure. Therefore, without the inclusion of covariates, the spatial structure would be needed to distinguish neighbouring points from distant ones. However, the inclusion of covariates can reduce or even eliminate the spatial correlation between two points. The way to evaluate this is by fitting the proposed model and evaluating its residuals spatially. If an autocorrelation is detected in the residuals of a model, this tells us that we forgot an important predictor or process or that the relationship is different from the one we modelled. This type of spatial correlation will disappear if the key predictors are included in the model (Korner-Nievergedt *et al*., 2015). In addition to the inclusion of key variables, it is important to emphasize that the sampling design, good area coverage and the appropriate size of the grid cells can also influence the autocorrelation; since this can vary depending on the spatial scales (Qi and Wu, 1996; Overmars *et al*., 2003). In our analyses, conditional on the selected covariates, no spatial correlation was apparent. Hence, we considered it unnecessary to pursue it further.

A full Bayesian model is composed of two components: the data likelihood and the prior distribution for the model parameters. Special attention to the priors is recommended in two situations. First, when there is some critical information previously available, which should be respected during inference and prediction. Second, when the complexity of the model, combined with a lack of informative data, makes the Bayesian inference unstable and dependent on prior specifications. Neither of these conditions applied to this study. Therefore, we used flat priors as the general default. Last, since posterior predictive diagnostics displayed an acceptable fit, we decided not to explore this line of inquiry further. A posterior distribution can be useful as an a priori distribution for subsequent studies, so these can become increasingly accurate.

4.2. Ecological aspects

Although seasonal and interannual patterns tend to be complex, cetacean species-habitat relationships appear to remain relatively stable over time, and distribution patterns covary with preferred habitat patterns for various species (Ballance *et al*., 2006). In our study, we detected patterns that are consistent in cetacean species richness in the Santos Basin. Our results indicate that the main predictors for the spatial distribution of cetacean species richness were distance to the 1000 m isobath, latitude, rugosity, slope and temperature. That is, species richness was highest near the 1000 m isobath, in regions with abrupt topography and in the northern portion of the study area. Additionally, temperature indicated a seasonality component within this distribution, especially considering the enhanced richness in the southern portion during the cold season.

Spatially, it is expected that cetacean species richness increases with proximity to the shelf break and slope (Davis *et al*., 1998; Di Tullio *et al*., 2016). This is because the subsurface processes that enhance productivity predominate in this region, such as the shelf-slope fronts in the South Brazil Bight, that are year-round, but are best defined from April through September (Belkin and Cornillon, 2007). Furthermore, in this area, there is a depth gradient that may influence a greater diversity of prey or these

bathymetric features facilitate the capture of prey, such as the two submarine canyons that exist in the shelf-break of the study area (see Moors-Murphy, 2014). Our study emphasizes the importance of the outer continental shelf and slope to a diverse community of cetaceans found in the western South Atlantic Ocean; this study also provides information on species richness at the other depths within the study area, thereby producing a more complete overview. In the northern portion, even in regions closest to the coast, it is possible to visualize areas with higher average species richness. The Cabo Frio upwelling system occurs in this region mainly in the hot seasons induced by persistent northern winds, which, together with the narrower and deeper continental shelf, results in SACW dominance and limits the influence of coastal waters. The primary productivity linked to upwelling of SACW in the northern portion of our study area (Cabo Frio region) and along shelf-break results in greater species richness.

Temperature is by far the most common descriptor in distribution models and is among the most important determinants of general species distributions (Melo-Merino *et al*., 2020). Regarding cetacean species richness distribution, temperature can predict seasonal and geographic changes (Whitehead *et al*., 2008). When comparing only the absolute number of sighted species during the warm and cold seasons, important differences in the number of species were not apparent. However, when using our prediction maps to visualize species richness patterns for the entire area, there was a general increase in species richness throughout the area during the cold period, with the most pronounced increase occurring in the southernmost region. This general increase is partially explained by the presence of migratory species during the cold period. Furthermore, the hydrological characteristics of this portion of the SB can change in the cold period due to a stronger influence of the Malvinas Current and inputs of nutrient-rich waters from continental sources such as the La Plata river and Pato's Lagoon (Möller *et al.*, 2018; Marta-Almeida *et al*., 2021). In the warm period, the area is dominated by oligotrophic tropical waters, therefore a dominance of tropical species is expected. During the cold period, both tropical and cold-water species are likely to form the cetacean community in the study area (DiTullio *et al*., 2016). Stranding data showed greater cetacean species richness in the Aug-Sep months, which Prado *et al*. (2022)

suggest being an effect of a higher occurrence/abundance of cetaceans at the region driven by productive waters.

In general, including both environmental and spatiotemporal covariates improve the accuracy of the models in terms of how they estimate spatial patterns and quantify underlying species–environment relationships (Brodie *et al*., 2020). While patterns can be detected at larger scales, the temporal scales in dynamic covariates can influence the model accuracy, especially for highly mobile species (Scales *et al*., 2017). When cloud obstruction leads to significant information loss in remotely sensed data, as in our study, weekly or monthly environmental data averages may therefore be preferable (Scales *et al*., 2017). However, it is possible that the ecological influence of certain dynamic ecological variables, such as ocean fronts or chlorophyll concentrations, could not be detected at these resolutions. Nevertheless, by aggregating all species, as in our study, patterns among individual species are suppressed, and since each marine species has biological and ecological constraints, only consistent and general patterns are observed.

The present work increases the number of cetacean species observed during large surveys in the SB to 24 (see Zerbini *et al*., 2004; Di Tullio *et al*., 2016). Noteworthy are new records of *Balaenoptera borealis*, *Delphinus delphis*, *Lagenodelphis hosei*, *Peponocephala electra*, and *Kogia sima*. Our study area is traditionally an important region for cetacean diversity, where several species have been studied, including large whales (Williamson, 1975; Lodi *et al*., 2015; Santos *et al*., 2020; Rossi-Santos *et al*., 2022) and delphinids (Siciliano *et al*., 2007; Azevedo *et al*., 2007; Melo *et al*., 2010). Regarding conservation status, most of the species of the Santos Basin were recently evaluated by ICMBio (2022). Four species are classified as under some degree of threat. *Balaenoptera physalus* and *B. borealis* are listed as “endangered” and *Physeter macrocephalus* and *Sotalia guianensis* are listed as “vulnerable”. Eight species have a “data deficient” status, *Megaptera novaeangliae* is “near threatened”, and the remaining were classified as being of “least concern”.

The deepest areas, commonly used by sperm whales and the two endangered baleen whales recorded in our study, as well the northernmost portion, widely used by humpback whales, were highlighted above as the richest areas and are therefore

important for conservation. In addition, the coastal region, although having relatively lower species richness, faces higher anthropogenic impacts and is used by species of special conservation concern, such as *S. guianensis*, sighted during the cold period in our study. The Franciscana *Pontoporia blainvillei*, which was not detected in our ship sampling but occurs in coastal regions similar to *S. guianensis* (Sucunza *et al*., 2020; Prado *et al*., 2022), is critically endangered. As pointed out by Orme *et al*. (2005), hotspots of species richness, species threat and endemism typically do not share the same geographical distribution. Information about these three aspects regarding their utility as conservation tools varies (Orme *et al.*, 2005). As has been the case for other taxa, it has been suggested that hotspots of cetacean species richness indicate regions that deserve increased protection (Ballance, 2018). Marine mammals are currently among the most studied taxa within species distribution models (Melo-Merino *et al*., 2020) and are considered sentinels for ecosystem variability. The outputs of species distribution models can be used to support management decisions and conservation measures in a marine spatial planning context. Furthermore, combining models with layers of human footprint can add value to the use of models for conservation purposes (Marshall *et al*., 2014).

Data availability

The data that support this study are property of Petrobras and the summary of these, allowed for sharing, are available at the address https://sispmcprd.petrobras.com.br/sispmc.

Authors' contribution

ACM: conceptualization, methodology, formal analysis, visualization, investigation, writing—original draft, writing—review & editing; PGK: conceptualization, methodology, formal analysis, visualization, investigation, writing—original draft, writing—review & editing and supervision; LLW: writing—review and editing,

resources, project administration; LDR: conceptualization, methodology, formal analysis, investigation, writing—review & editing and supervision.

Acknowledgements

We thank Petrobras for kindly authorizing the use of data collected during the Projeto de Monitoramento de Cetáceos na Bacia de Santos (PMC-BS). The PMC-BS is one of the monitoring programs required by Brazil's federal environmental agency, IBAMA, for the environmental licensing process of oil production and transport by Petrobras in the Santos Basin presalt province (process no. 02001.114279/2017-80, ACCTMB no. 657/2015). We also thank all marine mammal observers who helped to gather data during PMC-BS; the crew from Sea Route/Samburá, especially captain Fábio Costa; and the Socioambiental Consultores Associados 'ground' team, especially José Olimpio da Silva Jr. and Aline Gonçalves. We appreciate the valuable comments of L. Bugoni and M. Camargo on an earlier version of the manuscript. The first author is grateful to the Coordination for the Improvement of Higher Education Personnel (CAPES), which provided financial support during her PhD. This research was developed under the guidance of the second and last authors. The Brazilian National Research Council (CNPq) provided a research fellowship to L. Dalla Rosa (PQ 309548/2019-5).

References

Aho, K., Derryberry, D.W., Peterson, T. 2014. Model selection for ecologists: The worldviews of AIC and BIC. Ecology 95(3), 631-636.

Akaike, H. 1973. Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika 60(2), 255-265.

Azevedo, A.F., Oliveira, A.M., Viana, S.C., van Sluys, M. 2007. Habitat use by marine tucuxis (Sotalia guianensis)(Cetacea: Delphinidae) in Guanabara Bay, south-eastern Brazil. J. Mar. Biolog. Assoc. U.K. 87(1), 201-205.

Ballance, L.T., Pitman, R.L., Fiedler, P.C. 2006. Oceanographic influences on seabirds and cetaceans of the eastern tropical Pacific: A review. Prog. Oceanog. 69(2-4), 360-390.

Ballance, L.T. 2018. Cetacean ecology. In: Wursig, B., Thewissen, J.G.M., Kovacs, K. (Eds.), Encyclopedia of marine mammals, third ed. Academic Press.

Belkin, I.M., Cornillon, P.C. 2007. Fronts in the world ocean's large marine ecosystems. ICES CM 500(130), 21.

Blangiardo, M., Cameletti, M. 2015. Spatial and spatio-temporal Bayesian models with R-INLA, first ed. John Wiley & Sons, Chichester.

Brandini, F.P., Tura, P.M., Santos, P.P.G.M. 2018. Ecosystem responses to biogeochemical fronts in the South Brazil Bight. Prog. Oceanogr. 164, 52-62.

Brodie, S.J., Thorson, J.T., Carroll, G., Hazen, E.L., Bograd, S., Haltuch, M.A., Holsman, K.K., Kotwicki, S., Samhouri, J.F., Willis-Norton, E., Selden, R.L. 2020. Trade-offs in covariate selection for species distribution models: a methodological comparison. Ecography 43(1), 11-24.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L. 2004. Advanced distance sampling: estimating abundance of biological populations. Oxford, Oxford University Press.

Burnham, K.P., Anderson, D.R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach, second ed. Springer-Verlag, New York.

Burnham, K.P., Anderson, D.R. 2004. Multimodel inference: Understanding AIC and BIC in model selection. Sociol. Methods Res. 33(2), 261-304.

Campos, E.J.D., Miller, J.L., Muller, T.J., Peterson, R.G. 1995. Physical oceanography of the Southwest Atlantic Ocean. Oceanography 8, 87–91.

Cañadas, A., Sagarminaga, R., Garcıa-Tiscar, S. 2002. Cetacean distribution related with depth and slope in the Mediterranean waters off southern Spain. Deep-Sea Res. I: Oceanogr. Res. Pap. 49(11), 2053-2073.

Committee on Taxonomy. 2021. List of marine mammal species and subspecies. Society for Marine Mammalogy, www.marinemammalscience.org.

Dalla Rosa, L., Ford, J.K., Trites, A.W. 2012. Distribution and relative abundance of humpback whales in relation to environmental variables in coastal British Columbia and adjacent waters. Cont. Shelf Res. 36, 89-104.

Dalpaz, L., Paro, A.D., Daura-Jorge, F.G., Rossi-Santos, M.R., Norris, T.F., Ingram, S.N., Wedekin, L.L. 2021. Better together: analysis of integrated acoustic and visual methods when surveying a cetacean community. Mar. Ecol. Prog. Ser. 678, 197–209.

Davis, R.W., Fargion, G.S., May, N., Leming, T.D., Baumgartner, M., Evans, W.E., Hansen, L.J., Mullin, K. 1998. Physical habitat of cetaceans along the continental slope in the northcentral and western Gulf of Mexico. Mar. Mamm. Sci. 14(3), 490-507.

Davis, R.W., Ortega-Ortiz, J.G., Ribic, C.A., Evans, W.E., Biggs, D.C., Ressler, P.H., Cady, R.B., Leben, R.R., Mullin, K.D., Wursig, B. 2002. Cetacean habitat in the northern oceanic Gulf of Mexico. Deep-Sea Res. I: Oceanogr. Res. Pap. 49(1), 121-142.

Dey, S.P., Vichi, M., Fearon, G., Seyboth, E., Findlay, K.P., Meynecke, J.O., De Bie, J., Lee, S.B., Samanta, S., Barraqueta, J.M., Roychoudhury, A.N., Mackey, B. 2021. Oceanographic anomalies coinciding with humpback whale super-group occurrences in the Southern Benguela. Sci. Rep. 11(1), 20896.

Di Tullio, J.C., Gandra, T.B., Zerbini, A.N., Secchi, E.R. 2016. Diversity and distribution patterns of cetaceans in the subtropical southwestern Atlantic outer continental shelf and slope. PLoS One 11(5), e0155841.

Diggle, P.J., Ribeiro, J.P.J., 2007. Model-based Geostatistics, first ed. Springer, New York.

Dorazio, R.M., Royle, J.A., Söderström, B., Glimskär, A. 2006. Estimating species richness and accumulation by modeling species occurrence and detectability. Ecology 87(4), 842-854.

Ekau, W., Knoppers, B. 2003. A review and redefinition of the large marine ecosystems of Brazil. In: Sherman, K., Hempel, G. (Eds.). Large Marine Ecosystems of the World: Trends in Exploitation, Protection and Research. Elsevier Science, Amsterdam.

ESRI (Environmental Systems Research Institute, Inc.). 2018. ArcGIS [GIS software]. Version 10.6.1. Redlands, CA.

Franklin, J., 2010. Mapping Species Distributions: Spatial Inference and Prediction, first ed. Cambridge University Press, New York.

Gelman, A., Hill, J. 2006. Data analysis using regression and multilevel/hierarchical models, first ed. Cambridge University Press, Cambridge.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B. 2013. Bayesian Data Analysis, third ed. Chapman and Hall/CRC, Boca Raton.

Gelman, A., Vehtari, A., Simpson, D., Margossian, C.C., Carpenter, B., Yao, Y., Kennedy, L., Gabry, J., Burkner, P.C., Modrák, M. 2020. Bayesian workflow. arXiv preprint arXiv:2011.01808.

Gelman, A., Su, Y.S., Yajima, M., Hill, J., Pittau, M.G., Kerman, J., Zheng, T., Dorie, V. 2021. arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.12-2. http://CRAN.R-project.org/package=arm/

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8, 993–1009.

ICMBio (Instituto Chico Mendes de Conservação da Biodiversidade). 2018. Livro vermelho da fauna Brasileira ameaçada de extinção. Brasília: ICMBio/MMA.

ICMBio (Instituto Chico Mendes de Conservação da Biodiversidade). 2020. Guia Ilustrado

de Identificação de Cetáceos e Sirênios do Brasil. Brasília: ICMBio/CMA.

Kaschner, K., Tittensor, D.P., Ready, J., Gerrodette, T., Worm, B. 2011. Current and Future Patterns of Global Marine Mammal Biodiversity. PLoS ONE 6(5), e19653.

Kéry, M. 2010. Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses, first ed. Academic Press, Burlington.

Kinas, P.G., Andrade, H.A. 2007. Bayesian statistics for fishery stock assessment and management: a synthesis. Pan-Am. J. Aquat. Sci. 2(2), 103-112.

Kinas, P.G., Andrade, H.A. 2020. Introdução à análise Bayesiana (com R), second ed. Consultor Editorial, Porto Alegre.

Korner-Nievergelt, F., Roth, T., Von Felten, S., Guélat, J., Almasi, B., Korner-Nievergelt, P. 2015. Bayesian data analysis in ecology using linear models with R, BUGS, and Stan, first ed. Academic Press, London.

Lodi, L., Borobia, M. 2013. Baleias, botos e golfinhos do Brasil: guia de identificação, first ed. Technical Books Editora, Rio de Janeiro.

Lodi, L., Tardin, R.H., Hetzel, B., Maciel, I.S., Figueiredo, L.D., Simão, S.M. 2015. Bryde's whale (Cetartiodactyla:Balaenopteridae) occurrence and movements in coastal areas of southeastern Brazil. Zoologia (Curitiba) 32, 171-175.

Marshall, C.E., Glegg, G.A., Howell, K.L., 2014. Species distribution modelling to support marine conservation planning: the next steps. Mar. Policy 45, 330–332.

Marta-Almeida, M., Dalbosco, A., Franco, D., Ruiz-Villareal, M. 2021. Dynamics of river plumes in the South Brazilian Bight and South Brazil. Ocean Dyn. 71, 59-80.

Melo, C.L.C., Santos, R.A., Bassoi, M., Araújo, A.C. 2010. Feeding habitats of delphinids (Mammalia:Cetacea) from Rio de Janeiro State, Brazil. J. Mar. Biolog. Assoc. U.K. 90(8), 1509-1515.

Melo-Merino, S.M., Reyes-Bonilla, H., Lira-Noriega, A. 2020. Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. Ecol. Model. 415, 108837.

MMA (Ministério do Meio Ambiente do Brasil). 2018. Zona Costeira e Marinha: Áreas e ações prioritárias para conservação da Biodiversidade Brasileira. Brasília: Ministério do Meio Ambiente.

Möller, O.O., Piola, A.R., Freitas, A.C., Campos, E.J.D. 2008. The effects of river discharge and seasonal winds on the shelf off southeastern South America. Cont. Shelf Res. 28, 1607-1624.

Moors-Murphy, H.B. 2014. Submarine canyons as important habitat for cetaceans, with special reference to the Gully: a review. Deep-Sea Res. II: Top. Stud. Oceanogr. 104, 6-19.

Nunes, G.T., Bugoni, L. 2018. Local adaptation drives population isolation in a tropical seabird. J. Biogeogr. 45(2), 332-341.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., …, Owens, I.P. 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436(7053), 1016-1019.

Overmars, K.D., De Koning, G.H.J., Veldkamp, A. 2003. Spatial autocorrelation in multi-scale land use models. Ecol. Model. 164(2-3), 257-270.

Pebesma, E.J., Bivand, R.S. 2005. Classes and methods for spatial data in R. R News 5 (2), https://cran.r-project.org/doc/Rnews/.

Pereira, R.C., Gomes, A.S. 2009. Biologia Marinha. Editora Interciência, Rio de Janeiro.

Prado, J., Daudt, N., Perez, M., Castilho, P., Monteiro, D. 2022. Intensive and wide-ranging beach surveys uncover temporal and spatial stranding patterns of marine megafauna. ICES J. Mar. Sci. 0, 1-15.

Qi, Y., Wu, J. 1996. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. Landsc. Ecol. 11(1), 39-49.

R Core Team, 2021. R: a language and environment for statistical computing. http://.R-project.org/

Redfern, J.V., et al. (19 authors). 2006. Techniques for cetacean–habitat modeling. Mar. Ecol. Prog. Ser. 310, 271–295.

Ribeiro Jr, P.J., Diggle, P.J. 2020. geoR: Analysis of Geostatistical Data. R package version 1.7-5.2. http://CRAN.R-project.org/package=geoR/

Roberts, J.J., Best, B.D., Dunn, D.C., Treml, E.A., Halpin, P.N. 2010. Marine Geospatial Ecology Tools: an integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. Environ. Model. Softw. 25, 1197–1207.

Rodrigues, L.D.S., Kinas, P.G., Cardoso, L.G. 2022. Optimal setting time and season increase the target and reduce the incidental catch in longline fisheries: a Bayesian beta mixed regression approach. ICES J. Mar. Sci. 79, 1245–1258.

Rossi-Santos, M.R., Filun, D., Soares-Filho, W., Paro, A.D., Wedekin, L.L. 2022. “Playing the beat”: Occurrence of Bio-duck calls in Santos Basin (Brazil) reveals a complex acoustic behaviour for the Antarctic minke whale (Balaenoptera bonaerensis). PLoS One 17(9), e0255868.

Royle, J.A., Kéry, M. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 88(7), 1813-1823.

Rufener, M.C., Kinas, P.G., Nóbrega, M.F., Oliveira, J.E.L. 2017. Bayesian spatial predictive models for data-poor fisheries. Ecol. Model. 348, 125-134.

Santos, M.C. de O., Siciliano, S., Souza, S.P., Pizzorno, J.L.A. 2020. Occurrence of southern right whales (Eubalaena australis) along southeastern Brazil. J. Cetacean Res. Manage. 2, 153-156.

Scales, K.L., Hazen, E.L., Jacox, M.G., Edwards, C.A., Boustany, A.M., Oliver, M.J., Bograd, S.J. 2017. Scale of inference: on the sensitivity of habitat models for wide-ranging marine predators to the resolution of environmental data. Ecography 40(1), 210-220.

Schwarz, G. 1978. Estimating the dimension of a model. Ann. Stat. 6(2), 461-464.

Siciliano, S., Ramos, R.M.A., Di Beneditto, A.P.M., Santos, M.C.O., Fragoso, A.B., Brito, J.L., Azevedo, A.F., Vicente, A.F.C., Zampirolli, E., Alvarenga, F.S., Barbosa, L., Lima, N.R.W. (2007). Age and growth of some delphinids in south-eastern Brazil. J. Mar. Biolog. Assoc. U.K. 87(1), 293-303.

Soberón, J., Nakamura, M. 2009. Niches and distributional areas: concepts, methods, and assumptions. Proc. Natl. Acad. Sci. U.S.A. 106(Supplement 2), 19644-19650.

Stachowitsch M., Rose N.A., Parsons E.C.M. 2019. State of the cetacean environment report (SOCER). IWC Global Compendium 2014-2018.

Sucunza, F., Danilewicz, D., Andriolo, A., Azevedo, A., Secchi, E., Zerbini, A. 2020. Distribution, habitat use, and abundance of the endangered franciscana in southeastern and southern Brazil. Mar. Mammal Sci. 36(2), 421-435.

Szewczyk, T.M., McCain, C.M. 2019. Disentangling elevational richness: a multi-scale hierarchical Bayesian occupancy model of Colorado ant communities. Ecography 42(5), 977-988.

Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Berghe, E.V., Worm, B. 2010. Global patterns and predictors of marine biodiversity across taxa. Nature 466(7310), 1098–1101.

Whitehead, H., McGill, B., Worm, B. 2008. Diversity of deep-water cetaceans in relation to temperature: implications for ocean warming. Ecol. Lett. 11, 1198-1207.

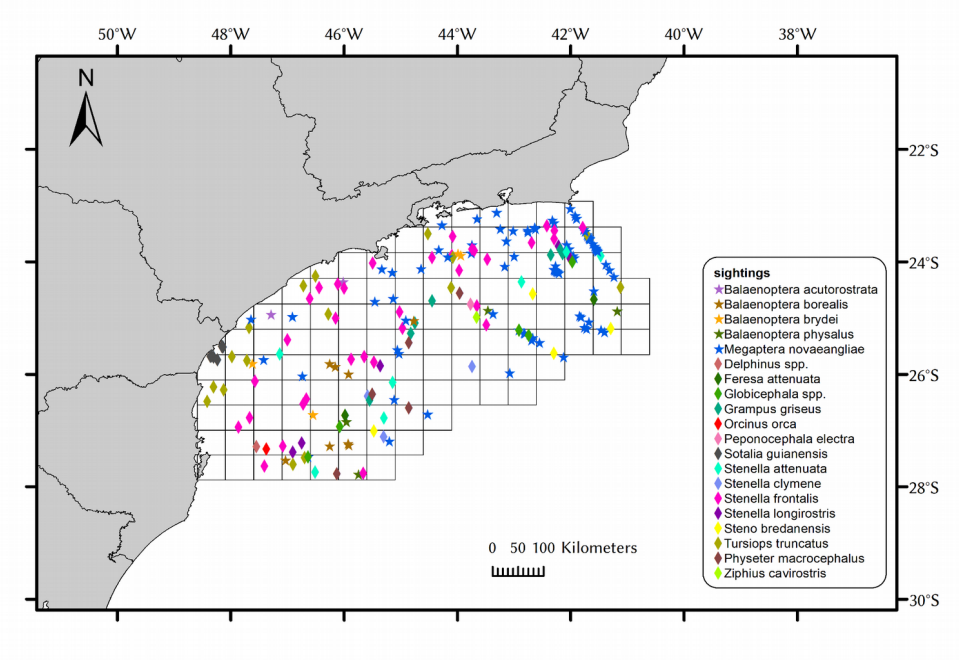
Williamson, G.R. 1975. Minke whales off Brazil. Sci. Rep. Whales Res. Inst. 27, 37-59.

Whitehead, H., O'Brien, K., Worm, B. 2010. Diversity of deep-water cetaceans and primary productivity. Mar. Ecol. Prog. Ser. 408, 1-5.

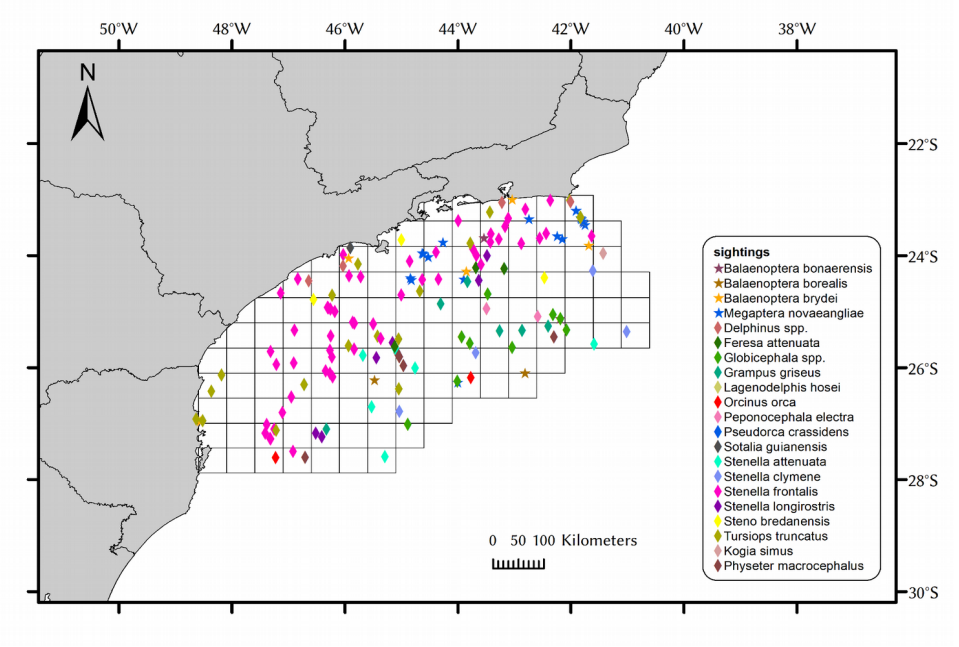
Zerbini, A.N., Secchi, E.R., Bassoi, M., Dalla Rosa, L., Higa, A., Sousa, L., Moreno, I.B., Moller, L.M., Caon, G. 2004. Distribução e abundância relativa de cetáceos na Zona Econômica Exclusiva da região sudeste-sul do Brasil. Série de Documentos Revizee-Score Sul. São Paulo: Instituto Oceanográfico – USP.

Zuur, A.F., Ieno, E.N., Smith, G.M. 2007. Analysing ecological data, first ed. Springer, New York.

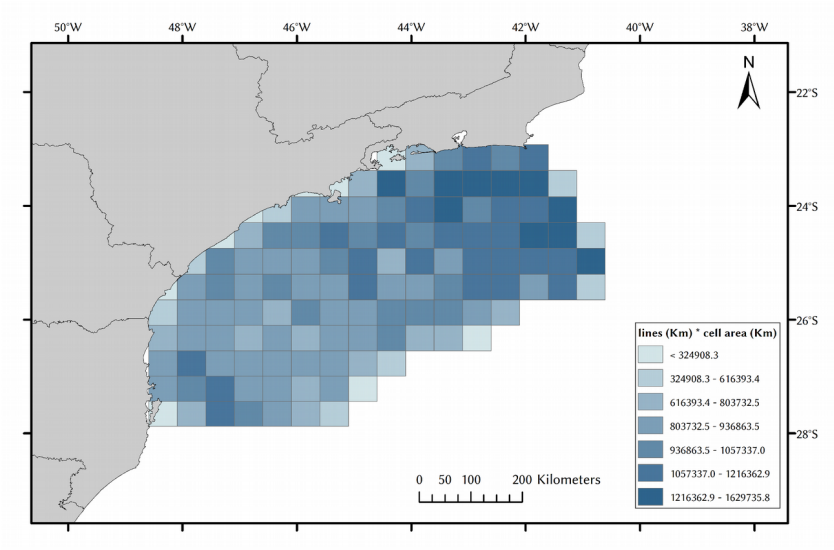
Supplementary Figures and Tables



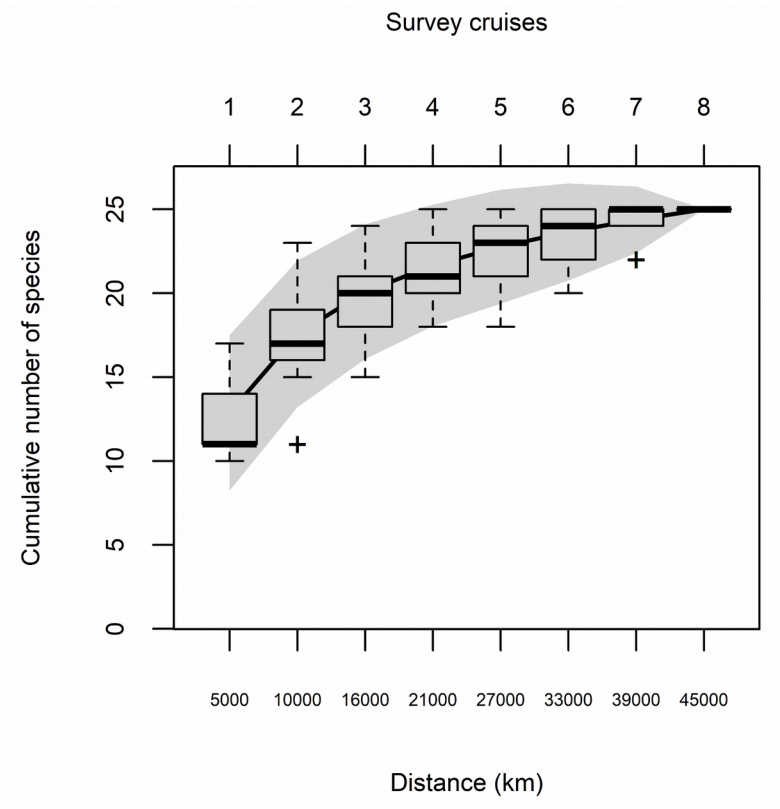
Supplementary Figure 1. Map of the gridded (0.5 x 0.5°) study area. Dots represent species sighting positions in the cold period (April – September).



Supplementary Figure 2. Map of the gridded (0.5 x 0.5°) study area. Dots represent species sighting positions in the warm period (October – March).



Supplementary Figure 3. Ship-based cetacean survey effort on the gridded study area. Grid cells are 0.5 x 0.5°, and darker shader indicates greater searching effort.



Supplementary Figure 4. Species accumulation curve for species richness data recorded according to distance travelled (lower x-axis) and the surveys (upper x-axis). Continuous black line indicate the mean, the shaded area indicate the confidence intervals, and the boxplots indicate the distribution of the estimated cumulative number of species.

Supplementary Table 1. Summary of cetacean species sighted per period.

Specie	Cold	Warm	Total
Megaptera novaeangliae	88	18	106
Stenella frontalis	34	55	89
Tursiops truncatus	17	19	36
Grampus griseus	9	6	15
Globicephala spp.	5	9	14
Stenella attenuata	7	7	14
Stenella longirostris	6	7	13
Physeter macrocephalus	5	6	11
Balaenoptera borealis	8	2	10
Balaenoptera brydei	4	6	10
Steno bredanensis	4	6	10
Sotalia guianensis	7	1	8
Stenella clymene	3	4	7
Delphinus spp.	1	4	5
Feresa attenuata	2	3	5
Balaenoptera physalus	4	-	4
Orcinus orca	1	2	3
Peponocephala electra	1	2	3
Balaenoptera acutorostrata	2	-	2
Balaenoptera bonaerensis	-	2	2
Lagenodelphis hosei	-	1	1
Pseudorca crassidens	-	1	1
Kogia simus	-	1	1
Ziphius cavirostris	1	-	1
Total	209	162	371
Total species	20	21	24

Supplementary Appendix A

Environmental predictors

To evaluate cetacean species richness in relation to habitat suitability, we used 17 potential predictors. Static physiographic covariates include depth, slope, distance to land, and distance to the 50 m, 200 m, and 1000 m isobaths. Dynamic environmental data include sea surface temperature, chlorophyll-a concentrations, sea surface salinity, sea surface height, sea level anomalies, zonal current velocity, meridional current velocity, and distance to SST fronts. The study area was divided into a 0.5° x 0.5° grid with each grid cell assigned the latitude and longitude at its midpoint. Environmental data were downloaded directly from the source using a bounding box whose extent covered the study area. Data was subsequently processed using ArcGIS 10.6.1. (ESRI, 2018) on an Albers equal area projection. A simple nearest-neighbour interpolation process was used for cells containing missing data.

Bathymetry data were obtained from the General Bathymetric Chart of the Oceans (GEBCO). The GEBCO_2019 gridded data product, which is available at https://www.gebco.net, provides global coverage on a 15 arc-second grid. We calculated the slope of sea depth (in degrees) using the *raster calculator* tool as an index of the heterogeneity of bathymetry. Rugosity was derived from the bathymetry layer using *focal statistics* and *raster calculator* tools based on Jenness (2004), which designates a position of a specific point relative to its local neighbourhood and whose values range from 0 to 1. This index was used to indicate seafloor complexity. The coastline, and the 50 m, 200 m, and 1000 m contour lines were originated from the bathymetry using contour tools. Euclidean distances were calculated between the midpoint of grid cells and line coordinates using the *spatial analyst* tool.

Chlorophyll-a concentration and sea surface temperature, which are commonly related to marine productivity, were extracted from the MODIS sensor of AQUA satellite monthly mean data provided by the Ocean Biology Processing Group (OBPG, at https://www.oceancolor.gsfc.nasa.gov). These images consisted of a Level-3 product at a ~9 km spatial resolution. Salinity, sea surface height, zonal and meridional current velocity data, which are linked to the chemistry and hydrodynamics of water masses,

were obtained as monthly averages from the HYbrid Coordinate Ocean Model (HYCOM, available at https://www.hycom.org). The project, part of the U.S. Global Ocean Data Assimilation Experiment (GODAE) provides global data that are interpolated to a uniform 1/12° grid, converted to a netcdf, and interpolated to 41 levels. Sea level anomaly data, which is the difference between measured sea surface height (SSH) and the mean SSH, were extracted from the Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO, available at https://www.aviso.altimetry.fr) program as monthly averages at a 0.25° resolution.

Fronts were identified by the single-image edge detection algorithm (Cayula and Cornillon, 1992) in the SST monthly raster images using the Marine Geospatial Ecology Tools (Roberts *et al*., 2010) in ArcGIS. Weak fronts with a mean temperature difference of less than 0.375° were not included. We then calculated the Euclidean distances between the midpoint of grid cells and the closest polyline front.

All environmental predictors were standardized to enable the coefficients of different variables to be comparable to each other and to reduce the correlation between model coefficients (Kinas and Andrade, 2020). Multicollinearity was checked by pair plots and Pearson's correlation index. Graphical inspections of the response variable and all environmental covariates were performed. Most explanatory variables were transformed according to the data characteristics to improve the fit (Zuur *et al*., 2007).

References

Cayula, J.F., Cornillon, P. 1992. Edge detection algorithm for SST images. J. Atmos. Ocean. Technol. 9,67–80.

ESRI (Environmental Systems Research Institute, Inc.). 2018. ArcGIS [GIS software]. Version 10.6.1. Redlands, CA.

Kinas, P.G., Andrade, H.A. 2007. Bayesian statistics for fishery stock assessment and management: a synthesis. Pan-Am. J. Aquat. Sci. 2(2), 103-112.

Jenness, J.S. 2004. Calculating landscape surface area from digital elevation models. Wild. Soc. Bull. 32, 829-839.

Roberts, J.J., Best, B.D., Dunn, D.C., Treml, E.A., Halpin, P.N. 2010. Marine Geospatial Ecology Tools: an integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. Environ. Model. Softw. 25, 1197–1207.

Zuur, A.F., Ieno, E.N., Smith, G.M. 2007. Analysing ecological data, first ed. Springer, New York.

Supplementary Appendix B

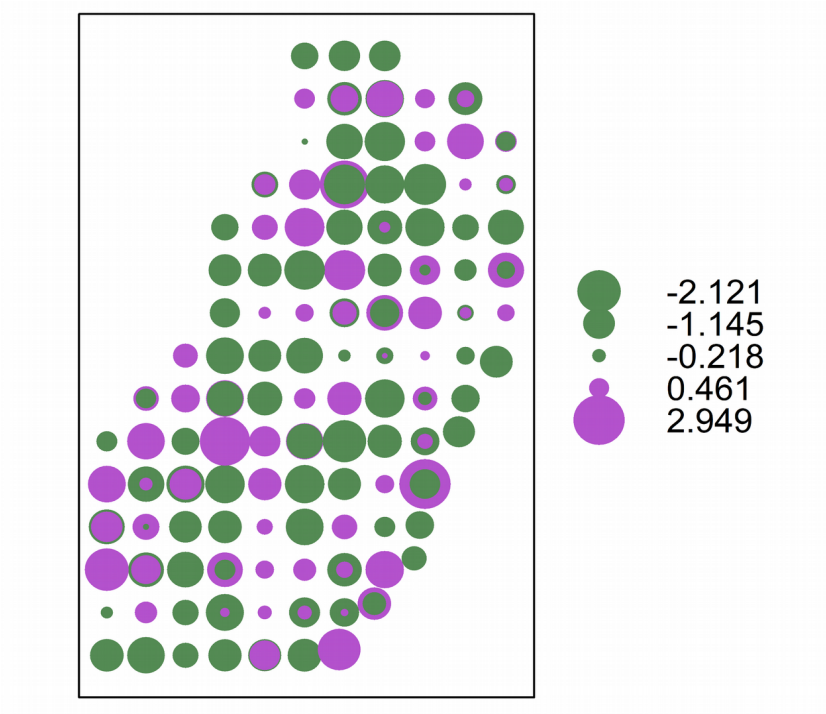
Stepwise procedure and model checking

To select the model that best fit the data, we performed a stepwise procedure with the R package stats (R Core Team, 2021), in conjunction with AIC criteria (Akaike, 1973), which provides an asymptotically efficient selection. Due to collinearity issues that may impact model fitting, the distance to the 200 m isobath and that to the 1000 m were tested separately in models B and C, respectively. Salinity was tested separately from sea surface height and chlorophyll-a in model D. The distance to the 50 m isobath variable was omitted from models because it was highly correlated with distance to the coast.

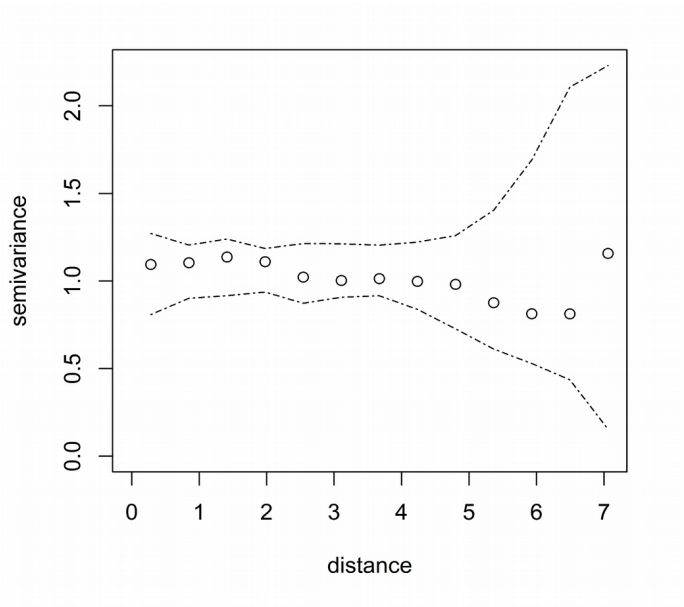
Table B1 summarizes the most relevant models selected by the stepwise procedure. In addition to AIC, we calculated the BIC criteria (Schwarz, 1978) for each selected model in steps A: D2 as an auxiliary measure to assess the consistency (Burnham and Anderson, 2004). According to the AIC and BIC, model C1 (Table B1) was the best. The period (cold/warm), which was included in the models as a factor, was not selected as a relevant predictor. Instead, *SST* seems to be a more explanatory variable of temporal patterns.

Supplementary Table B1. Summary of main candidate Poisson models selected by stepwise regression including Akaike information criterion (AIC) and Bayesian information criterion (BIC). Predictors acronyms are d_c = distance to coast, d_200 = distance to 200 m isobath, d_1000 = distance to 1000 m isobath, elev = elevation, lat = latitude, long = longitude, rug = rugosity, slo = slope, chl = chlorophyll-a concentration, d_f = distance to SST fronts, sla = sea level anomaly, ssh = sea surface height, sss = sea surface salinity, sst = sea surface temperature, vel_u = zonal current velocity, vel_v= meridional current velocity, eff = effort, per = period (warm/cold).

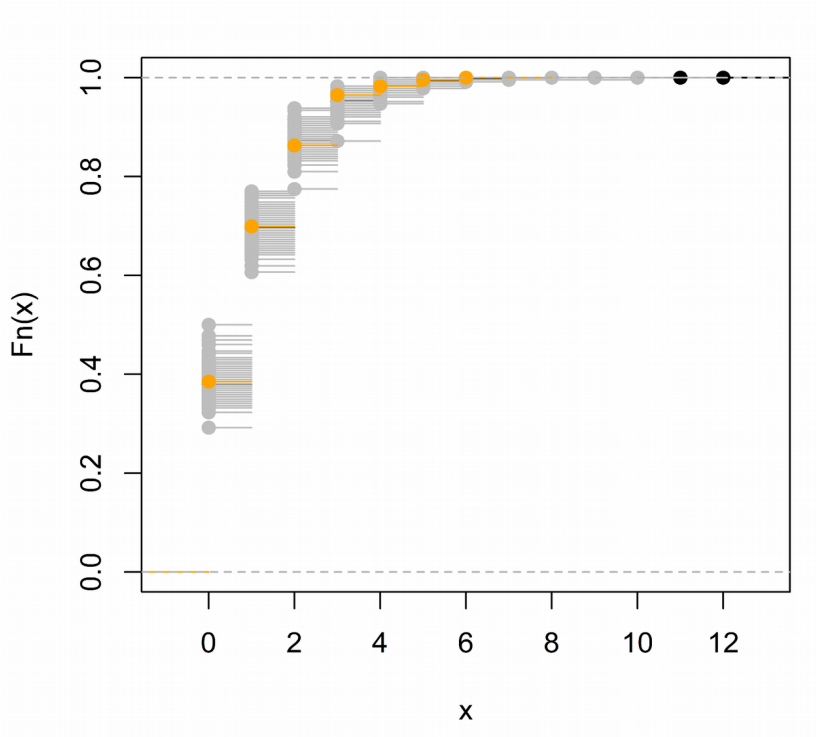
Model	Predictors	Selection	Selected variables	AIC	BIC
A	*null*	-	-	652.4	655.8
B1	sqrt(d_c), sqrt(d_200), elev, lat, long, rug, log(slo), log(chl), log(chl2), sqrt(d_f), sla, ssh, sst, sst2, vel_u, vel_v, log(eff), per	backward	log(eff) + sqrt(d_200) + sst + sst2 + lat + rug + long	605.7	633.2
B2		forward	log(eff) + sqrt(d_200) + sst + sst2 + lat + rug	604.1	629.7
C1	sqrt(d_c), sqrt(d_1000), elev, lat, long, rug, log(slo), log(chl), log(chl2), sqrt(d_f), sla, ssh, sst, sst2, vel_u, vel_v, log(eff), per	backward	**log(eff) + sqrt(d_1000) + sst + sst2+ lat + rug + log(slo)**	**601.4**	628.8
C2		forward	idem C1	601.4	628.8
D1	sqrt(d_c), sqrt(d_1000), elev, lat, long, rug, log(slo), log(chl2), sqrt(d_f), sss, sst, sst2, vel_u, vel_v, log(eff), per	backward	idem C1	601.4	628.8
D2		forward	idem C1	601.4	628.8



Supplementary Figure B1. Spatial distribution of residuals of the GLM Poisson model for species richness. Green points indicate positive and blue points indicate negative residuals. The size of the points corresponds to the absolute values of the residuals.



Supplementary Figure B2. Variogram with Monte Carlo envelopes of deviance residuals of the selected GLM Poisson model.



Supplementary Figure B3. Empirical cumulative distribution function (ECDF) plot comparing the distribution of species richness of simulated predictions from the model (grey) and of the observed data (orange).

Supplementary Table B2. Proportion of zero counts in observed and predicted values (adapted from Korner-Nievergelt *et al*. 2015).

	observed	predicted	
		Q0.01	Q0.99
propzeros	0.38	0.29	0.46

References

Akaike, H. 1973. Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika 60(2), 255-265.

Burnham, K.P., Anderson, D.R. 2004. Multimodel inference: Understanding AIC and BIC in model selection. Sociol. Methods Res. 33(2), 261-304.

R Core Team, 2021. R: a language and environment for statistical computing. http://.Rproject. org/

Schwarz, G. 1978. Estimating the dimension of a model. Ann. Stat. 6(2), 461-464.

CAPÍTULO III

Habitat-based BART models for cetaceans in the western South Atlantic: current and future distribution under climate change scenarios

(artigo publicado na revista Marine Environmental Research – IF 3.2, CAPES Qualis Biodiversidade A1)

DOI: https://doi.org/10.1016/j.marenvres.2026.107964

Habitat-based BART models for cetaceans in the western South Atlantic: current and future distribution under climate change scenarios

Ana Carolina Martins[a,b,*], Paul Gerhard Kinas[b], Leonardo Liberali Wedekin[c], Luciano Dalla Rosa[a,d]

[a] Programa de Pós-Graduação em Oceanografia Biológica, Universidade Federal do Rio Grande – FURG, 96203-900, Rio Grande, RS, Brazil

[b] Laboratório de Estatística Ambiental, Instituto de Matemática, Estatística e Física, Universidade Federal do Rio Grande – FURG, 96203-900, Rio Grande, RS, Brazil

[c] Socioambiental Consultores Associados, 88015-200, Florianópolis, SC, Brazil

[d] Instituto de Oceanografia, Universidade Federal do Rio Grande – FURG, 96203-900, Rio Grande, RS, Brazil

Corresponding author: tel: +55(53)984193893; e-mail: ana.carolbm@hotmail.com

Key-words: Bayesian Additive Regression Trees, community, occupancy model, Santos Basin, South Brazil Bight, species distribution modeling

Abstract

The Santos Basin, Southwestern Atlantic Ocean, is a region of high cetacean diversity. We analyzed the largest systematic ship-based cetacean sightings dataset for the area, collected between 2015 and 2019, to assess current distributions and potential responses to climate change for major taxonomic groups: migratory whales, Bryde's whales, small delphinids, large delphinids, and sperm and beaked whales. Bayesian Additive Regression Trees (BART) were employed to model species occurrence as a function of static and dynamic environmental covariates. Models demonstrated good predictive performance (AUC = 0.76–0.96; TSS = 0.40–0.86), highlighting the importance of depth, isobaths, rugosity, SST, and chlorophyll-a in driving habitat suitability. Predicted distributions indicate that migratory whales are concentrated in the northern SB and along the continental slope, whereas Bryde's whales occupy shallow coastal areas.

Small delphinids favor shallow waters and the outer shelf, whereas large delphinids, sperm and beaked whales occur mainly along the 1000-m isobath and in deeper offshore areas. Future projections under medium- (SSP2-4.5) and high-emission (SSP5-8.5) scenarios suggest a coastal shift and contraction of suitable habitat for migratory whales, a potential range expansion for Bryde's whales and small delphinids, and relative stability for large delphinids and sperm and beaked whales. Our results provide a robust, group-level assessment of cetacean habitat preferences and potential climate-driven distributional changes, thereby informing adaptive conservation and management strategies in the SB under future climate scenarios.

1. Introduction

Reliable assessments of species distributions are crucial across various fields, including ecological research and management. A key objective of modern ecological research is to refine methodologies that improve exploration, inference, and prediction in ecological systems while addressing model selection and associated uncertainties (Tredennick et al. 2021). By integrating robust statistical frameworks and transparent model selection processes, ecologists can better understand ecological patterns and processes, ultimately supporting more effective conservation and management strategies (Guisan and Thuiller 2005).

In recent years, more than two dozen statistical and machine learning methods have been utilized in species distribution modeling (SDM; Norberg et al. 2019, Melo-Merino et al. 2020, Rodrigues et al. 2023). These models link species occurrence data with environmental variables, striving to accurately represent the range of a species' ecological needs, uncover underlying ecological processes, and approximate its ecological niche (Guisan and Thuiller 2005; Redfern 2006). Many widely used SDM approaches, when considered individually, lack robust frameworks for addressing uncertainty. One common strategy has been the use of ensemble methods (Lobo et al. 2021, Tardin et al. 2025), combining predictions from multiple models to reduce reliance on any single algorithm. While this approach can improve predictive robustness, it can be time-consuming and computationally intensive. In contrast,

Bayesian methods remain underrepresented (Carlson 2020). They offer conceptual advantages, like adherence to the likelihood principle and the admissibility of Bayesian estimators, and represent a potentially more efficient solution for practitioners by consolidating predictive and inferential strengths within a single model (Gelman et al. 2020). Bayesian additive regression trees (BART; Chipman et al. 2010) exemplify this: BART is an ensemble of regression trees (Hill et al. 2020), integrating the conceptual familiarity and strengths of classification tree methods with a Bayesian formulation that explicitly accounts for model uncertainty. As a fully non-parametric alternative to traditional linear and additive regression models, tree-based models employ classification trees for categorical responses and regression trees for numeric ones. These models aim to simplify complex datasets by identifying the most effective empirical classifier. A potential advantage of tree-based models in studying cetacean-habitat relationships is their ability to naturally capture non-additive interactions among predictors, along with their ease of interpretation. Only predictors that effectively partition the data and explain variation in the response variable are included in the final model (Redfern 2006). Comparative studies have demonstrated that BART outperforms traditional methods in balancing calibration and discrimination, while also offering superior explanatory power (Baquero et al. 2021, Plant et al. 2021). Thus, BART emerges as a promising tool for broad applications in SDMs, combining practical efficiency with these conceptual strengths, and is only beginning to be applied in cetacean distribution research.

Cetaceans are currently among the most studied taxa in SDMs (Melo-Merino et al. 2020, Rodrigues et al. 2023). As predominantly higher-trophic-level predators, cetaceans play a vital role in maintaining ecosystem balance by regulating prey populations and contributing to nutrient cycling (Roman et al. 2014). These species can serve as valuable indicators of environmental health and mid- to long-term ecological changes, including responses to climate variability and change. However, their effectiveness as indicators depends on sufficient understanding of species distribution and behavior across space and time (Fossi et al. 2012, Hazen et al. 2019). Bathymetry and sea surface temperature (SST), the most commonly used environmental predictors in SDMs (Melo-Merino et al. 2020, Pasanisi et al. 2024), have documented importance

in determining cetacean occurrences, either directly or indirectly, by influencing prey availability (Forcada 2018). However, the effects of derived variables (e.g. slope, rugosity, and SST fronts) and their interactions with other factors remain less understood, though they are generally assumed to be closely linked to prey availability.

Despite cetaceans being a well-studied taxon in SDMs, knowledge of their distribution dynamics in the Southwestern Atlantic Ocean (SWAO) remains limited (Rodrigues et al. 2023, Pasanisi et al. 2024). One of the most studied regions along the Brazilian coast is the Santos Basin (SB) and the broader South Brazil Bight, where research has addressed various aspects of cetacean ecology and biology, including habitat use, acoustics, feeding habits, bioaccumulation of contaminants, and life history parameters (e.g. Moreno et al. 2005; Siciliano et al. 2007; Melo et al. 2010; Dorneles et al. 2013; Santos et al. 2019; Dalpaz et al. 2021; Pires et al. 2021; Correa et al. 2023; Lima et al. 2024). However, despite these efforts, important knowledge gaps persist regarding community-level distribution patterns in this area. Some studies have investigated cetacean distribution and habitat use in the SB, focusing on specific species, spatial zones, or time periods (e.g. Di Tullio et al. 2016; Maricato et al. 2022; Dalpaz et al. 2023; Tardin et al. 2025). Martins et al. 2023 have contributed to identifying key areas of species richness — particularly along the 1000 m isobath and regions of complex topography — and revealed the influence of environmental variables and migratory behavior on distribution patterns. However, most analyses have focused on a limited number of species or taxonomic groups and often rely on broad-scale or correlative approaches. In this context, our study introduces a novel approach by applying a multi-species framework using Bayesian Additive Regression Trees (BART), which allows for complex, non-linear relationships between predictors. We also incorporate future environmental projections to assess potential climate-induced shifts in cetacean distribution, offering finer spatial resolution and predictive insights beyond traditional distribution models.

Climate change has the potential to influence cetacean distribution, migration patterns, and overall ecosystem dynamics, posing challenges for future conservation efforts (Grose et al. 2020). Certain species may adapt to these changes or even thrive,

while others may face significant stress, losing suitable habitats and, thus, heightening their risk of extinction (van Weelden et al. 2021). Therefore, identifying species and populations most vulnerable to these environmental shifts is essential. Research quantifying the impacts of climate change on biodiversity heavily relies on species distribution modeling (Peterson et al. 2011). This modeling allows the use of high-resolution data layers that represent the future projections of physical, chemical, and biological conditions of the global ocean, considering different scenarios (Assis et al. 2024). The climate change scenarios discussed by the Intergovernmental Panel on Climate Change (IPCC 2021) are based on the "Shared Socioeconomic Pathways" (SSPs), which represent different future trajectories for greenhouse gas emissions, socioeconomic changes, and climate impacts. Among these, two scenarios are frequently highlighted in previous studies: the "middle-of-the-road development" (SSP 2-4.5), which envisions a balanced progression with moderate mitigation efforts, and the "fossil fuel-driven development" (SSP 5-8.5), which depicts a future heavily reliant on fossil fuels, resulting in more severe climate consequences. Within this framework, our objectives were (i) to model and compare current cetacean distribution patterns in the Santos Basin based on identifiable predictor variables, and (ii) to predict potential shifts in these assemblage's distribution ranges, under environmental changes derived from global warming scenarios.

2. Methods

2.1. Data acquisition and processing

2.1.1. Cetacean surveys

The Santos Basin (SB) was sampled from the shoreline to the boundary of the Brazilian Exclusive Economic Zone during eight systematic ship surveys conducted as part of the "Projeto de Monitoramento de Cetáceos da Bacia de Santos" (PMC-BS) project. These seasonal surveys covered approximately 44,000 km of zigzag-shaped transect lines between 2015 and 2019. Survey periods and routes are detailed in Table 1 and Fig. 1. The transect lines were intended to cover the SB in a systematic and homogeneous way, so that the cetacean counts could be extrapolated to the entire study

area. The transects were oriented perpendicular to the coast, crossing the depth contours of the study area, and were shifted 60-90 km between surveys in order to avoid spatial biases in the sampling. The vessel traveled along the transect lines at an average speed of 7 knots.

Table 1. Survey periods and summaries for search effort by cruise.

Cruise	Year	Period	Effort(h)	Effort (km)
1	2015 - 2016	20 Nov - 5 Jan	389.9	4,937.1
2	2016	10 May - 20 Jun	410.9	5,462.5
3	2016	27 Sep - 6 Nov	420.1	5,500.2
4	2017	8 Mar - 18 Apr	434.9	5,541.9
5	2017	25 Aug - 7 Oct	435.7	5,583.8
6	2018	12 Jan - 20 Feb	444.3	5,863.4
7	2018	1 Sep - 20 Oct	477.2	5,930.5
8	2019	13 Mar - 26 Apr	425.3	5,502.3
Total	-	-	3,438.3	44,321.7

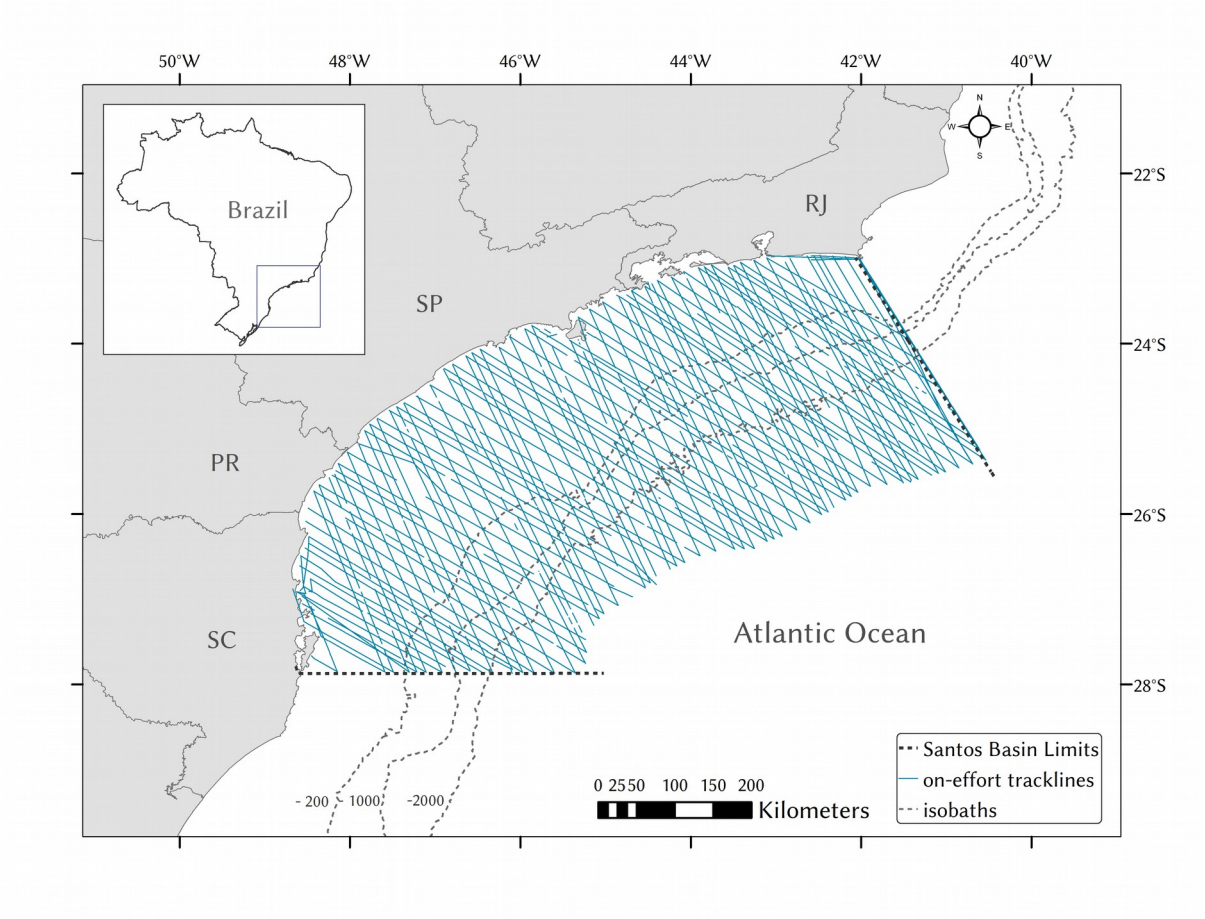


Figure 1. Study area and survey tracklines along the Santos Basin during 2015 - 2019. Brazilian states: Rio de Janeiro (RJ), São Paulo (SP), Paraná (PR) and Santa Catarina (SC). Bathymetric contours (isobaths) of 200 m, 1000 m, and 2000 m are shown.

2.1.2. Effort and sighting data

Searches were conducted during daylight hours, with three observers stationed at the vessel's highest point, approximately 7 m above sea level, following standard line-transect distance sampling protocols (Buckland et al. 2004; Dalla Rosa et al. 2012; Dalpaz et al. 2021). Observers rotated periodically with two others off effort. They scanned for cetacean groups using both the naked eye and binoculars (7x50 and 15x50 with digital stabilization). Observers on rest aimed to capture photographs to aid in species identification. Sampling effort was strictly regimented, and the time and location of all observations were recorded. For each detected cetacean group, the

horizontal angle relative to the trackline and the vertical angle to the horizon were measured using an angle board and reticulated binoculars, respectively. These measurements allowed for the cetacean group's geographic position (registered with a GPS) to be corrected incorporating the perpendicular distance from the transect, in accordance with the established distance sampling methodology (Buckland et al. 2004). Sighting data included time, location, species identification, group size and composition, presence of offspring, and observed behaviors.

2.1.3. Data preparation

Effort tracklines were divided into 10-km segments for analysis (Dalla Rosa et al. 2012). Five distinct binary presence/absence datasets were created for species groups based on the occurrence of 1) mysticetes and 2) odontocetes; further subdivided into: 1.1) migratory whales and 1.2) Bryde's whales (the only resident whale and feeding strictly in our study area); and 2.1) small delphinids, 2.2) large delphinids, and 2.3) sperm and beaked whales (see species composition in Table 2). Species were grouped according to their phylogeny and ecological-behavioral traits (Roberts et al. 2016). Each sighting recorded during on-effort survey periods was matched to the centroid of the nearest surveyed segment (Dalla Rosa et al. 2012). Sightings recorded outside of on-effort periods were excluded from the analysis. All geospatial data processing was done using ArcGIS 10.6.1 (ESRI, 2018) and the Marine Geospatial Ecology Tools toolbox (Roberts et al. 2010). Statistical analyses were performed in the R 4.3.0 (R Core Team 2023).

Table 2. Summary of sightings and number of individuals by taxonomic group and subgroup.

Group	Subgroup	Scientific name	Common name	N sightings	N individuals
Mysticeti	**Migratory whales**	*Balaenoptera acutorostrata*	Common minke whale	2	2
		Balaenoptera bonaerensis	Antarctic minke whale	2	3
		Balaenoptera borealis	Sei whale	10	28
		Balaenoptera physalus	Fin whale	4	9
		Megaptera novaeangliae	Humpback whale	106	215
	Total	-	-	124	257
Mysticeti	**Bryde's whales**	*Balaenoptera brydei*	Bryde’s whale	10	15
	Total	-	-	10	15
Odontoceti	**Small delphinids**	*Delphinus delphis*	Common dolphin	5	67
		Lagenodelphis hosei	Fraser’s dolphin	1	30
		Sotalia guianensis	Guiana dolphin	8	42
		Stenella sp.	Unidentified *Stenella*	10	122
		Stenella attenuata	Pantropical spotted dolphin	14	372
		Stenella clymene	Clymene dolphin	7	130
		Stenella frontalis	Atlantic spotted dolphin	89	2825
		Stenella longirostris	Spinner dolphin	13	1339
		Steno bredanensis	Rough-toothed dolphin	10	50
		Tursiops truncatus	Common bottlenose dolphin	36	847
	Total	-	-	193	5824
Odontoceti	**Large delphinids**	*Feresa attenuata*	Pygmy killer whale	5	60
		Globicephala macrorhynchus	Short-finned pilot whale	14	186
		Grampus griseus	Risso’s dolphin	15	313
		Orcinus orca	Killer whale	3	4
		Peponocephala electra	Melon-headed whale	3	620

		Pseudorca crassidens	False killer whale	1	15
	Total	-	-	41	1198
Odontoceti	**Sperm & beaked whales**	*Kogia sima*	Dwarf sperm whale	1	2
		Physeter macrocephalus	Sperm whale	11	25
		Ziphius cavirostris	Cuvier's beaked whale	1	4
	Total	-	-	13	31
		Total	-	**381**	7325

2.1.4 Environmental data

Habitat predictors included a suite of static physiographic data and dynamic environmental covariates and were obtained from remotely sensed satellite data and ocean model products (Table 3). These were chosen based on their ecological relevance for cetacean distribution (Cañadas et al. 2002; Dalla Rosa et al. 2012; Melo-Merino et al. 2020; Pasanisi et al. 2024). Prior to model fitting, collinearity among predictors was assessed using pairwise Pearson correlation coefficients. Variables showing high correlation ($|r| > 0.7$) were not included together in the same model, and redundant predictors were excluded from the initial candidate set to enhance ecological interpretability. Variables with coarser spatial resolution were resampled using Cubic Convolution Interpolation in ArcGIS 10.6.1 to match the initial resolution of other environmental covariates (0.05 degrees) and facilitate subsequent scale selection. Weekly environmental variables were matched to the coordinates corresponding to the center of each surveyed segment.

In addition to the current oceanographic layers, used in BART to generate sets of regression trees and to construct posterior distributions of estimated classification probabilities (see Section 2.2.1. below), a second set of layers representing climatological averages of dynamic variables was obtained for use in the prediction phase. For current predictions, decadal climatologies (2002 - 2022) were obtained from the same sources as the weekly layers (Table 3). For future projections, environmental layers were sourced from the Bio-ORACLE database version 3.0 (Assis et al. 2024),

which is based on a multi-model ensemble derived from CMIP6. Projections were selected for two future time periods - 2050 (2040–2050) and 2090 (2080–2090) - under two Shared Socioeconomic Pathway (SSP) scenarios: a medium (SSP2-4.5) and a high-range (SSP5-8.5) emission scenario (IPCC 2021).

Table 3. Candidate environmental covariates used in the Bayesian Additive Regression Trees (BART) models.

Predictor variable	**Unit**	**Temporal resolution**	**Acronym**	**Source**
Latitude	degree	static	lat	Field data
Longitude	degree	static	long	Field data
Depth	m	static	depth	GEBCO
Slope	degree	static	slope	GEBCO
Rugosity	index	static	rug	GEBCO
Distance to land	m	static	d_coast	GEBCO
Distance to 200 m isobath	m	static	d_200	GEBCO
Distance to 1000 m isobath	m	static	d_1000	GEBCO
Temperature	°C	weekly	sst	NASA
Chlorophyll-a	mg/m3	weekly	chl	NASA
Salinity	psu	weekly	sss	HYCOM
Sea Level Anomaly	m	weekly	sla	AVISO
Zonal current velocity	m/s	weekly	vel_u	HYCOM
Meridional current velocity	m/s	weekly	vel_v	HYCOM
Distance to SST fronts	m	weekly	dist_front	NASA

2.2. Species distribution modeling

2.2.1. Model development and evaluation

To evaluate the distribution of cetacean groups based on oceanographic predictors, BART models were run with default parameters of the *embarcadero* R package (Carlson 2020), using a binary response variable. The model employed 200 trees and 1000 MCMC iterations, retaining 800 iterations after discarding the first 20% as burn-in. This package includes an automated variable selection procedure, in which variable importance is measured by counting the number of times a given variable is used by a tree split across the full posterior draw of trees. The best set of explanatory variables is returned based on the model presenting the lowest average root-mean-square error (RMSE) (Chipman et al. 2010, Carlson 2020).

The prediction results can be classified into four categories: (1) true positives, when occurrences are correctly predicted; (2) true negatives, when non-occurrences are correctly predicted; (3) false positives, when non-occurrences are incorrectly predicted as occurrences; and (4) false negatives, when occurrences are incorrectly predicted as non-occurrences. Model performance was evaluated using the area under the curve (AUC) of the receiver operating characteristic curve (ROC) and the True Skill Statistics (TSS). Model discrimination was interpreted using commonly applied thresholds in species distribution modelling studies (Norberg et al. 2019), with AUC values >0.7 generally indicating acceptable discriminatory ability and higher values reflecting increasing predictive accuracy, and TSS values >0.4 indicating useful discriminatory performance. Once automated model selection and evaluation were completed, the selected models were used to generate spatially explicit predictions of cetacean occurrence probability across the study area. These predicted probabilities were interpreted as proxies for habitat suitability (Guisan and Thuiller 2005). In this context, areas with higher predicted probabilities are assumed to represent more suitable environmental conditions for the presence of the species or group, whereas lower values indicate less suitable habitats.

2.2.2. Projection of models into the future

To assess potential future shifts in the distribution ranges of cetacean groups under global climate change, the models were projected into two timeframes: the near future (2050) and the far future (2090), and in two SSP scenarios of greenhouse gas emission: a medium (SSP2-4.5) and a high-range (SSP5-8.5). The occurrence models were used to forecast spatial variation in the probability of occurrence for each species group in response to projected changes in climate.

3. Results

3. 1. Model selection and current occurrence prediction

The BART models demonstrated consistent discriminatory performance across all cetacean groups (Table 4). AUC values ranged from 0.76 to 0.96 and TSS from 0.40 to 0.86, indicating moderate to excellent predictive capacity.

Both environmental and spatial covariates contributed to model performance. Depth, distance to isobaths, and rugosity were common static predictors, whereas SST and chlorophyll-a were frequently selected dynamic covariates. Longitude, slope, latitude, sea level anomaly, salinity, and distance to SST fronts were also retained following the automated BART variable selection procedure, whereas current velocity variables were not supported by the selection process. The relative contribution of predictors for each cetacean group is presented in Table 4 and illustrated through partial dependence plots in Supplementary figures S1 - S5. These plots show the marginal influence of individual predictors on the model outcome, after accounting for the average effects of all other explanatory variables. Greater uncertainty is expected in regions with sparse data coverage, as reflected by wider credible intervals. Accordingly, such patterns should be interpreted with caution.

The current predicted distributions of migratory whales and large delphinids, based on the best-fit models A and F, are shown in Figure 2. Models B, D and G were selected as the best-fit models for Bryde's whales, small delphinids and sperm and beaked whales, respectively. Because environmental layers with future projections of

sea level anomaly were unavailable, and distance to SST fronts is not suitable for use with decadal climatological layers, Figure 2 presents the current predictions for Models C, E, and H, allowing direct comparison with future projections. Supplementary figures S6 - S10 show the 95% credible interval posterior width (a measure of spatial uncertainty), highlighting areas of highest uncertainty across all predictions.

Table 4. Summary of BART models selected through automated variable importance procedures, including the Area Under the Receiver Operating Characteristic Curve (AUC) and True Skill Statistics (TSS). Predictor acronyms are defined in Table 3.

Model	Group	Covariates	AUC	TSS	Cutoff
A	1.1. migratory whales	lat + long + d_coast + d_200 + sst	0.91	0.68	0.06
B	1.2. Bryde's whales	depth + slope + rug + d_coast + sst + chl + sla + sss	0.96	0.86	0.01
C	1.2. Bryde's whales	depth + slope + rug + d_coast + sst + chl + sss	0.95	0.83	0.01
D	2.1. small delphinids	depth + rug + d_1000 + sst + sla	0.76	0.40	0.06
E	2.1. small delphinids	depth + rug + d_1000 + sst	0.76	0.40	0.06
F	2.2. large delphinids	depth + d_1000 + chl	0.87	0.63	0.01
G	2.3. sperm and beaked whales	long + depth + slope + rug + d_1000 + chl + d_front	0.91	0.75	0.01
H	2.3. sperm and beaked whales	long + depth + slope + rug + d_1000 + chl	0.91	0.75	0.01

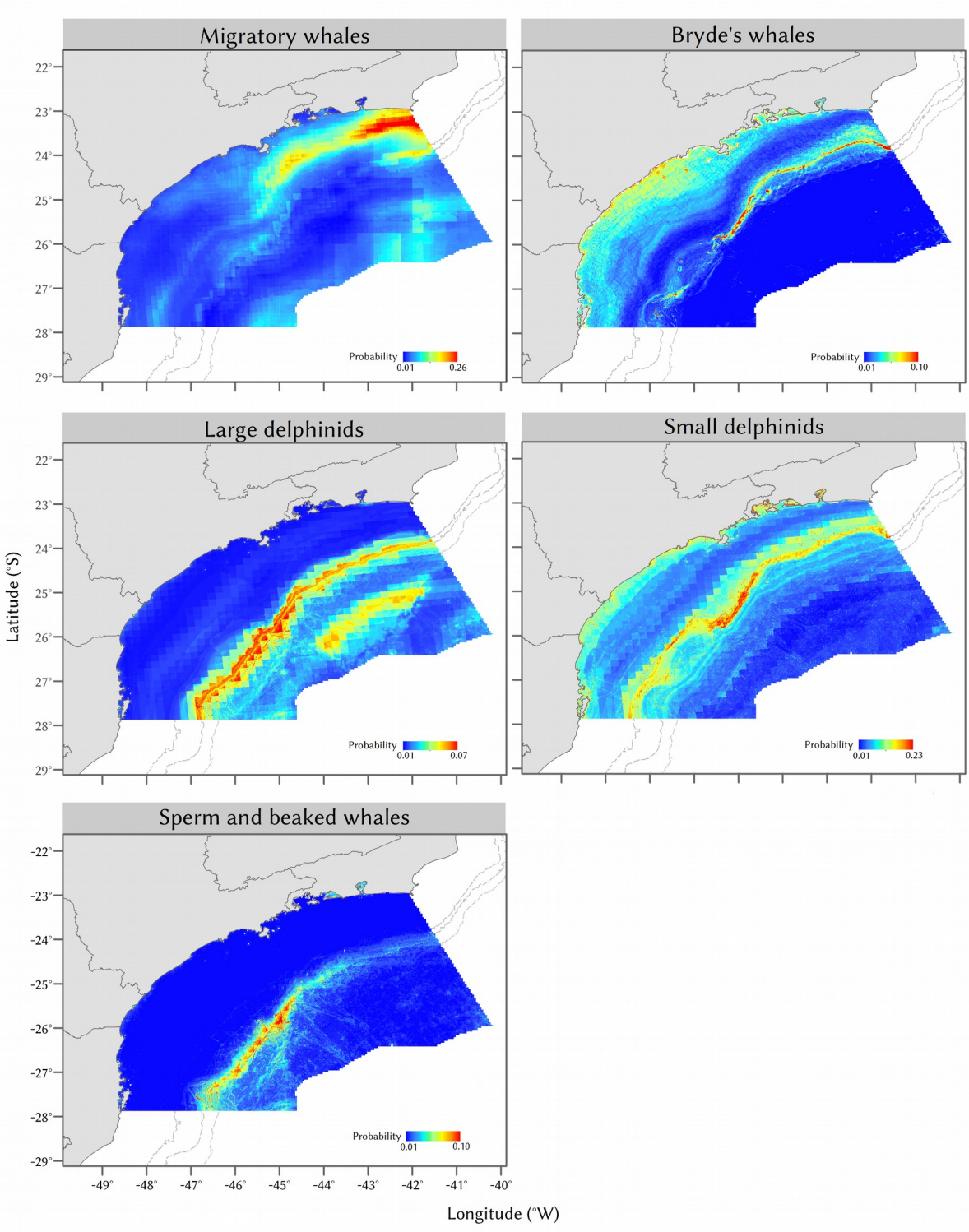


Figure 2. Current predicted distribution of cetacean groups generated by the BART model.

3.1.1. Mysticetes

For migratory whales, the model's predicted current distribution (Figure 2) indicates the highest probability of occurrence in the northern portion of the SB region, and along the continental slope region. Suitable areas are also observed in more coastal zones, particularly off the states of Rio de Janeiro (RJ) and São Paulo (SP); as well as in deeper offshore areas. Partial dependence plots (Supplementary Figure 1) show that occurrence probability peaks at latitudes below 24°S and around 26°S. In terms of longitude, the highest probabilities are concentrated near 42°W. Higher probabilities are also associated with proximity to the 200 m isobath, sea surface temperatures between 20-23°C, and intermediate depths.

The output map for Bryde's whales (Figure 2) indicates that this group predominantly occupies coastal and shallow areas, with increased probabilities of occurrence along the entire coastal region and the shelf-break, particularly in the northern sector where isobaths are closer to shore. This spatial pattern is also reflected in the partial dependence plots (Supplementary Figure 2). Occurrence probabilities increase with higher values of SST, chlorophyll-a concentration, salinity, and slope, as well as with negative sea level anomaly values; although all these estimates are accompanied by wide credible intervals.

3.1.2. Odontocetes

Within the odontocetes, the small delphinids subgroup is predicted to occur primarily in shallow waters and over the outer continental shelf and slope (Figure 2), with lower probabilities in deeper offshore areas. Occurrence probabilities are also higher at positive values of sea level anomaly. Rugosity and SST are also associated with occurrence predictions, although the credible intervals for these predictors are wide (Supplementary Figure 3).

The large delphinids subgroup is primarily associated with the 1000-m isobath across the study area, favoring deeper offshore regions. This pattern is particularly pronounced in the northern SB, where depths between 2000 and 3000 m seem to

represent particularly suitable habitat (Figure 2). In addition to bathymetric features, chlorophyll-a was also important in predicting the occurrence of this group (Table 4, Supplementary Figure 4).

Similar to the large delphinids, the sperm and beaked whale subgroup predominantly inhabits oceanic regions over the continental slope, especially areas surrounding the 1000-m isobath. However, unlike large delphinids, this group is more likely to occur in the southern half of the study area (Figure 2). Slope and chlorophyll-a show positive associations with whale presence, whereas rugosity, distance to 1000-m isobath, and distance to SST fronts exhibit negative relationships, although these estimates are accompanied by wide credible intervals (Supplementary Figure 5).

3.2. Projected future distributions

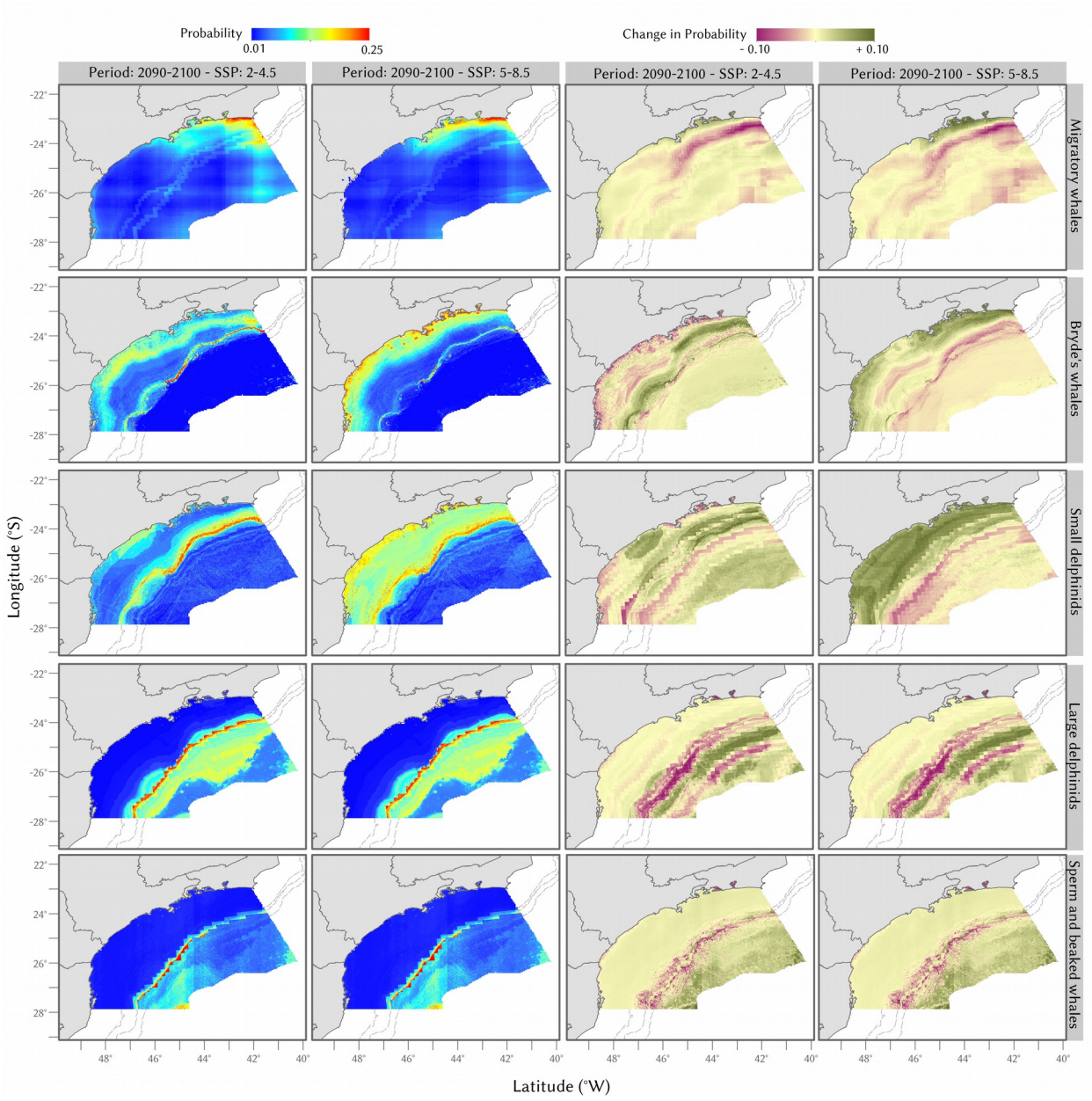


Figure 3. Maps depicting the future (period: 2090 - 2100) predicted distribution of the cetacean groups (migratory whales, Bryde's whales, small delphinids, large delphinids, sperm and beaked whales) and changes in probability, as obtained through BART models.

3.2.1 Mysticetes

The predictions suggest a spatial shift in the distribution of the migratory whales group, with higher predicted probabilities of occurrence in coastal waters and a decline in deeper offshore regions (Figure 3, Supplementary Figure 11). This shift is accompanied by an overall decrease in occurrence probability across the BS region under both medium-impact and high-impact scenarios, which becomes more pronounced in the far future (2090), especially under the high-impact scenario (SSP5-8.5).

For the Bryde's whales, the maps suggest a potential shoreward expansion of areas with higher occurrence probabilities (Figure 3), with expansion prevailing over contraction, particularly pronounced under the high-impact scenario (SSP5-8.5) in the far future (2090-2100).

3.2.2. Odontocetes

BART models predict little variation in the overall distribution patterns of small delphinids in the near future (by 2050, Supplementary Figure 11) under the medium-impact scenario (SSP2-4.5). However, in the far future, especially under the high-impact scenario (SSP5-8.5) the model predicts an expansion of relatively high-probability occurrence areas across the entire continental shelf up to the shelf break (Figure 3).

For large delphinids, BART model predictions suggest that distribution patterns will remain relatively stable in both the near future (2040-2050) and the distant future (2090-2100), regardless of the climate change scenario. However, compared to the baseline (Figure 2), the projections suggest a smoothed expansion in occurrence probability (Figure 3). While areas with the highest probabilities show a relative decrease, regions with lower probabilities over the continental slope exhibit a relative increase.

For the sperm and beaked whales group, the BART model indicates relatively minor spatial changes compared to the other groups, with overall distribution patterns remaining largely similar under both near-future (Supplementary Figure 11) and far-

future (Figure 3) projections, regardless of scenario. Nonetheless, a possible shift of the highest occurrence probabilities towards deeper offshore waters cannot be ruled out.

4. Discussion

In this study, we analyzed the most comprehensive dataset of ship-based cetacean sightings available for the Santos Basin, in the Southwestern Atlantic Ocean - a region of high relevance for cetaceans and a key area for Brazil's marine mammal diversity (Martins et al. 2023). We present an assessment of cetacean distribution and their potential responses to climate change in the SWAO, focusing on habitat preferences of major cetacean groups (migratory whales, Bryde's whales, small delphinids, large delphinids, sperm and beaked whales). To investigate the patterns and drivers of cetacean occurrence in the study area, we employed Bayesian Additive Regression Trees (BART), a decision tree-based modeling technique that has only recently been applied to SDMs. The models exhibited robust discriminatory performance across cetacean groups. By identifying habitat preferences, we were able to map the current and projected future distributions based on predicted probabilities of occurrence. The main findings and their potential implications for conservation are discussed below.

4.1. Mysticetes

The SB may function as a transit area, but also as a region of temporary residence for large migratory whales. With our approach, we identified environmental variables that can influence the presence of migratory baleen whales in this region. As expected, the best-fit model included spatial/bathymetric covariates (latitude, longitude, distance to coast and distance to 200 m isobath), and sea surface temperature. For cetaceans, a "shelf tendency" measure (distance from the coast or water depth) is a key variable influencing preferred habitats and community structure (Cañadas et al. 2002; MacLeod et al. 2007). But this pattern is mainly driven by the dominant species in this group sample, the humpback whale. In contrast with other balaenopterids (blue, fin, sei

and minke) that prefer deeper waters, humpback whales exhibit a preference for shelf habitats, particularly during their breeding and calving seasons (Forcada 2018; Pavanato et al. 2018), sometimes with high topographic variation. The slope is usually an area of intense use by some mysticete species, such as sei and fin whales (Gregr and Trites 2001; Weir et al. 2020), although humpbacks may also utilize slope waters (Di Tullio et al. 2016).

The higher probability of occurrence of migratory whales in the northern SB is likely driven by humpback whales (the most common migratory baleen whale in our sample) and the proximity to the Abrolhos Bank (16°40' – 19°30' S and 38°00' – 39°30' W), with environmental variables playing a secondary role. For the SWAO humpback whale population, the Abrolhos Bank – a region of the Brazilian coast with a broader continental shelf north of the Santos Basin – has been identified as the main breeding ground, hosting the highest concentrations along the coast (Andriolo et al. 2010; Wedekin et al. 2010; Pavanato et al. 2018). While the SB is located within the species' breeding range, it is primarily used as a migratory corridor by whales en route to the Abrolhos Bank (Landine et al. 2024). Nonetheless, increasing records of humpback whales in the region indicate a growing presence and raise concerns, particularly for juveniles, about potential conflicts with maritime traffic (Bortolotto et al. 2016). At the same time, the area may also function as a breeding ground, as indicated by records of males' song, competitive groups and mother-calf pairs. In contrast, the migratory and breeding patterns of other baleen whales, such as sei, fin, and minke whales, remain poorly understood in the SWAO, although growing evidence supports a preference for deeper oceanic regions (Baracho-Neto et al. 2025).

Sea water temperature is the most common descriptor in cetacean distribution models and is considered a key determinant of general species distributions (Melo-Merino et al. 2020; Pasanisi et al. 2024). In this study, the highest probabilities of presence for the migratory whale group were associated with SSTs ranging from 21-23°C, which are within the thermal range considered optimal in breeding areas (Rasmussen et al. 2007). This finding, together with other influential spatial and bathymetric predictors identified by the BART models, provides a basis for forecasting

the potential consequences of climate change on cetacean distributions within the study area, especially for breeding purposes. Calves are typically found in warmer wintering waters, which have been linked to enhanced development and reproductive success in humpback whales (Rasmussen et al. 2007). These projections should be interpreted cautiously, as decadal averages do not capture seasonal variation, and higher occurrences in the northern region may reflect both proximity to Abrolhos Bank and local cooler waters.

A more comprehensive understanding of the potential impacts of climate change on cetaceans is critical for informing effective conservation strategies. A recent review of climate-driven distributional shifts in cetaceans (van Weelden et al. 2021) suggests that the ranges of humpback, fin, and common minke whales are likely to remain stable or shift poleward. In contrast, Antarctic minke whales are expected to experience a contraction in habitat availability, coupled with a southward shift. Our study predicts a potential contraction in SB areas characterized by the highest probabilities of migratory whale occurrence. Furthermore, predictions indicate a spatial shift in distribution, with higher occurrence probabilities concentrated in more coastal waters. Notably, the northern coastal region of the SB is projected to show one of the lowest relative increases in SSTs under future climate scenarios, as indicated by Bio-ORACLE 3.0 data. According to von Hammerstein (2022), both intermediate and high-emission climate change scenarios may compromise the breeding habitat of humpback whales in the SWAO, primarily due to projected shifts in temperature away from their optimal thermal range.

Our study was conducted at a regional (meso) scale, focusing on species-habitat relationships at this finer resolution, in contrast to broader-scale approaches (e.g. species' distributional or geographical ranges). An advantage is that our input data were collected systematically in this region, eliminating potential spatial biases or sampling gaps. On the other hand, distribution models based uniquely on environmental variables, such as those presented here, often have limited explanatory power in capturing the full complexity of species distributions (Barry and Elith 2006). Consequently, it is unlikely that physiological tolerances represent a limiting factor within the study region.

Environmental conditions present in the area, even if suboptimal, would still be sufficient to support the cetacean groups' survival. The patterns observed and described here reflect habitat preferences and active selection, rather than mere restriction or unsuitability.

Our best-fit model for Bryde's whales included both static (depth, slope, rugosity, and distance to coast) and most of the dynamic variables considered (SST, chlorophyll-a, sea level anomaly, and salinity). Unlike migratory whale species, Bryde's whales are observed year-round in the SB and are therefore expected to be more closely associated with variables linked to prey availability, as feeding is the predominant behavior recorded in coastal areas of southeastern Brazil (Lodi et al. 2015). Our findings regarding the explanatory variables influencing Bryde's whale occurrence were partially consistent with previous studies in the region. Dalpaz et al. (2023) identified distance to the coast and slope as important static predictors in their most parsimonious generalized least squares (GLS) model for encounter rates, while current speed was selected as the key dynamic variable. Predictive approaches such as BART, applied here, are expected to identify a broader set of influential predictors than parametric models (Bleich et al. 2014), given their ability to capture complex interactions. Dalpaz et al. (2023) also identified key areas along the coastline and the shelf-break extension that align with those highlighted here and reported seasonal shifts in distribution. During warmer periods, Bryde's whales were predominantly observed in the northern portion of the study area, near Cabo Frio, a region influenced by upwelling events that enhance productivity and increase prey availability. In colder seasons, their distribution shifted southward, suggesting a behavioral response to seasonal variability in oceanographic conditions and prey distribution. In the Santos Basin, prey resources in these upwelling-influenced areas include small pelagic fishes such as the Brazilian sardine (Sardinella brasiliensis) (Dalpaz et al. 2023). Climate projections for the southwestern Atlantic indicate ongoing warming trends, increased stratification, and potential alterations in circulation and productivity (Franco et al. 2022), as well as projected shifts in larval fish distribution and exploited fish stocks (Costa et al. 2023). Such environmental changes may modify the spatial distribution, recruitment dynamics,

and local availability of these forage species, potentially influencing Bryde's whale habitat use through bottom-up ecosystem processes.

Climate change is expected to drive a general expansion in the distribution range of the Bryde's whale, mainly towards the inner shelf, according to the review by van Weelden et al. (2021). In this review, the authors also highlight that stranding records suggest climate change may increasingly favor warm-water species over cold-water species. Our predictions are consistent with these findings, forecasting an expansion of areas with higher probabilities of Bryde's whale occurrence within the SB under future scenarios. However, as Bryde's whales are strongly linked to prey availability, any potential expansion in their distribution will likely depend on corresponding responses of small pelagic fish populations under future climate conditions.

4.2. Odontocetes

The continental slope, an area of intense use by mysticetes, is also critically important for toothed whales, representing the region of highest cetacean species richness in the SB (Martins et al. 2023). This zone provides essential habitats and feeding grounds characterized by steep bathymetric gradients. Subsurface processes that enhance productivity, such as the shelf-slope fronts, predominate in this area (Belkin and Cornillon 2007). Additionally, the pronounced depth gradients are likely to support greater prey diversity. Bathymetric features such as canyons, ridges, and irregular topographies also provide navigational cues that facilitate prey capture (Moors-Murphy 2014) or may channel bottom nutrient-rich water promoting enhanced biodiversity (Castro et al. 2023), as exemplified by the two submarine canyons present within the study area. In our predictive models for odontocete groups, variables related to bathymetry, including distance to 1000 m isobath, depth, rugosity, and slope, were consistently selected, highlighting their relevance to habitat use by cetaceans. Protection of slope habitats and associated oceanographic hotspots could enhance the resilience of cetacean populations to climate change, consistent with recent assessments of the Brazilian MPA network (Tardin et al. 2025).

Dynamic covariates are essential for capturing the temporal and spatial variability of species' habitats (Chavez-Rosales et al. 2019; Pasanisi et al. 2024). In our study, chlorophyll-a was an important predictor for large delphinids as well as for sperm and beaked whales, whereas SST was particularly relevant for small delphinids. Incorporating such covariates into SDMs is critical for assessing potential climate change impacts across cetacean groups. Additionally, distance to oceanic fronts and sea level anomaly were retained in the models for sperm and beaked whales and for small delphinids, respectively, highlighting the influence of oceanographic processes on distribution patterns.

For small delphinids, our results indicate a higher probability of occurrence in shallower waters compared to other odontocete groups. Although some species in this group exhibit broader depth ranges or even offshore preferences (e.g. Fertl et al. 2003), our findings align with the general pattern for this group, which is frequently associated with coastal areas and the continental slope. SST is commonly linked to small delphinid species in SDMs (MacLeod et al. 2007, Chavez-Rosales et al. 2019). While sea surface height and related anomalies can influence cetacean habitats by affecting oceanographic conditions and prey availability, such variables are less frequently selected in SDMs when multiple species or groups are analyzed (Chavez-Rosales et al. 2019). In this study, considering small delphinids as a group, SLA emerged as a significant predictor, with higher probabilities of occurrence associated with positive SLA values, typically indicative of anticyclonic eddies (Silveira et al. 2000). Although anticyclonic eddies generally do not promote nutrient upwelling, they can affect local prey dynamics by creating convergence zones or modifying the behavior of marine organisms. Notably, in subtropical gyres, anticyclonic eddies may be more productive than cyclonic eddies, as they can modulate winter mixing (Dufois et al. 2016). Nevertheless, the mechanisms linking eddy-driven productivity to cetacean distribution are complex, and further research is needed to clarify these ecological relationships.

Regarding future projections, Tardin et al. (2025) analyzed the potential distribution of *Sotalia guianensis* under the latest CMIP6 climate scenarios. In the SB region, their models predict the maintenance or even expansion of suitable habitat across all

scenarios. However, when considering the entire Brazilian coast and distribution range of this threatened species, range contractions are predicted. The authors note that mitigation measures, such as the establishment and effective management of Marine Protected Areas (MPAs), could help reduce the severity of climate change impacts on Guiana dolphins. Similar patterns were observed for *Steno bredanensis* (Ferreira et al. 2024). Additionally, van Weelden et al. (2021) projected an expansion in the distribution range of the most common small delphinid species under future climate conditions. Our projections are consistent with these studies, also indicating a potential expansion of suitable areas for the small delphinid group under future climate scenarios in the SB region.

Large delphinids are most commonly found in oceanic waters (Roberts et al. 2016). In our model, this group - along with beaked and sperm whales - was strongly associated with the 1000 m isobath. For deep-diving odontocetes, such as beaked and sperm whales, our models predicted distributions concentrated in deep waters with high-relief bathymetry, consistent with evidence of high prey density in these areas, as reported for the North Atlantic and Gulf of Mexico (Roberts et al. 2016). In oceanic waters, large delphinids exhibited the highest probabilities of occurrence in patchy distributions concentrated further north, whereas sperm and beaked whales were more commonly found in southern areas. Similarly, Di Tullio et al. (2016) observed sperm whales primarily in the southern portion of the study area, particularly in deep waters over the continental slope. Despite the considerable trophic gap between primary productivity and deep-diving predators - which occupy higher levels of the trophic web - the occurrence of these species is also influenced by dynamic variables such as primary productivity (MacLeod et al. 2007), which plays a key role in determining prey availability and distribution.

Our results indicate minor changes in the future distribution ranges of these two last odontocete groups, aligning with the projections of van Weelden et al. (2021), who suggested that their ranges would remain largely stable, with some potential for expansion. Similarly, changes in the distribution of sperm whales in the North Atlantic over the past 30 years resembled a range expansion rather than a shift (Ramirez-

Martinez et al. 2024). These authors also highlighted the strong association between sperm whales and depth, suggesting it as a reliable proxy for prey availability. However, in the context of ocean warming and shifts in pelagic fish distributions, they propose that sperm whales may be adapting to environmental changes by remaining in the same areas while targeting different prey species (Ramirez-Martinez et al. 2024).

4.3. Modeling considerations: potentialities and limitations

An important aspect of this study was the strong ability of the BART model to predict species occurrence with limited presence data. Bayesian models, in particular, offer distinct advantages when data volumes are limited due to their use of MCMC simulation techniques and regularizing priors (Gelman et al. 2020). These features enhance computational robustness and improve inferential accuracy, making Bayesian approaches especially effective under such conditions.

Species distribution models (SDMs) estimate occurrence areas and support decision-making for species and groups, often under significant uncertainty. Decision analysis under uncertainty is a topic of great relevance and practical applicability within the Bayesian approach to data analysis, which utilizes posterior distributions inferred from the data. From this perspective, BART stands out as a distinctive tool. In addition to being effective for SDMs, BART's Bayesian formulation allows the predictive flexibility of tree-based models to be combined with a principled treatment of uncertainty. Unlike traditional ensembles, which improve robustness but generally do not provide coherent inferential outputs, BART delivers posterior estimates that can be used to assess and compare alternative actions under uncertainty. Its final output, posterior distributions, enables a comprehensive application of Bayesian theory to support decision-making in ecological management.

An additional advantage of our dataset is the availability of true absences, which is uncommon in most SDM studies that rely on pseudo-absences or background points. The presence of true zeros provides stronger inferential support, allows the use of models beyond presence-only methods such as MaxEnt (Phillips et al. 2006), and enhances the reliability of predictions (Pasanisi et al. 2024).

Most studies assessing the effects of climate change on cetaceans focus on individual species. Considering the entire cetacean community of Santos Basin, as well as groups of species with taxonomic and biological similarities, provides a broader perspective on the impacts of climate change and can support management and conservation efforts based on an ecosystem approach. By aggregating species, species-specific patterns are diluted, revealing only consistent and generalized trends across the group. Nevertheless, since related species often share similar ecological requirements and dispersal capacities, the distribution of higher taxonomic groups can still exhibit distinct patterns and constraints. Such patterns reflect the cumulative distributions of the species they encompass (LeDuc 2018). In addition to that, pooling data across species with few records increases statistical power and reduces data limitations.

Uncertainty in the output models was generally high, particularly in areas identified as having the highest habitat suitability. This outcome was expected, as the variance in predicted probabilities tends to increase as values rise from zero to 0.5. Despite the extensive survey effort and broad spatial coverage achieved through the PMC-BS, the number of sightings remained low for some groups and species. Increasing sample sizes through additional data collection, and explicitly accounting for imperfect detection during model development, would likely improve predictive accuracy and reduce uncertainty. Nevertheless, the regional spatial scale and systematic sampling employed in this study provide a more robust approach than analyzing the entire species range, as broader-scale studies often face spatial gaps and sampling biases in occurrence data, potentially leading to biased current and future projections. Either way, we recommend further validation of the models using independent external datasets to enhance their robustness and applicability.

Model uncertainty arises from multiple sources, including sampling, process, and model selection errors (Redfern et al. 2006). In addition, most of the environmental predictor layers used in our analysis were derived from remote sensing or model-based products, which inherently carry their own uncertainties. These may result from random noise, systematic bias, or spatially clustered errors (Barry and Elith 2006), all of which can propagate through to the final model outputs.

Seasonal variation may influence species distributions, particularly for migratory species, and may affect non-migratory species through changes in food availability (Dalpaz et al. 2023; Martins et al. 2023). However, the projected future environmental layers used in this study represent decadal averages and do not provide monthly or seasonal data (Assis et al. 2024). In addition, some variables identified as important - such as sea level anomaly and distance to SST fronts - were unavailable or unsuitable for future projections and could not be incorporated. Consequently, neither seasonality nor these variables could be explicitly modeled, representing limitations that future studies could address using seasonal or more comprehensive environmental datasets.

Our study is intended solely to identify areas with the highest likelihood of occurrence and to project potential future distributions of cetacean groups based on available data and climate scenarios. Predictions of future population trends and estimates of cetacean densities are beyond the scope of this study. The recovery of some whale populations from historical exploitation may also contribute to future shifts in their distribution patterns.

The broader implications of climate change may include increased interspecific competition, genetic alterations, ecosystem-level alterations, and new conservation challenges (van Weelden et al. 2021). Shifts in species' niches may occur in response to changing climate conditions, habitat availability, or evolutionary pressures (Soberon 2011). Further research is essential to determine vulnerabilities and conservation actions.

Conclusion

This study provides the first comprehensive assessment of cetacean distribution patterns and their projected responses to climate change in the Santos Basin, revealing distinct ecological trends among major taxonomic groups. Environmental predictors for each group were identified using BART models applied to systematically collected regional data, allowing robust assessment of habitat preferences and potential changes. Migratory whales, particularly humpbacks, may undergo contractions and shifts of suitable habitats toward coastal areas; whereas Bryde's whales and small delphinids

may expand their ranges; and large delphinids, sperm, and beaked whales are expected to remain relatively stable. These findings highlight the ecological importance of the region as a migratory corridor, feeding ground, and potential breeding habitat, while also underscoring emerging risks of overlap with human activities. The projections presented here can inform adaptive conservation strategies, spatial prioritization, and long-term monitoring, while acknowledging the uncertainties inherent to climate projections and sampling limitations.

Authors' contribution

ACM: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing - original draft, writing - review & editing; PGK: conceptualization, formal analysis, investigation, methodology, visualization, writing - review & editing and supervision; LLW: investigation, project administration, resources, writing - review & editing; LDR: conceptualization, formal analysis, investigation, methodology, visualization, writing - review & editing and supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support this study are property of Petrobras and the summary of these, allowed for sharing, are available at the address https://sispmcprd.petrobras.com.br/sispmc.

Acknowledgements

We thank Petrobras for kindly authorizing the use of data collected during the Projeto de Monitoramento de Cetáceos na Bacia de Santos (PMC-BS). The PMC-BS is one of the monitoring programs required by Brazil's federal environmental agency,

IBAMA, for the environmental licensing process of oil production and transport by Petrobras in the Santos Basin (process no. 02001.114279/2017-80, ACCTMB no. 657/2015). We also thank all marine mammal observers who helped to gather data during PMC-BS, and the Socioambiental Consultores Associados. We appreciate the valuable comments of L. Bugoni and M. Camargo on an earlier version of the manuscript. The first author is grateful to the Coordination for the Improvement of Higher Education Personnel (CAPES), which provided financial support during her PhD. This research was developed under the guidance of the second and last authors. L. Dalla Rosa thanks the research fellowship provided by Brazilian National Research Council (CNPq)(PQ 309548/2019-5).

References

Andriolo, A., Kinas, P.G., Engel, M.H., Martins, C.C.A., Rufino, A.M., 2010. Humpback whales within the Brazilian breeding ground: distribution and population size estimate. Endang. Species Res. 11 (3), 233-243. https://doi.org/10.3354/esr00282

Assis, J., Fernández Bejarano, S.J., Salazar, V.W., Schepers, L., Gouvea, L., Fragkopoulou, E., Leclercq, F., Vanhoorne, B., Tyberghein, L., Serrão, E.A., Verbruggen, H., De Clerck, O., 2024. Bio-ORACLE v3.0. Pushing marine data layers to the CMIP6 Earth System Models of climate change research. Glob. Ecol. Biogeogr. 33 (4), e13813. https://doi.org/10.1111/geb.13813

Baracho-Neto, C.G., de Oliveira, L.L., Wedekin, L.L., Cremer, M.J., Zerbini, A.N., 2025. Movements of the endangered sei whale *Balaenoptera borealis* in the Southwestern Atlantic Ocean. Mar. Ecol. Prog. Ser. 763,157-169. https://doi.org/10.3354/meps14878

Baquero, R.A., Barbosa, A.M., Ayllón, D., Guerra, C., Sánchez, E., Araújo, M.B., Nicola, G.G., 2021. Potential distributions of invasive vertebrates in the Iberian Peninsula under projected changes in climate extreme events. Divers. Distrib. 27 (11), 2262-2276. https://doi.org/10.1111/ddi.13401

Barry, S., Elith, J., 2006. Error and uncertainty in habitat models. J. Appl. Ecol. 43 (3), 413-423. https://doi.org/10.1111/j.1365-2664.2006.01136.x

Belkin, I.M., Cornillon, P.C., 2007. Fronts in the world ocean's large marine ecosystems. ICES CM 2007/D:21. ICES Annu. Sci. Conf., Finland.

Bleich, J., Kapelner, A., George, E.I., Jensen, S.T., 2014. Variable selection for BART: an application to gene regulation. Ann. Appl. Stat. 8 (3), 1750-1781. https://doi.org/10.1214/14-AOAS755

Bortolotto, G.A., Kolesnikovas, C.K.M., Freire, A.S., Simões-Lopes, P.C., 2016. Young humpback whale *Megaptera novaeangliae* feeding in Santa Catarina coastal waters, Southern Brazil, and a ship strike report. Mar. Biodivers. Rec. 9, 1-6. https://doi.org/10.1186/s41200-016-0043-4

Buckland, S., Anderson, D., Burnham, K., Laake, J., Borchers, D., Thomas, L., 2004. Advanced distance sampling: Estimating abundance of biological populations. Oxford University Press, Oxford.

Cañadas, A., Sagarminaga, R., García-Tiscar, S., 2002. Cetacean distribution related with depth and slope in the Mediterranean waters off southern Spain. Deep-Sea Res. Part I: Oceanogr. Res. Pap. 49 (11), 2053-2073. https://doi.org/10.1016/S0967-0637(02)00123-1

Carlson, C.J., 2020. embarcadero: Species distribution modelling with Bayesian additive regression trees in R. Methods Ecol. Evol. 11 (7), 850-858. https://doi.org/10.1111/2041-210X.13389

Castro, G.M., Vargens, R.P., Carlos-Júnior, L.A., Cardoso, F.C., Salomon, P.S., Tenório, M.M.B., Bastos, A.C., Oliveira, N., Ghisolfi, R.D., Cordeiro, R.T.S., Moura, R.L., 2023. Incised valleys drive distinctive oceanographic processes and biological assemblages within rhodolith beds. PLoS ONE 18 (11), e0293259. https://doi.org/10.1371/journal.pone.0293259

Chavez-Rosales, S., Palka, D.L., Garrison, L.P., Josephson, E.A., 2019. Environmental predictors of habitat suitability and occurrence of cetaceans in the western North Atlantic. Ocean. Sci. Rep. 9 (1), 5833. https://doi.org/10.1038/s41598-019-42288-6

Chipman, H.A., George, E.I., McCulloch, R.E., 2010. BART: Bayesian additive regression trees. Ann. Appl. Stat. 4 (1), 266-298. https://doi.org/10.1214/09-AOAS285

Corrêa, T.P., Wedekin, L.L., Cremer, M.J., 2023. Occurrence and habitat use of pantropical spotted dolphins (*Stenella attenuata*) at the edge of their distribution in the southwestern Atlantic Ocean. Mar. Mamm. Sci. 40 (2), e13094. https://doi.org/10.1111/mms.13094

Costa, M.D., Wilson, K.A., Dyer, P.J., Pitcher, R., Muelbert, J.H., Richardson, A.J., 2021. Potential future climate-induced shifts in marine fish larvae and harvested fish communities in the subtropical southwestern Atlantic Ocean. Clim. Change 165 (3), 66. https://doi.org/10.1007/s10584-021-03097-x

Dalla Rosa, L., Ford, J.K.B., Trites, A.W., 2012. Distribution and relative abundance of humpback whales in relation to environmental variables in coastal British Columbia and adjacent waters. Cont. Shelf Res. 36, 89-104. https://doi.org/10.1016/j.csr.2012.01.017

Dalpaz, L., Paro, A.D., Daura-Jorge, F.G., Rossi-Santos, M., Norris, T.F., Ingram, S.N., Wedekin, L.L., 2021. Better together: Analysis of integrated acoustic and visual methods when surveying a cetacean community. Mar. Ecol. Prog. Ser. 678, 197-209. https://doi.org/10.3354/meps13898

Dalpaz, L., Cypriano-Souza, A.L., Lodi, L., Wedekin, L., Daura-Jorge, F., 2023. Bryde's whales in South Brazil Bight: Evidence of low genetic diversity and seasonal habitat use. Mar. Biol. 170 (8), 94. https://doi.org/10.1007/s00227-023-04241-0

Di Tullio, J.C., Gandra, T.B.R., Zerbini, A.N., Secchi, E.R., 2016. Diversity and distribution patterns of cetaceans in the subtropical Southwestern Atlantic outer continental shelf and slope. PLoS ONE 11 (5), e0155841. https://doi.org/10.1371/journal.pone.0155841

Dorneles, P.R., Sanz, P., Eppe, G., Azevedo, A.F., Bertozzi, C.P., Martínez, M.A., Secchi, E.R., Barbosa, L.A., Cremer, M., Alonso, M.B., Torres, J.P.M., Lailson-Brito, J., Malm, O., Eljarrat, E., Barceló, D., Das, K., 2013. High accumulation of PCDD, PCDF, and PCB congeners in marine mammals from Brazil: a serious PCB problem. Sci. Total Environ. 463, 309-318. https://doi.org/10.1016/j.scitotenv.2013.06.015

Dufois, F., Hardman-Mountford, N.J., Greenwood, J., Richardson, A.J., Feng, M., Matear, R.J., 2016. Anticyclonic eddies are more productive than cyclonic eddies in subtropical gyres because of winter mixing. Sci. Adv. 2, e1600282. https://doi.org/10.1126/sciadv.1600282

ESRI, 2018. ArcGIS Desktop: Release 10.6. Redlands, CA: Environmental Systems Research Institute. URL https://www.esri.com

Ferreira, J.D.N., Vale, M.M., Maricato, G., Maciel, I., Lodi, L., Daura-Jorge, F.G., Millman, L., Marcondes, M., Tardin, R.H., 2024. Climate change is likely to jeopardize a marine apex predator (*Steno bredanensis*, Cetacea) and its prey on the Coast of Brazil. SSRN 5069671. https://doi.org/10.2139/ssrn.5069671

Fertl, D., Jefferson, T.A., Moreno, I.B., Zerbini, A.N., Mullin, K.D., 2003. Distribution of the Clymene dolphin *Stenella clymene*. Mammal Rev. 33 (3), 253-271. https://doi.org/10.1046/j.1365-2907.2003.00033.x

Forcada, J., 2018. Distribution. In: Würsig B., Thewissen J.G.M., Kovacs K.M (Eds). Encyclopedia of Marine Mammals. 3rd ed., Academic Press, London, pp 259-262.

Fossi, M.C., Casini, S., Caliani, I., Panti, C., Marsili, L., Viarengo, A., Giangreco, R., di Sciara, G.N., Serena, F., Ouerghi, A., Depledge, M.H., 2012. The role of large marine vertebrates in the assessment of the quality of pelagic marine ecosystems. Mar. Environ. Res. 77, 156-158. https://doi.org/10.1016/j.marenvres.2012.03.003

Franco, B.C., Defeo, O., Piola, A.R., Barreiro, M., Yang, H., Ortega, L., Gianelli, I., Castello, J.P., Vera, C., Buratti, C., Pájaro, M., Pezzi, L.P., Möller, O.O., 2020. Climate change impacts on the atmospheric circulation, ocean, and fisheries in the southwest South Atlantic Ocean: a review. Clim. Change 162(4), 2359-2377. https://doi.org/10.1007/s10584-020-02783-6

Gelman, A., Vehtari, A., Simpson, D., Margossian, C.C., Carpenter, B., Yao, Y., Kennedy, L., Gabry, J., Burkner, P.C., Modrak, M., 2020. Bayesian workflow. arXiv:2011.01808. https://doi.org/10.48550/arXiv.2011.01808

Gregr, E.J., Trites, A.W., 2001. Predictions of critical habitat for five whale species in the waters of coastal British Columbia. Can. J. Fish. Aquat. Sci. 58 (7), 1265-1285. https://doi.org/10.1139/f01-078

Grose, S.O., Pendleton, L., Leathers, A., Cornish, A., Waitai, S., 2020. Climate change will re-draw the map for marine megafauna and the people who depend on them. Front. Mar. Sci. 7, 547. https://doi.org/10.3389/fmars.2020.00547

Guisan, A., Thuiller, W., 2005. Predicting species distribution: Offering more than simple habitat models. Ecol. Lett. 8 (9), 993-1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Hazen, E.L., Abrahms, B., Brodie, S., Carroll, G., Jacox, M.G., Savoca, M.S., Scales, K.L., Sydeman, W.J., Bograd, S.J., 2019. Marine top predators as climate and ecosystem sentinels. Front. Ecol. Environ. 17 (10), 565-574. https://doi.org/10.1002/fee.2125

Hill, J., Linero, A., Murray, J., 2020. Bayesian additive regression trees: A review and look forward. Annu. Rev. Stat. Its Appl. 7 (1), 251-278. https://doi.org/10.1146/annurev-statistics-031219-041110

IPCC, 2021. Climate change 2021 — The physical science basis. Chemistry Int. 43 (4), 22-23. https://doi.org/10.1515/ci-2021-0407

Landine, A., Zerbini, A.N., Danilewicz, D., Sucunza, F., Andriolo, A., 2024. Humpback whales in the Southwest Atlantic Ocean: investigating their breeding movements by satellite tracking. Mar. Mamm. Sci. 40 (4), e13146. https://doi.org/10.1111/mms.13146

LeDuc, R., 2018. Biogeography. In: Würsig B., Thewissen J.G.M., Kovacs K.M (Eds). Encyclopedia of Marine Mammals. 3rd ed., Academic Press, London, pp 99-103.

Lima, G. dos S., Pedrobom, J.H., Suarez, C.A., Torres-Florez, J.P., Vidal, L.G., Domit, C., Menegario, A.A., 2024. Bioaccumulation of trace elements in marine mammals: new data and transplacental transfer on threatened species. Sci. Total Environ. 948, 174854. https://doi.org/10.1016/j.scitotenv.2024.174854

Lobo, A.J., Wedekin, L.L., Sobral-Souza, T., Le Pendu, Y., 2021. Potential distribution of Guiana dolphin (*Sotalia guianensis*): a coastal-estuarine and tropical habitat specialist. J. Mammal. 102 (1), 308-318. https://doi.org/10.1093/jmammal/gyaa153

Lodi, L., Tardin, R.H., Hetzel, B., Maciel, I.S.M., Figueiredo, L.D., Simão, S.M., 2015. Bryde's whale (Cetartiodactyla: Balaenopteridae) occurrence and movements in coastal areas of southeastern Brazil. Zoologia 32 (2), 171-175. https://doi.org/10.1590/S1984-46702015000200009

MacLeod, C.D., Weir, C.R., Pierpoint, C., Harland, E.J., 2007. The habitat preferences of marine mammals west of Scotland (UK). J. Mar. Biol. Assoc. U. K. 87 (1), 157-164. https://doi.org/10.1017/S0025315407055270

Maricato, G., Tardin, R., Lodi, L., Wedekin, L.L., Daura-Jorge, F.G., Maciel, I., Maria, T.F., Alves, M.A.S., 2022. Identifying suitable areas for common bottlenose dolphin in anthropized waters. Mar. Biol. 169(9), 110. https://doi.org/10.1007/s00227-022-04095-y

Martins, A.C., Kinas, P.G., Wedekin, L.L., Dalla Rosa, L., 2023. Spatial and seasonal patterns of cetacean species richness: a Bayesian approach. Deep-Sea Res. Part I: Oceanogr. Res. Pap. 196, 104046. https://doi.org/10.1016/j.dsr.2023.104046

Melo, C.L.C., Santos, R.A., Bassoi, M., Araújo, A.C., Lailson-Brito, J., Dorneles, P.R., Azevedo, A.F., 2010. Feeding habits of delphinids (Mammalia: Cetacea) from Rio de Janeiro State, Brazil. J. Mar. Biol. Assoc. U. K. 90 (8), 1509-1515. https://doi.org/10.1017/S0025315409991639

Melo-Merino, S.M., Reyes-Bonilla, H., Lira-Noriega, A., 2020. Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. Ecol. Model. 415, 108837. https://doi.org/10.1016/j.ecolmodel.2019.108837

Moors-Murphy, H.B., 2014. Submarine canyons as important habitat for cetaceans, with special reference to the Gully: a review. Deep-Sea Res. Part II: Top. Stud. Oceanogr. 104, 6-19. https://doi.org/10.1016/j.dsr2.2013.12.016

Moreno, I.B., Zerbini, A.N., Danilewicz, D., de Oliveira Santos, M.C., Simões-Lopes, P.C., Lailson-Brito, J., Azevedo, A.F., 2005. Distribution and habitat characteristics of

dolphins of the genus *Stenella* (Cetacea: Delphinidae) in the southwest Atlantic Ocean. Mar. Ecol. Prog. Ser. 300, 229-240. https://doi.org/10.3354/meps300229

Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., Araújo, M.B., Dallas, T., Dunson, D., Elith, J., Foster, S.D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N.A., Holt, R.D., Hui, F.K.C., …, Ovaskainen, O., 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecol. Monogr. 89 (3), e01370. https://doi.org/10.1002/ecm.1370

Pasanisi, E., Pace, D.S., Orasi, A., Vitale, M., Arcangeli, A., 2024. A global systematic review of species distribution modelling approaches for cetaceans and sea turtles. Ecol. Inform. 82, 102700. https://doi.org/10.1016/j.ecoinf.2024.102700

Pavanato, H.J., Mayer, F.P., Wedekin, L.L., Engel, M.H., Kinas, P.G., 2018. Prediction of humpback whale group densities along the Brazilian coast using spatial autoregressive models. Mar. Mamm. Sci. 34 (3), 734-754. https://doi.org/10.1111/mms.12492

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M., Araújo, M.B., 2011. Ecological Niches and Geographic Distributions (MPB-49). Princeton University Press, Princeton.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190 (3–4), 231-259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pires, C.R., Rossi-Santos, M.R., Paro, A.D., Wedekin, L.L., 2021. Whistles of the pantropical spotted dolphin (*Stenella attenuata*) in Santos Basin, western South Atlantic Ocean. J. Acoust. Soc. Am. 149 (5), 3241-3249. https://doi.org/10.1121/10.0004950

Plant, E., King, R., Kath, J., 2021. Statistical comparison of additive regression tree methods on ecological grassland data. Ecol. Inform. 61, 101198. https://doi.org/10.1016/j.ecoinf.2020.101198

R Core Team, 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org

Ramirez-Martinez, N.C., Víkingsson, G.A., Øien, N.I., Mikkelsen, B., Gunnlaugsson, T., Hammond, P.S., 2024. Distribution and habitat use of deep-diving cetaceans in the central and north-eastern North Atlantic. NAMMCO Sci. Publ. 13. https://doi.org/10.7557/3.7416

Rasmussen, K., Palacios, D.M., Calambokidis, J., Saborío, M.T., Dalla Rosa, L., Secchi, E.R., Steiger, G.H., Allen, J.M., Stone, G.S., 2007. Southern Hemisphere humpback whales wintering off Central America: Insights from water temperature into the longest mammalian migration. Biol. Lett. 3 (3), 302-305. https://doi.org/10.1098/rsbl.2007.0067

Redfern, J.V., Ferguson, M.C., Becker, E.A., Hyrenbach, K.D., Good, C., Barlow, J., Kaschner, K., Baumgartner, M.F., Forney, K.A., Ballance, L.T., Fauchald, P., Halpin, P., Hamazaki, T., Pershing, A.J., Qian, S.S., Read, A., Reilly, S.B., Torres, L., Werner, F., 2006. Techniques for cetacean-habitat modeling. Mar. Ecol. Prog. Ser. 310, 271-295. https://doi.org/10.3354/meps310271

Roberts, J.J., Best, B.D., Dunn, D.C., Treml, E.A., Halpin, P.N., 2010. Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. Environ. Model. Softw. 25 (10), 1197-1207. https://doi.org/10.1016/j.envsoft.2010.03.029

Roberts, J.J., Best, B.D., Mannocci, L., Fujioka, E.I., Halpin, P.N., Palka, D.L., Garrison, L.P., Mullin, K.D., Cole, T.V., Khan, C.B., McLellan, W.A., Pabst, D.A., Lockhart, G.G., 2016. Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico. Sci. Rep. 6 (1), 22615. https://doi.org/10.1038/srep22615

Rodrigues, L. dos S., Daudt, N.W., Cardoso, L.G., Kinas, P.G., Conesa, D., Pennino, M.G., 2023. Species distribution modelling in the Southwestern Atlantic Ocean: A systematic review and trends. Ecol. Model. 486, 110514. https://doi.org/10.1016/j.ecolmodel.2023.110514

Roman, J., Estes, J.A., Morissette, L., Smith, C., Costa, D., McCarthy, J., Nation, J., Nicol, S., Pershing, A., Smetacek, V., 2014. Whales as marine ecosystem engineers. Front. Ecol. Environ. 12 (7), 377–385. https://doi.org/10.1890/130220

Santos, M.C. de O., Laílson-Brito, J., Flach, L., Oshima, J.E.F., Figueiredo, G.C., Carvalho, R.R., Ventura, E.S., Molina, J.M.B., Azevedo, A.F., 2019. Cetacean movements in coastal waters of the southwestern Atlantic ocean. Biota Neotrop. 19 (2), e20180670. https://doi.org/10.1590/1676-0611-BN-2018-0670

Siciliano, S., Ramos, R.M.A., di Beneditto, A.P.M., Santos, M.C.O., Fragoso, A.B., Brito, J.L., Azevedo, A.F., Vicente, A.F.C., Zampirolli, E., Alvarenga, F.S., Barbosa, L., Lima, N.R.W., 2007. Age and growth of some delphinids in south-eastern Brazil. J. Mar. Biol. Assoc. U. K. 87 (1), 293-303. https://doi.org/10.1017/S0025315407053398

Silveira, I.C.A. da., Schmidt, A.C.K., Campos, E.J.D., Godoi, S.S.D., Ikeda, Y., 2000. A corrente do Brasil ao largo da costa leste brasileira. Rev. Bras. Oceanogr. 48, 171-183. https://doi.org/10.1590/S1413-77392000000200008

Soberon, J.M., 2010. Niche and area of distribution modeling: a population ecology perspective. Ecography 33 (1), 159-167. https://doi.org/10.1111/j.1600-0587.2009.06074.x

Tardin, R., Maricato, G., Kiszka, J.J., Cantor, M., Maciel, I., Melo-Santos, G., Vale, M.M., 2025. Optimistic climate mitigation scenario halves projected range loss in a neotropical dolphin. Ocean Coast. Manag. 269, 107800. https://doi.org/10.1016/j.ocecoaman.2025.107800

Tredennick, A.T., Hooker, G., Ellner, S.P., Adler, P.B., 2021. A practical guide to selecting models for exploration, inference, and prediction in ecology. Ecology 102 (6), e03336. https://doi.org/10.1002/ecy.3336

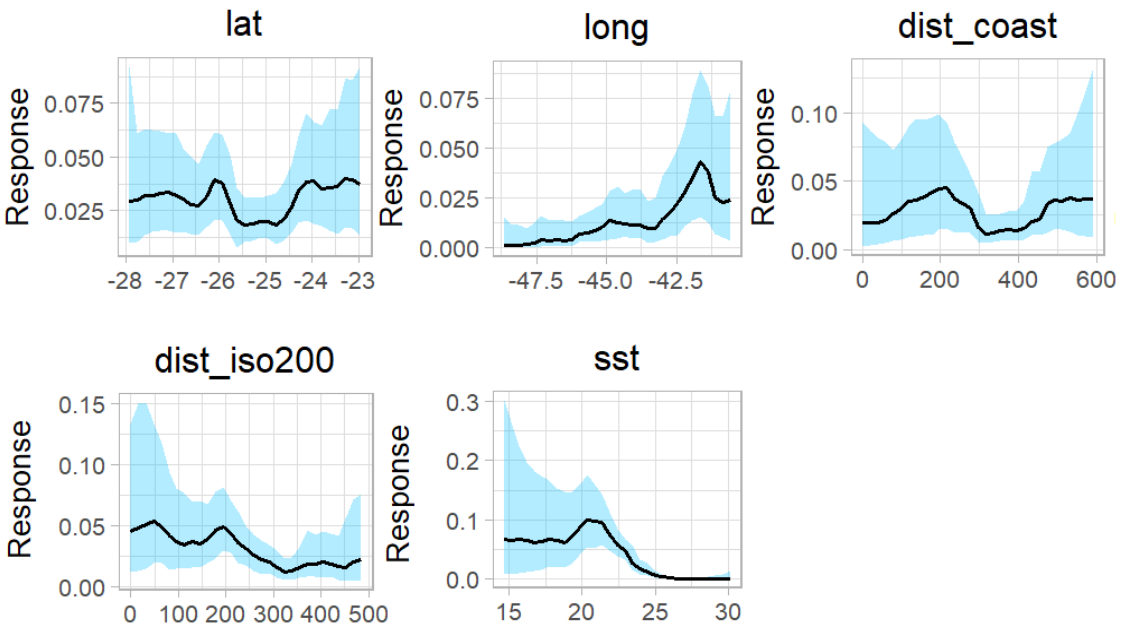
van Weelden, C., Towers, J.R., Bosker, T., 2021. Impacts of climate change on cetacean distribution, habitat and migration. Climate Change Ecol. 1, 100009. https://doi.org/10.1016/j.ecochg.2021.100009

von Hammerstein, H., Setter, R.O., van Aswegen, M., Currie, J.J., Stack, S.H., 2022. High-resolution projections of global sea surface temperatures reveal critical warming in humpback whale breeding grounds. Front. Mar. Sci. 9, 837772. https://doi.org/10.3389/fmars.2022.837772

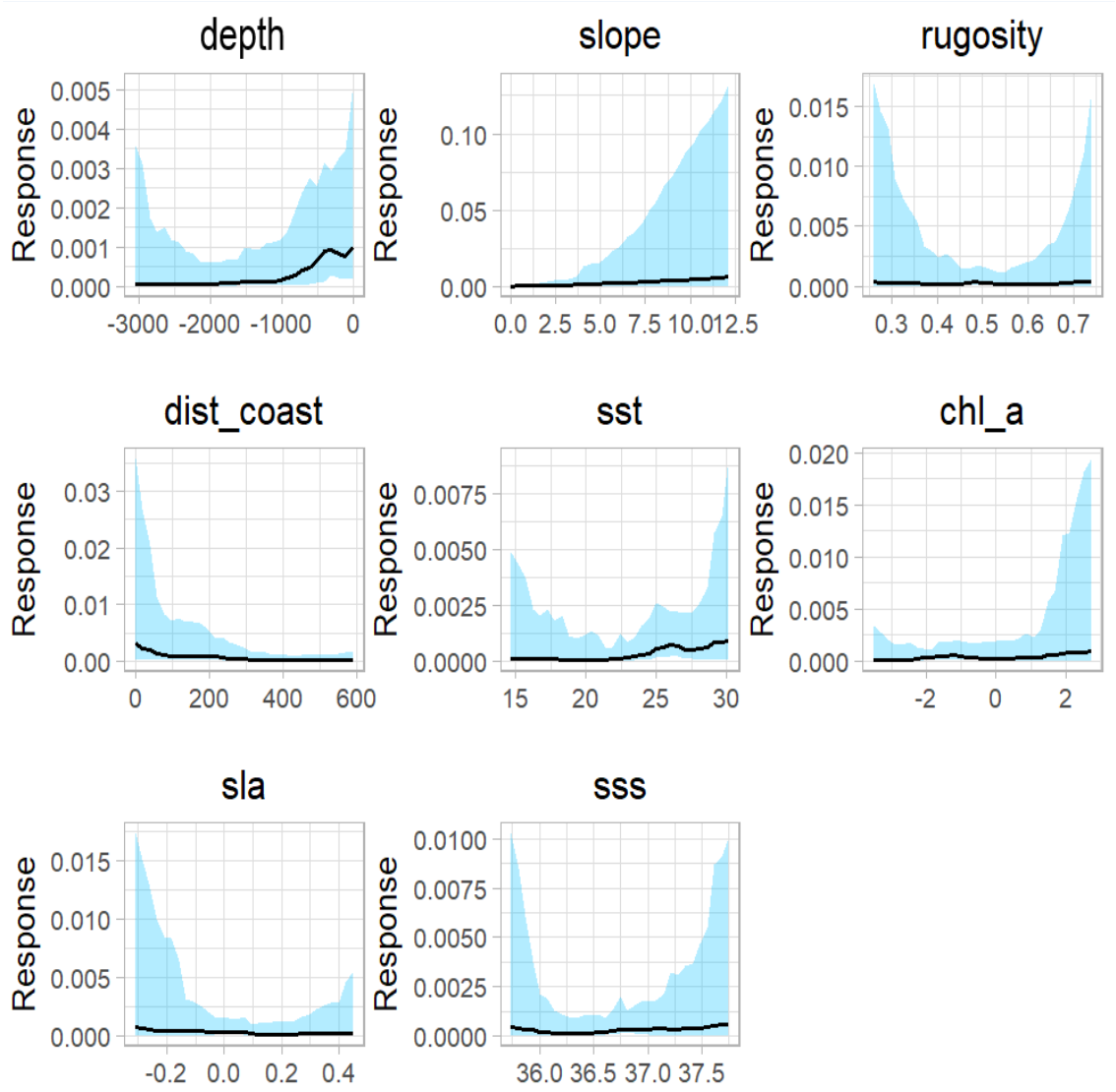
Wedekin, L.L., Neves, M.C., Marcondes, M.C.C., Baracho, C., Rossi-Santos, M.R., Engel, M.H., Simões-Lopes, P.C., 2010. Site fidelity and movements of humpback whales (*Megaptera novaeangliae*) on the Brazilian breeding ground, southwestern Atlantic. Mar. Mamm. Sci. 26 (4), 787-802. https://doi.org/10.1111/j.1748-7692.2010.00387.x

Weir, C.R., Oms, G., Baracho-Neto, C.G., Wedekin, L.L., Daura-Jorge, F.G., 2020. Migratory movement of a sei whale (*Balaenoptera borealis*) between Brazil and the Falkland Islands (Malvinas). Mar. Mamm. Sci. 36 (3), 1050-1057. https://doi.org/10.1111/mms.12687

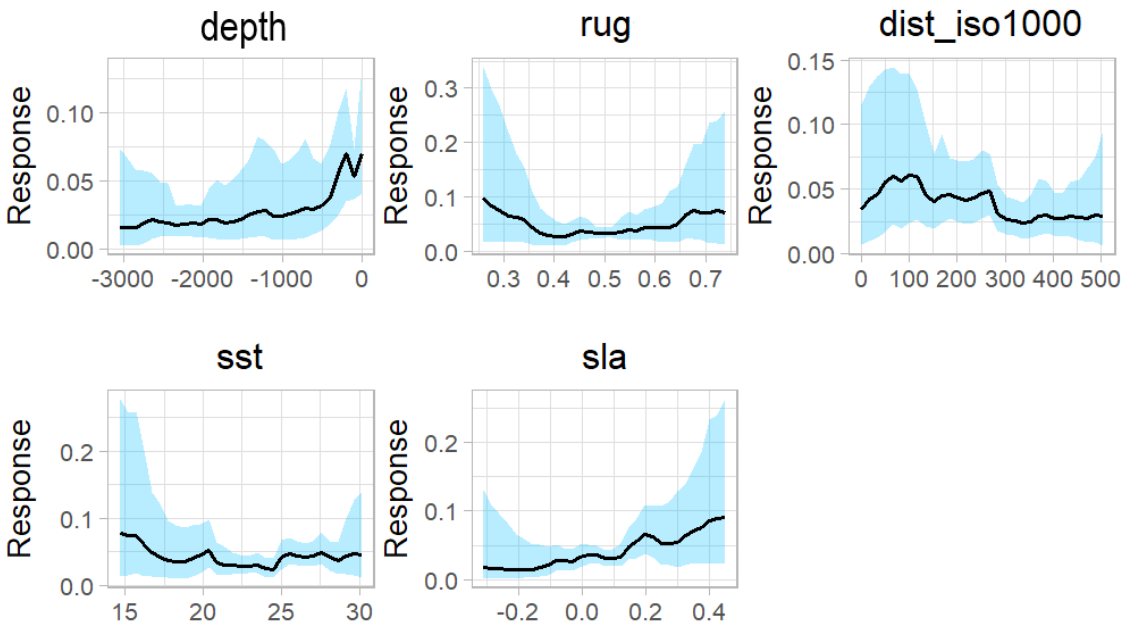
Supplementary material



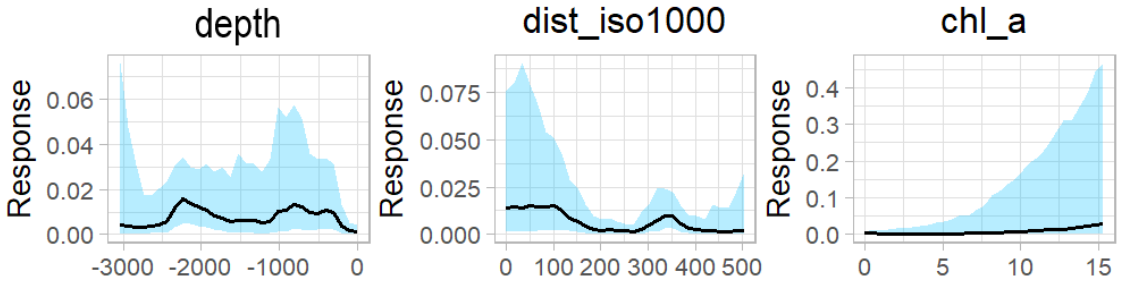
Supplementary figure S1. Partial dependence plots showing the modeled influence of each predictor on the probability of presence for mysticetes - migratory whales. Shaded light blue areas represent the 95% credible intervals.



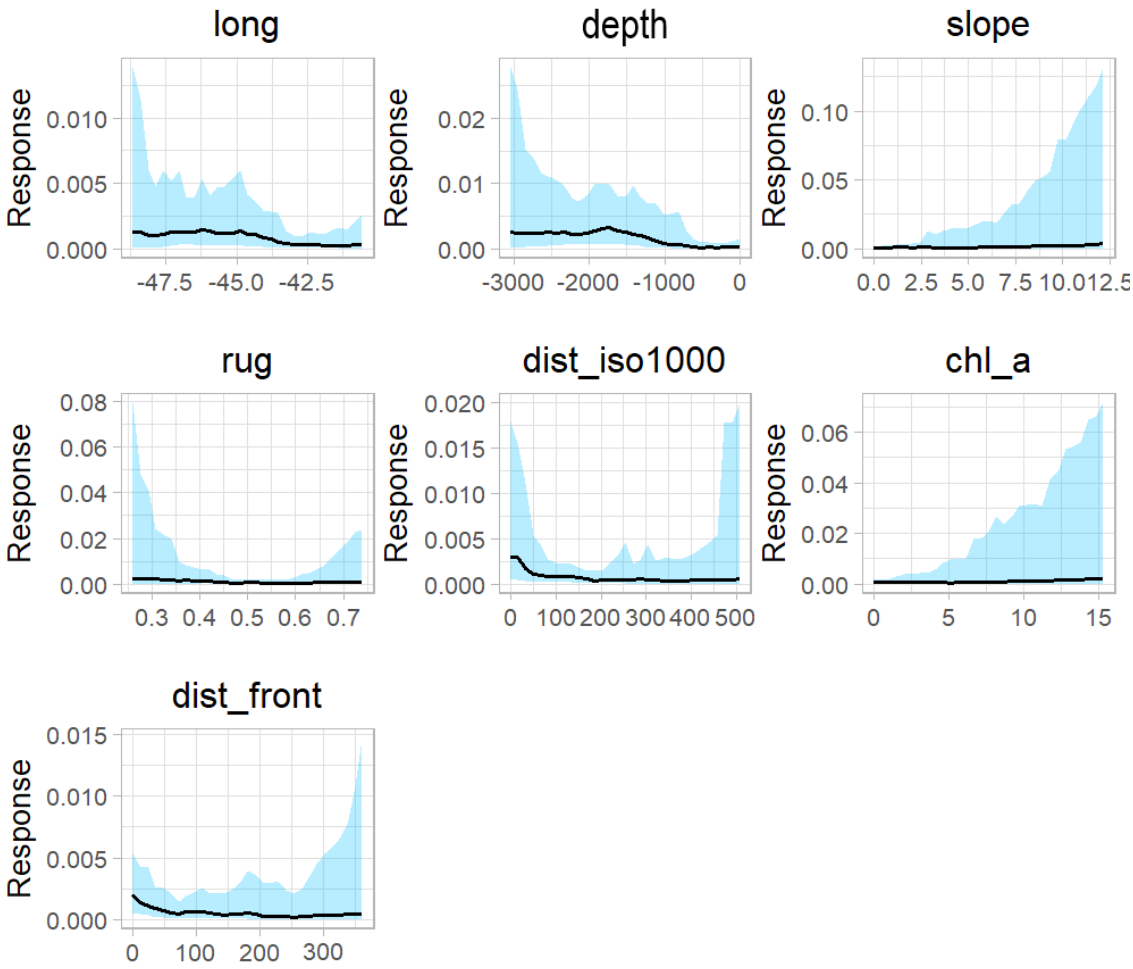
Supplementary figure S2. Partial dependence plots showing the modeled influence of each predictor on the probability of presence for mysticetes - Bryde's whales. Shaded light blue areas represent the 95% credible intervals.



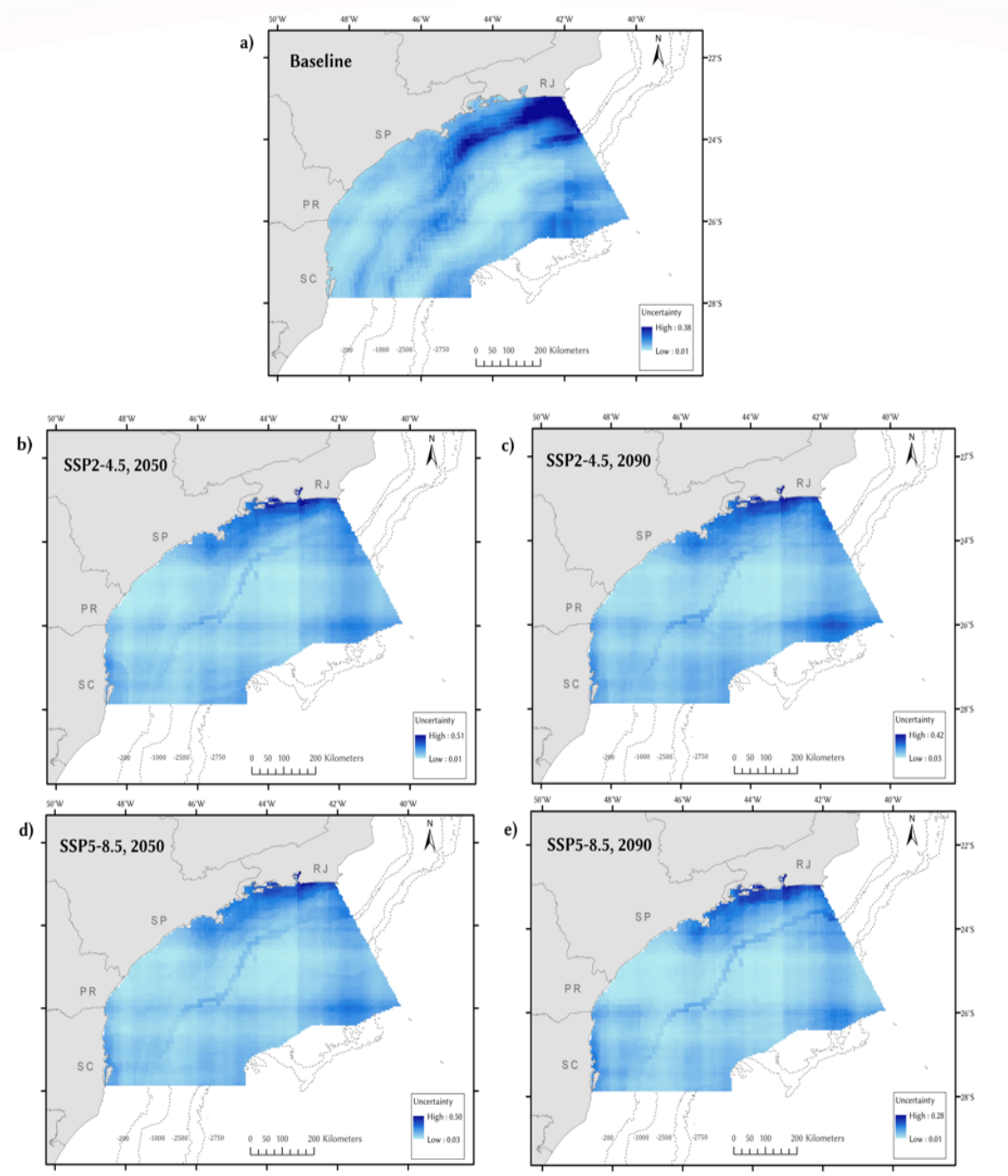
Supplementary figure S3. Partial dependence plots showing the modeled influence of each predictor on the probability of presence for odontocetes - small delphinids. Shaded light blue areas represent the 95% credible intervals.



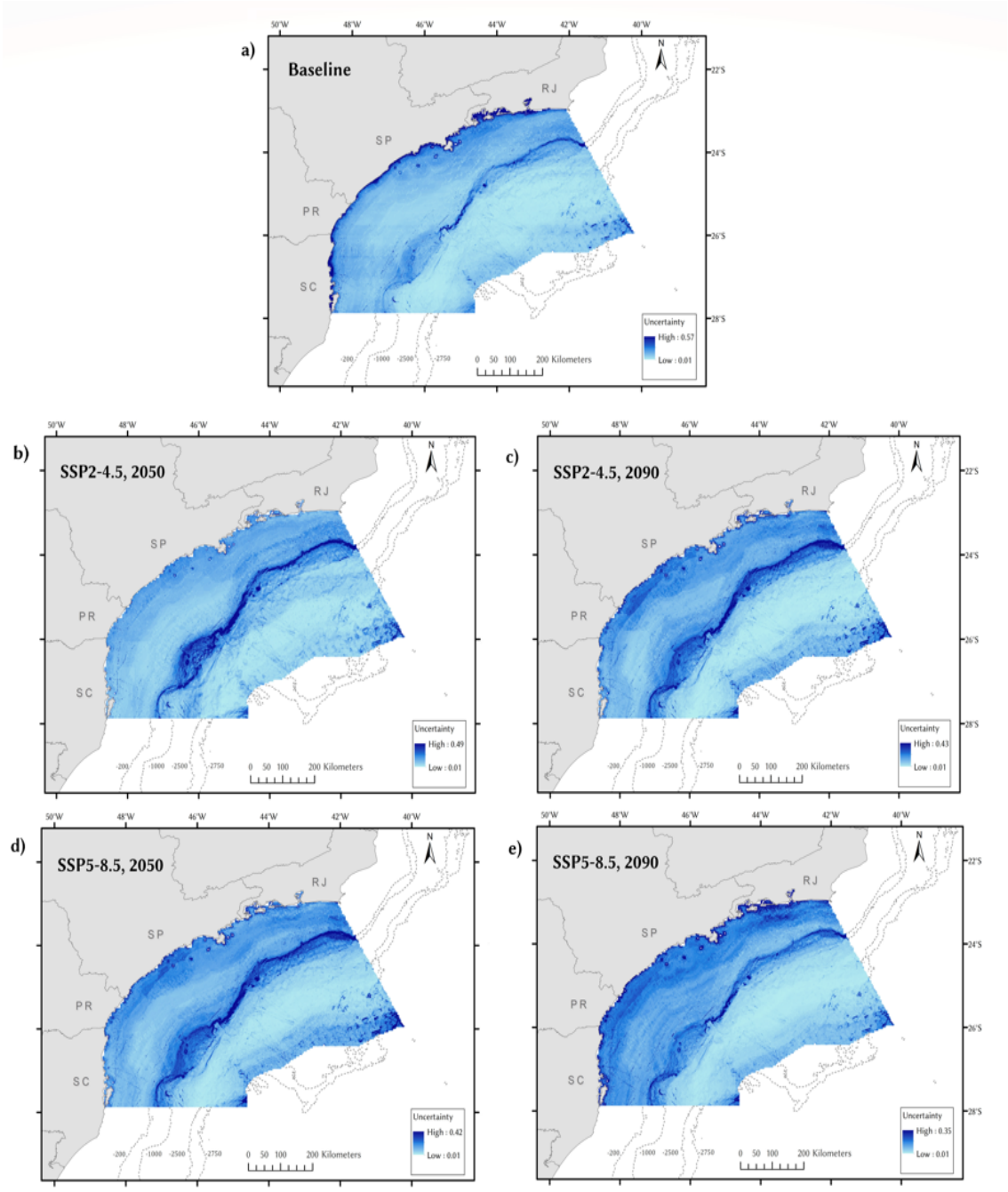
Supplementary figure S4. Partial dependence plots showing the modeled influence of each predictor on the probability of presence for odontocetes - large delphinids. Shaded light blue areas represent the 95% credible intervals.



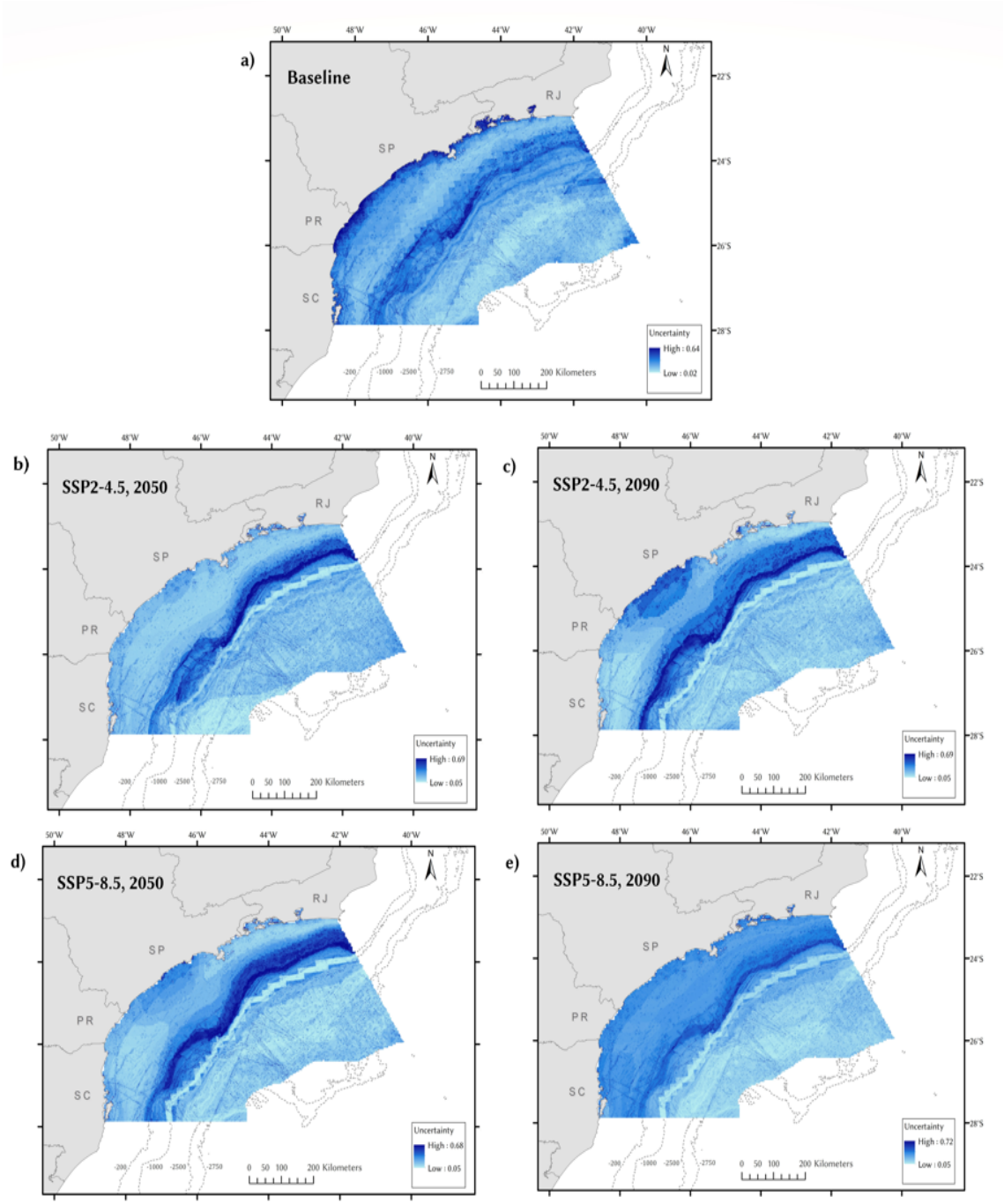
Supplementary figure S5. Partial dependence plots showing the modeled influence of each predictor on the probability of presence for odontocetes - sperm and beaked whales. Shaded light blue areas represent the 95% credible intervals.



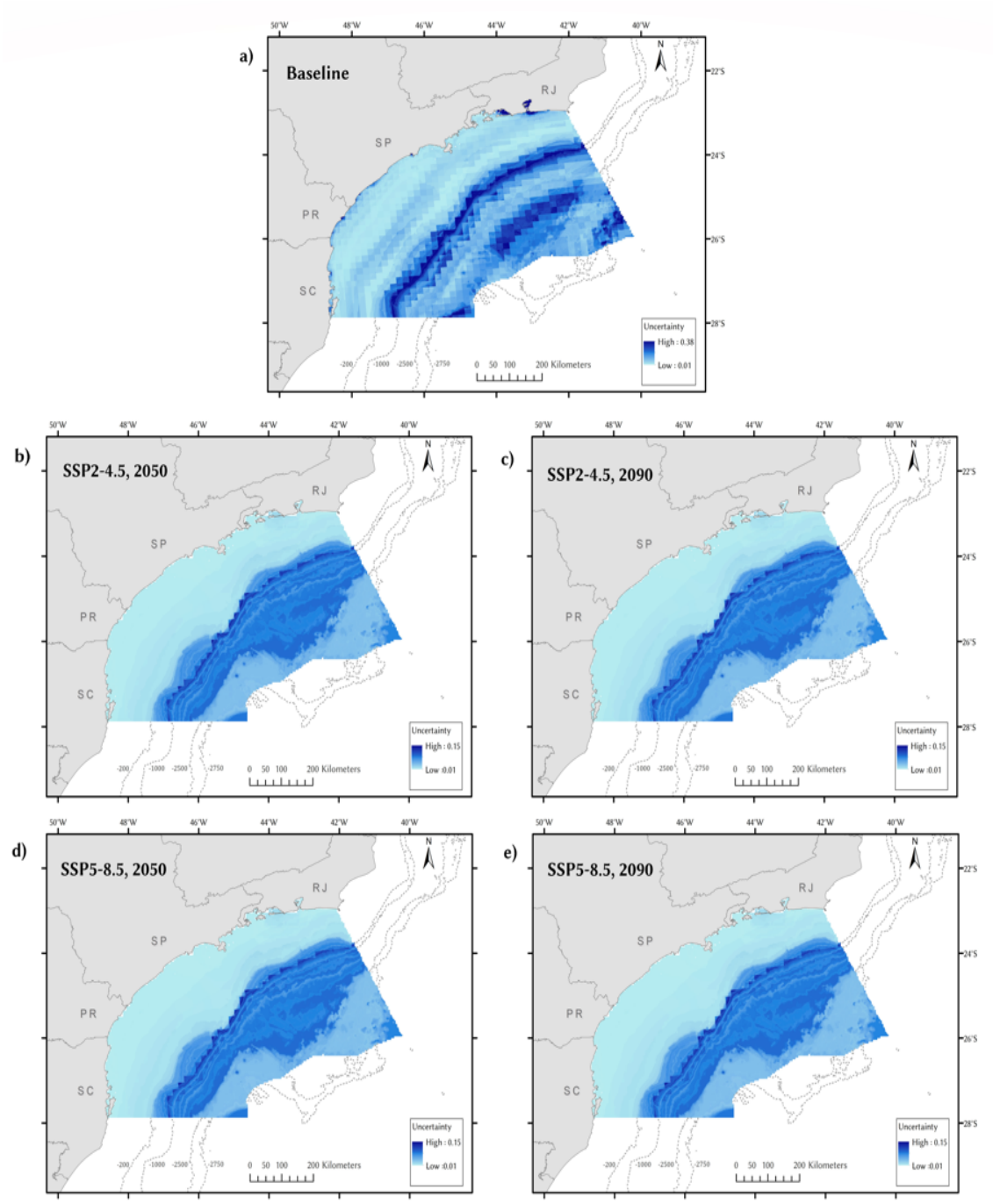
Supplementary figure S6. Maps representing uncertainties associated with baseline distribution (a) and future predicted distribution in the migratory whales group in relation to oceanographic variables in the southwestern Atlantic Ocean. Maps (b) and (c) represent predictions from SSP2-4.5 for 2050 and 2090, respectively. Maps (d) and (e) represent predictions from SSP5-8.5 for 2050 and 2090, respectively.



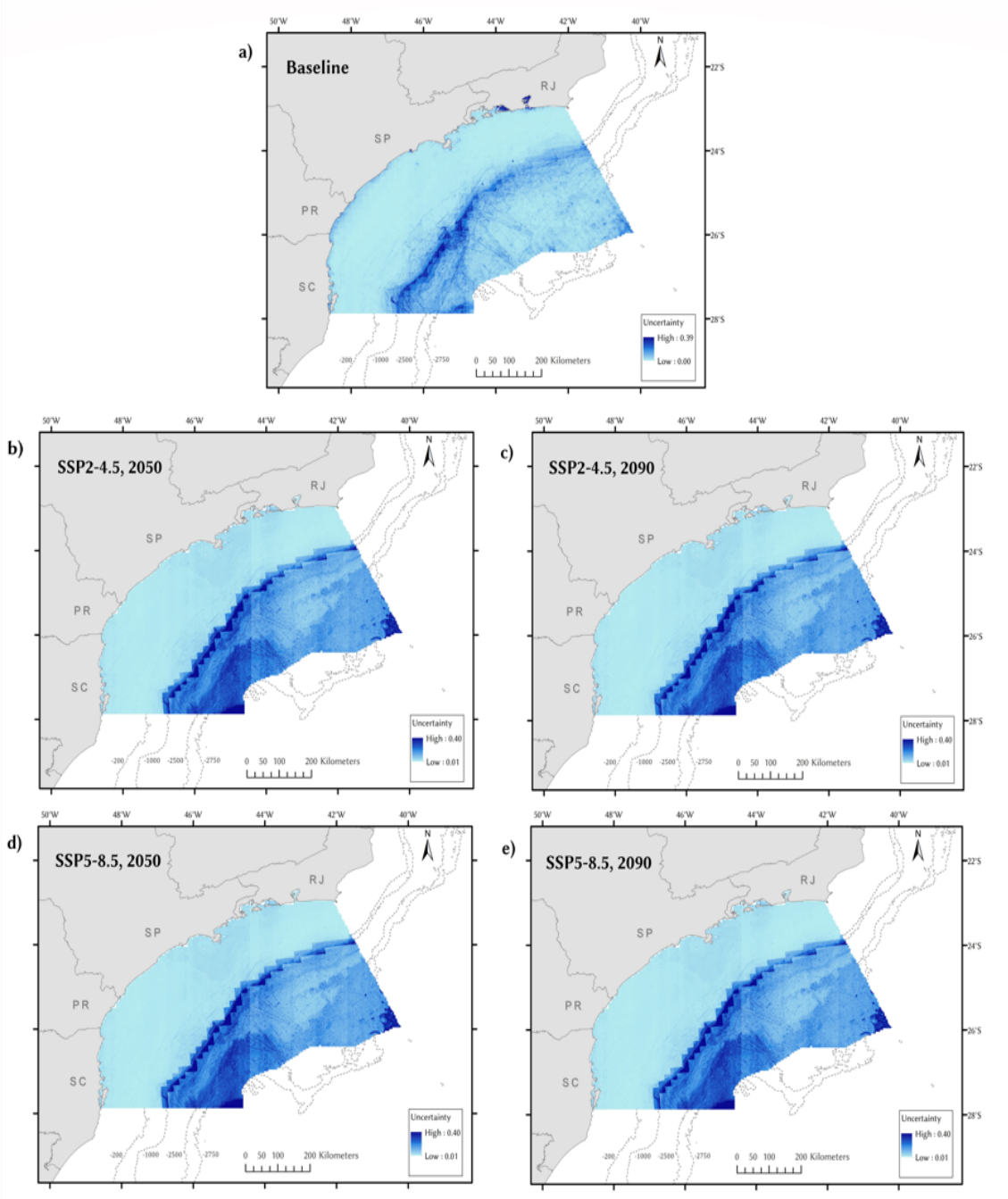
Supplementary figure S7. Maps representing uncertainties associated with baseline distribution (a) and future predicted distribution in the Bryde's whales group in relation to oceanographic variables in the southwestern Atlantic Ocean. Maps (b) and (c) represent predictions from SSP2-4.5 for 2050 and 2090, respectively. Maps (d) and (e) represent predictions from SSP5-8.5 for 2050 and 2090, respectively.



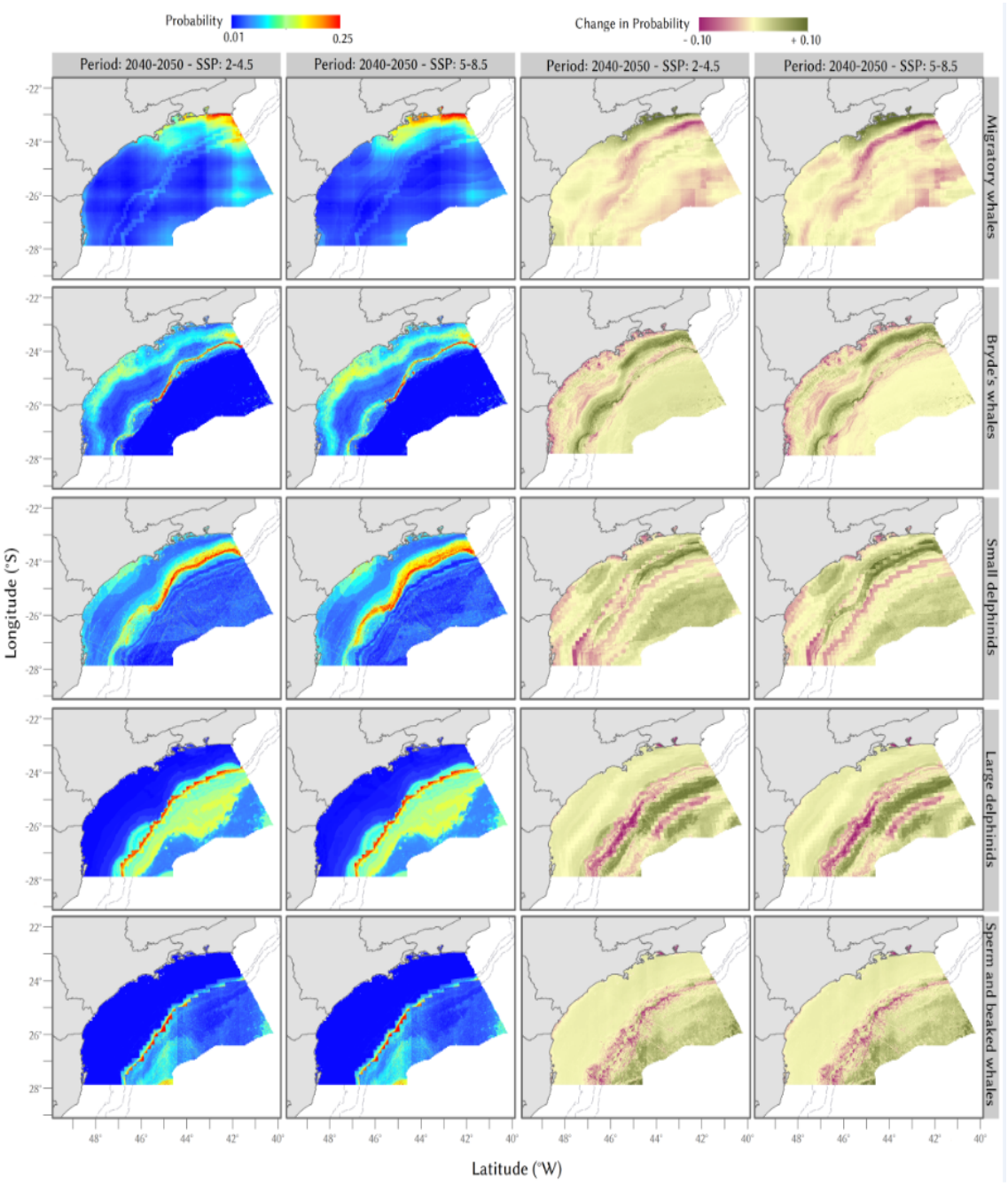
Supplementary figure S8. Maps representing uncertainties associated with baseline distribution (a) and future predicted distribution in the small delphinids group in relation to oceanographic variables in the southwestern Atlantic Ocean. Maps (b) and (c) represent predictions from SSP2-4.5 for 2050 and 2090, respectively. Maps (d) and (e) represent predictions from SSP5-8.5 for 2050 and 2090, respectively.



Supplementary figure S9. Maps representing uncertainties associated with baseline distribution (a) and future predicted distribution in the large delphinids group in relation to oceanographic variables in the southwestern Atlantic Ocean. Maps (b) and (c) represent predictions from SSP2-4.5 for 2050 and 2090, respectively. Maps (d) and (e) represent predictions from SSP5-8.5 for 2050 and 2090, respectively.



Supplementary figure S10. Maps representing uncertainties associated with baseline distribution (a) and future predicted distribution in the sperm and beaked whales group in relation to oceanographic variables in the southwestern Atlantic Ocean. Maps (b) and (c) represent predictions from SSP2-4.5 for 2050 and 2090, respectively. Maps (d) and (e) represent predictions from SSP5-8.5 for 2050 and 2090, respectively.



Supplementary figure S11. Maps depicting the future (period: 2040 - 2050) predicted distribution of the cetacean groups (migratory whales, Bryde's whales, small delphinids, large delphinids, sperm and beaked whales), as obtained through BART models.

CAPÍTULO IV

Cetacean abundance in the western South Atlantic: estimates from multi-species hierarchical distance sampling

(manuscrito ainda não submetido para revista científica)

Cetacean abundance in the western South Atlantic: estimates from multi-species hierarchical distance sampling

Ana Carolina Martins, Paul Gerhard Kinas, Leonardo Liberali Wedekin, Luciano Dalla Rosa

Abstract

Reliable estimates of population abundance are essential for biodiversity conservation and ecosystem-based management, yet they remain scarce for many marine regions and taxa, particularly in the Southern Hemisphere. We applied a multi-species hierarchical distance sampling (HDS) framework to produce the first basin-wide abundance estimates for major cetacean groups in the Santos Basin, southwestern Atlantic (Brazil). Systematic shipboard line-transect surveys conducted across four seasonal campaigns (2017–2019) were used to model detection and ecological processes separately, while explicitly accounting for imperfect detectability and spatially structured environmental drivers. Three functional groups were analysed: baleen whales, small delphinids, and large delphinids, allowing partial pooling of information and improved inference for rarely detected taxa. Detection was best described by half-normal functions, with scale parameters varying among groups, indicating substantial heterogeneity in detectability. Abundance was modelled as a function of sea surface temperature, depth, and distance to major isobaths. Model performance was evaluated using leave-one-out cross-validation. Posterior estimates revealed marked differences among groups and moderate seasonal variability. Baleen whales exhibited the strongest seasonal signal. Small delphinids were the most abundant group, and showed comparatively stable densities across seasons, whereas large delphinids presented lower and more uncertain abundance. This study provides the first basin-wide quantitative demographic baseline for cetaceans in the Santos Basin. By integrating detection uncertainty and environmental predictors within a unified Bayesian framework, these results offer spatially explicit abundance estimates that support conservation planning, environmental impact assessment, and long-term marine monitoring.

Key-words: Bayesian inference, density, distance sampling, hierarchical model, multi-species, Santos Basin, South Brazil Bight

1. Introduction

Abundance and density are fundamental demographic measures of a population, and understanding how they vary across space and time represents a central objective in ecology and conservation management. Distance sampling (DS) is a widely used framework for estimating density or population size because it explicitly models the decline in detection probability with increasing distance from the observer, thereby correcting abundance estimates for imperfect detection (Burnham et al., 1980; Buckland et al., 2004). Hierarchical distance sampling (HDS) extends classical DS by separating the observation process (detection) from the ecological process (abundance) and by allowing both processes to be modelled as functions of covariates, some of which vary across space. Such features are particularly advantageous when working with rare species, small sample sizes, or multi-species data sets (Sollmann et al., 2016).

Hierarchical models are well suited for ecological data with nested structure (e.g., individuals within species, species within groups). By partitioning variation across levels, they improve estimation and provide more reliable uncertainty quantification. Different summary measures emphasize different aspects of hierarchical variability: some quantify the proportion of variance attributable to higher levels of the model, while others describe the dispersion of group-level effects. Selecting appropriate measures is essential for correctly interpreting the dependence structure. In a Bayesian framework, hierarchical occupancy and abundance models have also proven effective at integrating imperfect detection and environmental covariates into density estimation, while allowing spatial structure to be explicitly modelled where necessary (Kéry, 2010; Pavanato et al., 2017). For multi-species distance-sampling data, the HDS framework can be extended so that each group of species has its own detection and abundance components, linked through shared hyperparameters. This design preserves group-specific inference while borrowing strength across species, enabling more robust

estimation of environmental effects and more reliable extrapolation across space and under-sampled taxa (Sollmann et al., 2016).

Cetaceans play fundamental ecological roles in marine ecosystems through trophic interactions, nutrient cycling, and ecosystem functioning (Roman et al., 2014; Ballance, 2018), while also acting as indicators of environmental change across multiple spatial and temporal scales (Hazen et al., 2019). Many cetacean populations experienced substantial declines throughout the twentieth century, largely due to intensive commercial whaling (Clapham and Baker, 2018). Although some populations have shown signs of recovery, contemporary threats including incidental capture in fisheries, vessel strikes, acoustic, chemical and plastic pollution, habitat degradation, and climate-driven environmental changes that may alter prey availability and distribution patterns continue to affect population distribution, abundance, and long-term viability (van Compernolle et al., 2025).

While the Santos Basin (SB, 23°–28° S), in the Southwestern Atlantic Ocean, is one of the most studied regions in Brazil with respect to cetaceans, there are still no published studies providing abundance estimates for the area. The basin supports a diverse cetacean assemblage distributed across marked environmental gradients, including both resident species and migratory whales that occur seasonally in the region (Di Tullio et al., 2016; Martins et al., 2026). The humpback whale (*Megaptera novaeangliae*), the most common mysticete species in the SB (Martins et al., 2023), has its wintering ground located between approximately 5° and 23° S (Andriolo et al., 2010), with the highest concentrations of groups observed over the Abrolhos Bank (16°–19° S) (Andriolo et al., 2010; Wedekin et al., 2010). Work on obtaining population estimates of humpbacks in this wintering ground region began in 1995 using mark-recapture data (Kinas and Bethlem, 1998), and estimates have since been updated using different methodological approaches (Seyboth et al., 2023), including hierarchical distance sampling (Pavanato et al., 2017). Recent estimates indicate that approximately 22,000 humpback whales occur within the wintering ground, suggesting that the population is approaching its pre-exploitation size (Zerbini et al., 2019; Bortolotto et al., 2021). Sei (*Balaenoptera borealis*), fin (*B. physalus*), common minke (*B.*

acutorostrata), Antarctic minke (*B. bonaerensis*), and Bryde's (*B. brydei*) whales also utilize the SB region (Martins et al., 2023). However, little is known about the abundance and population trends of these species in the SB, as well as throughout the Southern Hemisphere (IUCN, 2023). Although at least 15 species of delphinids have been observed in the SB (Martins et al., 2023), quantitative information on their population sizes are limited. Existing abundance estimates are largely species- and habitat-specific, such as those available for the for the Guiana dolphin (*Sotalia guianensis*) and the franciscana (*Pontoporia blainvillei*) in sheltered bays (Cremer and Simões-Lopes, 2008; Carvalho et al., 2025; De Almeida et al., 2023). However, no comparable studies have been conducted for these or other species in the adjacent coastal and offshore areas. There is an urgent need for quantitative studies to estimate the abundance and population trends of several cetacean species, as such information is essential for their effective conservation and management (ICMBio, 2011; IUCN, 2023).

This study employed an HDS framework to analyze three cetacean species groups (baleen whales, small delphinids, and large delphinids) while retaining a multi-species structure. Abundance was estimated across two hierarchical levels (detection and abundance) with environmental covariates incorporated into the abundance component. Considering three main cetacean groups in the Santos Basin, rather than individual species, allows for a macro-ecological perspective on abundance patterns and the factors influencing general population densities. Aggregating species with similar ecological requirements increases statistical power and allows robust estimation of density and abundance even for rarely detected species (Royle and Dorazio, 2008). While species-specific abundance patterns may be diluted, these macroanalyses provide a broad view of the community, capturing consistent trends across groups and reflecting the cumulative distributions of their constituent species (LeDuc, 2018). Such an approach facilitates ecosystem-based management and conservation by providing comprehensive information on the cetacean community.

2. Methods

2.1. Data acquisition and processing

2.1.1. Survey effort and sighting data

Cetacean surveys were conducted in the Santos Basin, from the coastline to the Brazilian Exclusive Economic Zone, during systematic shipboard campaigns of the "Projeto de Monitoramento de Cetáceos da Bacia de Santos" (PMC-BS). Line-transects were oriented perpendicular to the coast, crossing the main bathymetric contours to ensure broad spatial coverage. Transect positions were shifted in each survey to cover the area as homogeneously as possible. The vessel travelled at ~7 knots, and observations were made during daylight hours by rotating teams of three experienced marine mammal observers positioned about 7 m above sea level. A 'passing mode' approach was employed, whereby animals were counted along the transect line without stopping or deviating from the predetermined route. For each sighting, basic environmental data (i.e. Beaufort sea state and visibility), time, position, species, group size and composition, calf presence, and behaviour were recorded. In this study, we analysed data from four surveys (5–8) (Table 1), each conducted primarily during a different season, which are part of the eight systematic campaigns previously described in Martins et al. (2023).

Table 1. Survey periods and summaries for cetacean sightings and search effort by cruise.

Survey	Year	Period	N sightings	Effort (h)	Effort (km)	Mean segment length (km)
5	2017	25 Aug - 07 Oct	94	435,7	5583,8	10.0
6	2018	12 Jan - 20 Feb	44	444,3	5863,4	9.98
7	2018	01 Sep - 20 Oct	110	477,2	5930,5	9.99
8	2019	13 Mar - 26 Apr	58	425,3	5502,3	10.0
Total	-	-	306	1782,5	22880	-

Perpendicular detection distances were estimated following standard distance-sampling protocols (Buckland et al., 2004). Reticle binoculars were used to determine visual detection distances by counting the number of reticles between the horizon and each sighting. Radial distances were then calculated using a formula that accounts for the observer's height above sea level and the curvature of the Earth (Lerczak and Hobbs, 1998; Kinzey and Gerrodette, 2001). The horizontal sighting angle, measured with an angle board, was subsequently combined with the radial distance to derive the perpendicular distance of each cetacean group from the transect line.

Effort tracklines were segmented into 10-km units to allow the assignment of spatially explicit environmental covariates to each sampling unit for analysis. Three species-group datasets were constructed to support abundance estimation: baleen whales, small delphinids, and large delphinids (see species composition in Table 2), grouped based on phylogeny and ecological–behavioral traits. Each sighting was matched to the centroid of the nearest on-effort segment surveyed. Although sperm whales (*Physeter macrocephalus*) and beaked whales (*Kogia sima* and *Ziphius cavirostris*) are present in the study area (Martins et al., 2023), their low detection frequency precluded reliable estimation of abundance; consequently, they were not included in the quantitative analyses.

A polygon was constructed to represent the study area within the Santos Basin, extending from the shoreline to the boundary of the Brazilian Exclusive Economic Zone (EEZ; Flanders Marine Institute, 2018) and between latitudes 23°S and 28°S (Figure 1). This delineation encompasses the continental shelf, slope, and adjacent offshore areas within the study region. The polygon's area was subsequently calculated to define the spatial extent over which total abundance estimates were derived.

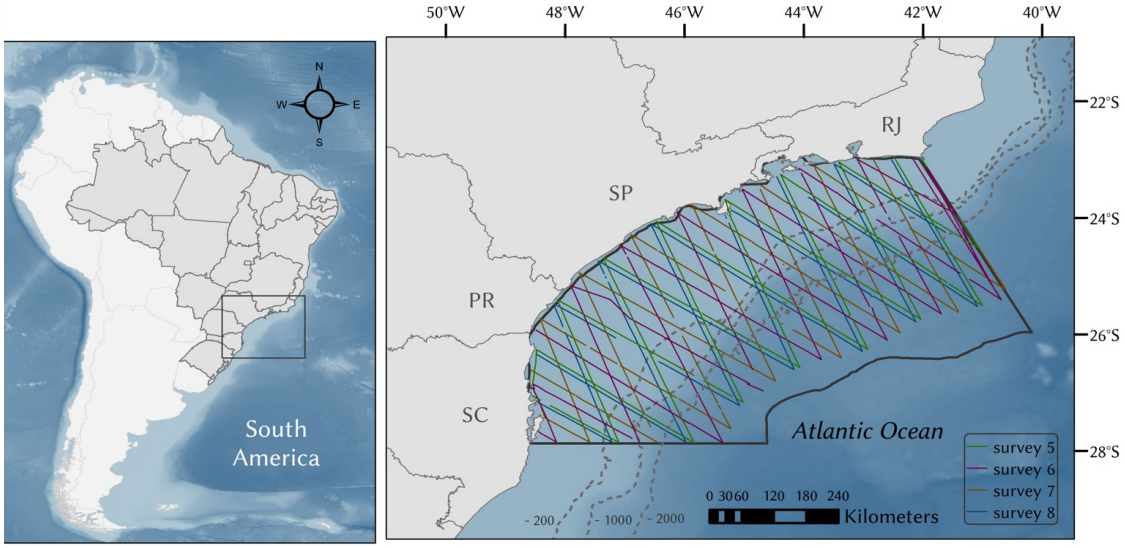


Figure 1. Spatial extent of the Santos Basin study area (southeastern Brazil), defined by a polygon spanning from 23°S to 28°S and extending from the shoreline to the Brazilian EEZ boundary. Brazilian states: Rio de Janeiro (RJ), São Paulo (SP), Paraná (PR) and Santa Catarina (SC). Bathymetric contours (isobaths) of 200 m, 1000 m, and 2000 m are shown.

2.1.2. Environmental data

Covariates were selected based on prior ecological knowledge and their relevance for cetacean distribution in the Santos Basin (Martins et al., 2023; Martins et al., 2026). Environmental variables including sea surface temperature (SST), chlorophyll-a, depth, and distance to isobaths were tested at the abundance level, while sea state (Beaufort scale) was evaluated as an explanatory variable for detection probability. Dynamic variables were obtained from NASA's Ocean Biology Processing Group (https://oceandata.sci.gsfc.nasa.gov), and bathymetric data were derived from GEBCO (https://www.gebco.net/data_and_products/gridded_bathymetry_data). Different combinations within this predefined set were evaluated.

Weekly values of environmental variables were associated with the latitude and longitude at the centre of each segment. Geoprocessing and sampling of covariates were performed using ArcGIS 10.6.1 (ESRI, 2018).

2.2. Analytical methods

2.2.1. *Multi-species hierarchical distance sampling model*

For each survey campaign, the optimal truncation distance was determined separately for each species group to remove excessively large perpendicular distance values. Following Buckland et al. (2001), truncation thresholds were set between 5% and 10% of the most distant observations, thereby balancing data retention and model robustness.

To meet sample size requirements for reliable detection function estimation (Buckland et al., 2001), a global detection function was initially fitted using data pooled across all survey campaigns (Supplementary Table 1). Two key functions were considered: the hazard rate and the half-normal. The choice of detection function considered both, the shape of the decline in detectability with distance, and model parsimony. Half-normal and hazard-rate functions were fitted for each species group and compared based on visual assessment of fit, information criteria, and the ecological plausibility of the resulting detection curves. The half-normal provided more stable fits for groups with limited detections and was therefore adopted for subsequent analyses. Consistent with standard distance sampling assumptions (Buckland et al., 2001), detection was modeled conditional on perfect detectability at zero distance (i.e., $g(0) = 1$). The resulting parameter estimates (σ) were then incorporated into informative priors in campaign-specific models to improve parameter stability in datasets with limited detections (Table 1).

The hierarchical multi-species distance sampling framework of Sollmann et al. (2016) was adapted for implementation in the *ubms* R package (Kellner et al., 2022) using the stan_distsamp() function. Detection probability was modeled as a function of species group and Beaufort sea state, while abundance included an offset for surveyed transect length, species, and environmental covariates. Weakly informative priors were employed to regularize parameter estimation without imposing strong assumptions (Table 3). Priors for the state model intercept and coefficients were centered at zero with sufficiently large standard deviations to allow a wide range of plausible parameter values, while priors for detection parameters were slightly more restrictive to stabilize

the estimation of the detection function. This modeling framework was fitted with three Markov chains, 4000 iterations each (including 1000 warmup iterations), ensuring adequate exploration of the posterior distribution. Weakly informative priors of this type are commonly applied in Bayesian hierarchical ecological models to promote stable and interpretable inference (Kéry and Royle, 2016). The inclusion of an offset term allows direct accounting for variable survey effort across transects, while species-specific effects capture heterogeneity in abundance and detection.

Table 2. Composition and summary of sightings for each cetacean group.

Group	Taxon	N sightings	N individuals
Baleen whales	*Balaenoptera acutorostrata*	1	1
	Balaenoptera bonaerensis	1	1
	Balaenoptera borealis	7	23
	Balaenoptera brydei	4	7
	Balaenoptera physalus	2	2
	Megaptera novaeangliae	82	169
	Unidentified whale	75	90
Total	-	172	293
Small delphinids	*Delphinus delphis*	2	35
	Lagenodelphis hosei	1	30
	Sotalia guianensis	3	11
	Stenella sp.	5	50
	Stenella attenuata	7	253
	Stenella clymene	5	72
	Stenella frontalis	50	1605
	Stenella longirostris	7	1055
	Steno bredanensis	3	24
	Tursiops truncatus	24	598
Total	-	107	3733

	Feresa attenuata	3	49
	Globicephala macrorhynchus	5	88
Large delphinids	*Grampus griseus*	6	79
	Orcinus orca	3	4
	Peponocephala electra	3	620
	Pseudorca crassidens	1	15
Total	-	21	855
	Total	**300**	**4881**

To formally describe the detection process, we modeled the probability of detecting a group at perpendicular distance x using a half-normal detection function:

$$g(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right),$$

where σ is the scale parameter of the detection function, determining the spread of the detection curve (Buckland et al., 2001). Perpendicular distances were grouped into K equal-width distance intervals $\left[b_{k,}\, b_{k+1}\right]$, each of width $v = w/K$, where w denotes the truncation distance. In our case, $K = 4$, resulting in four distance bins of equal width:

$$v = \frac{w}{4}.$$

The detection probability within the k-th distance interval was therefore calculated as:

$$p_k = \frac{\int_{b_k}^{b_{k+1}} g(x)\, dx}{v}.$$

To ensure abundance was expressed on a per-area basis, transect length was converted to an effective survey area $2\ \sigma_g L$, where L is transect length and σ_g is the detection

scale parameter. This term was included as a log-offset in the abundance model, allowing estimates to be interpreted directly as the number of groups per km^2.

At the transect level, the observed number of detected groups at site i (n_i) was linked to the latent number of groups N_i through a binomial observation model:

$$n_i \sim Binomial\ (N_i\ ,\ p_{t_i}),\ \text{where}\ p_{t_i} = \sum_{K\ =\ 1}^{K} p_{k_i}\ .$$

Latent abundance was modeled as:

$$N_{i,g} \sim Poisson\ (\lambda_{i,g})$$

with expected abundance expressed as a function of environmental covariates:

$$\log(\lambda_{i,g}) = \alpha_{0,g} + \alpha_g^T x_i.$$

Variation in detectability was modeled through the distance decay parameter:

$$\log(\sigma_i) = \beta_{0,g} + \beta_g^T y_i.$$

To extend this formulation to a multi-species community framework, all parameters were indexed by species group $g = 1, \ldots, 3$. Accordingly, $N_{i,g}$ e $\lambda_{i,g}$ describe the abundance of group g at site i, and species-specific intercepts and coefficients (e.g., $\alpha_{0,g}$, $\beta_{0,g}$) were assigned hierarchical hyperpriors. This partial pooling structure allows information sharing among species groups while preserving biologically meaningful variation in abundance and detection patterns (Sollmann et al., 2016).

Table 3. Prior distributions assigned for model parameters.

Parameter	Prior	Description
α_0	N(0, 1.5)	Prior for the log-scale abundance intercept
α_i	N(0, 1)	Priors for the effects of covariates on abundance
β_0	N(0, 1)	Prior for the log-scale detection intercept
β_i	N(0, 0.7)	Priors for the effects of covariates on detection
σ_g	N(μ_g, 0.5)	Priors based on the posterior median estimates from the global pooled detection function

Different combinations of the predefined covariate set were then evaluated to identify the most parsimonious relationships between environmental predictors and species abundance. This strategy enhances interpretability and reduces the likelihood of spurious associations, while still allowing the model to capture ecologically meaningful patterns (van de Pol et al., 2016).

Detection and abundance parameters were estimated separately for each survey campaign, with informative priors for σ derived from pooled data to improve parameter stability in datasets with limited numbers of detections. Model selection was conducted using the Leave-One-Out Information Criterion (LOOIC), which provides a robust measure of predictive fit for complex Bayesian models or sparse datasets and is considered more reliable than WAIC (Vehtari et al., 2017). Models with lower LOOIC values were deemed better suited to capture variation in cetacean abundance as a function of environmental covariates.

2.2.2. *Abundance estimates*

After fitting the hierarchical distance-sampling model to the number of groups, we extracted posterior samples of the state model coefficients to estimate group-specific densities and abundances. Expected densities (individuals x km^{-2}) were computed in two steps. First, group-densities resulted from the exponential transform of posterior draws from the regression coefficients of the linear predictor. These group-densities were then multiplied by the total surveyed area to obtain posterior draws of total abundance for each group. At a second step, to account for uncertainty and variability in group size, gamma-distributed random effects were applied to each posterior sample of abundance. Final posterior distributions of detection-corrected abundance were summarized using medians, median absolute deviation (Gelman et al., 2013) and 95% credible intervals.

3. Results

3.1. *Data overview*

The study considered four surveys carried out between 2017 and 2019 (Table 1), each spanning 25 to 50 days. The total survey effort amounted to 1782.5 hours of observation and 22,880 km of transects, with an average segment length of approximately 10 km. Cetacean sightings per survey ranged from 44 (Survey 6, summer) to 110 (Survey 7, spring).

In total, 300 sightings comprising 4881 individuals were recorded (Table 2). Baleen whales accounted for 172 sightings (293 individuals), with humpback whales (*Megaptera novaeangliae*) being the most frequently observed species in this group (82 sightings; 169 individuals). Small delphinids were the most frequently encountered group (107 sightings) and contained the highest number of individuals (3733); this group was dominated by *Stenella frontalis* (50 sightings; 1605 individuals) and *Stenella longirostris* (7 sightings; 1055 individuals), which together accounted for over 71% of all small delphinid individuals. Large delphinids were less frequently observed, totaling 21 sightings of 855 individuals, with *Peponocephala electra* being the most abundant species (3 sightings; 620 individuals). Sightings distribution varied across surveys and functional groups, reflecting seasonal and spatial differences in cetacean occurrence.

3.2. *Model selection and parameter estimation*

The inclusion of environmental covariates generally improved the fit of abundance models compared to the null model (Supplementary Table 2). Incorporating Beaufort improved LOOIC values but resulted in unstable MCMC chains ($\hat{R} > 1.03$ and low effective sample sizes for some parameters), except for survey 8. Therefore, we selected the model without Beaufort as the main model for these campaigns, ensuring reliable and interpretable abundance estimates, while the difference in LOOIC between models was small and is discussed in a sensitivity analysis. All selected models exhibited $\hat{R}$ values close to 1, indicating adequate MCMC convergence (Vehtari et al., 2021), along with satisfactory effective sample sizes. Posterior summaries of the

parameters describing abundance (state) and detection (observation) processes are presented in Supplementary tables 3 and 4, respectively.

3.3. *Abundance estimates of cetacean groups*

Detection functions were fitted using a half-normal model to perpendicular distances recorded for each group of species. Figure 2 shows histograms of the observed distances overlaid with the estimated detection curves. The curves indicate that detection probability declined at distinct rates with increasing distance from the transect line across all groups. Estimated scale parameters (σ) were 1.387 km for baleen whales; 0.731 km for small delphinids; and 0.874 km for large delphinids; reflecting differences in detectability among groups. Most detections occurred close to the transect line, as expected, and the fitted half-normal curves adequately emulated the shapes of the observed distance distributions, supporting the suitability of these detection functions for subsequent abundance estimation.

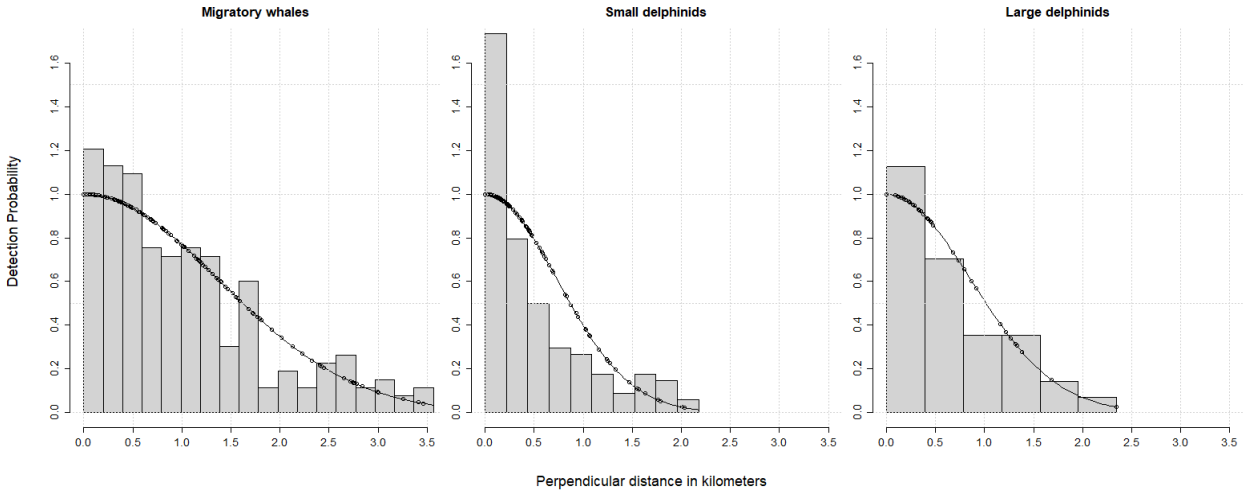


Figure 2. Half-normal detection functions fitted to perpendicular distances. Black lines indicate the detection function fitted to the posterior median; bars represent the observed distribution of perpendicular distances (in km).

Posterior estimates of group-specific abundance revealed clear differences among cetacean groups and moderate seasonal variability across surveys (Figure 3, Table 4). Baleen whales (a) exhibited the most marked variation, with posterior medians increasing from C5 (winter) to a peak in C7 (spring), followed by a decline in summer that continued through C8 (fall). Median abundance estimates varied from 483 in summer (C6) to 3598 in spring (C7). In contrast, small delphinids (b) maintained consistently high abundance across campaigns (18,718 – 36,101 individuals), with broadly overlapping credible intervals and only a slight increase during summer (C6). Large delphinids (c) showed intermediate abundance (1400 – 10,568) but wider credible intervals, reflecting higher uncertainty due to fewer detections. In general, these patterns, consistent with the group-specific overall mean abundance indicated by the dashed line in Figure 3, highlight stronger seasonal fluctuations for baleen whales and large delphinids, whereas small delphinids remained comparatively stable.

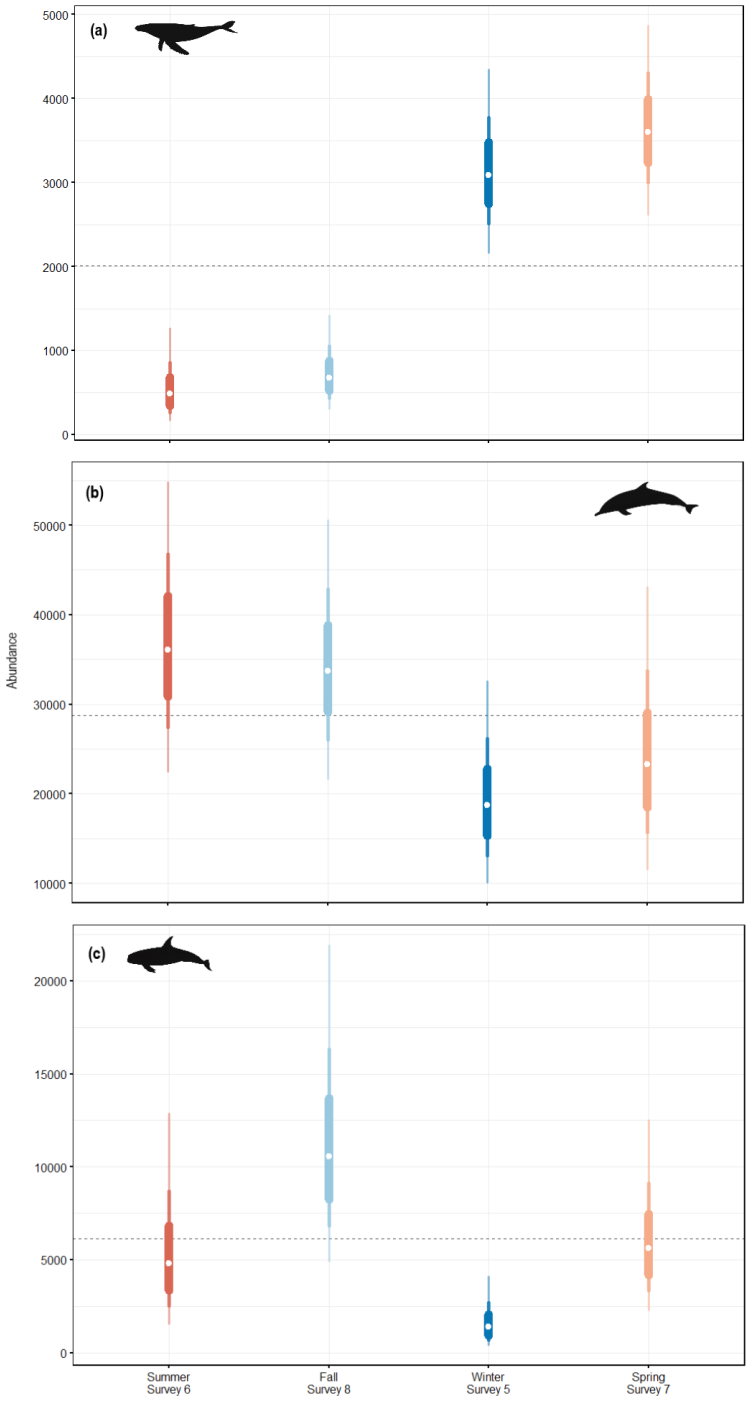


Figure 3. Posterior distributions of cetacean group-specific abundance estimates for the four seasonal survey campaigns (Winter = Survey 5, Summer = Survey 6, Spring = Survey 7, Fall = Survey 8). Vertical bars represent the 50% (thick) , 75% (medium), and 95% (slim) Bayesian credible intervals, with white points indicating the posterior medians. The group-specific overall mean abundance across all campaigns is displayed by the dashed horizontal lines. Panels correspond to (a) baleen whales, (b) small delphinids, and (c) large delphinids.

Table 4. Summary of parameters (Point estimate - median, CV_{MAD} - coefficient of variation based on median absolute deviation, and 95% credibility interval) estimated through the hierarchical multi-species distance sampling. E(s) is expected group size, N is abundance for the groups 1 = baleen whales; 2 = small delphinids; and 3 = large delphinids.

Survey	Parameter	Point estimate	CV_{MAD} (%)	95% CI
5	N_1	3087	11.94	2163 - 4343
	N_2	18,718	19.89	10,073 - 32,569
	N_3	1400	38.37	402 - 4113
	$E(s_1)$	1.84	3.93	1.65 - 2.07
	$E(s_2)$	21.52	3.36	19.49 - 23.63
	$E(s_3)$	4.15	16.33	2.57 - 6.54
6	N_1	483	33.82	168 - 1263
	N_2	36,101	15.46	22,478 - 54,838
	N_3	4810	34.74	1551 - 12,881
	$E(s_1)$	1.55	13.61	1.13 - 2.40
	$E(s_2)$	27.00	2.11	25.41 - 28.73
	$E(s_3)$	13.20	10.15	9.73 - 17.45
7	N_1	3598	10.57	2619 - 4870
	N_2	23,274	22.52	11,529 - 43,121
	N_3	5632	28.70	2281 - 12,505
	$E(s_1)$	1.86	3.64	1.51 - 2.33
	$E(s_2)$	22.89	2.27	21.43 - 24.40
	$E(s_3)$	11.82	6.38	9.82 - 14.19

	N_1	675	25.78	307 - 1419
	N_2	33,687	14.33	21,687 - 50,564
8	N_3	10,568	24.89	4913 - 21,926
	$E(s_1)$	1.10	6.00	1.00 - 1.50
	$E(s_2)$	25.92	2.97	23.75 - 28.28
	$E(s_3)$	47.78	3.39	43.16 - 52.59

4. Discussion

This study provides the first comprehensive estimates of cetacean abundances for the entire Santos Basin. By jointly estimating abundance for baleen whales, small delphinids, and large delphinids, our results establish a demographic and ecological baseline that is essential for detecting future changes in abundance, habitat use, and community structure. Such baseline information is particularly valuable in regions undergoing rapid environmental change and increasing human use, as it enables future assessments of population responses and supports evidence-based conservation planning. Our estimates suggest potential seasonal differences in abundance, particularly for baleen whales and large delphinids. These findings rely on a single survey implemented in each season within our modeling framework, and thus would benefit from additional analyses and monitoring efforts to further validate and refine these temporal patterns. Moreover, the inclusion of environmental covariates revealed preliminary ecological signals, indicating how temperature and seascape features may influence distribution patterns, although further sampling is needed to confirm and strengthen these associations. Importantly, the posterior distributions obtained here could help shape informative priors in future hierarchical models, contributing to progressively more precise and robust estimates as additional data are collected.

The Santos Basin may function as a transit area, but also as a region of temporary residence for large migratory whales. These species concentrate primarily in

the northern portion of the basin - near the southern limit of the humpback whale breeding range (Andriolo et al., 2010) - and along the continental slope (Martins et al., 2026). Our results indicate that more than 3,000 migratory baleen whales likely use the region. The seasonal variation in baleen whale abundance in the Santos Basin is largely driven by humpback whales, reflecting well-documented migratory patterns along the western South Atlantic. The Santos Basin may support a meaningful portion of the humpback whale population that breeds farther north, currently estimated at ~22,000 individuals (Bortolotto et al., 2021). Although humpback whales show signs of demographic recovery, with population size approaching pre-exploitation levels (Zerbini et al., 2019; Bortolotto et al., 2021), their long-term use of the Santos Basin may be affected by changing environmental conditions. Recent studies suggest a potential reduction in suitable habitat under projected environmental scenarios (von Hammerstein et al., 2022; Martins et al., 2026), highlighting the importance of continuous monitoring even for populations that appear demographically recovered. For other migratory species occurring in the region - sei, fin, and minke whales - the absence of abundance estimates along the Brazilian coast underscores the urgency of expanding survey efforts to better understand their regional status and trends, as well as their contribution to the total migratory assemblage. The Bryde's whale (*Balaenoptera brydei*), the only non-migratory baleen whale in Brazil, is commonly observed year-round in the Santos Basin (Dalpaz et al., 2023). Nevertheless, the surveys addressed here did not yield sufficient data to allow for an abundance estimate of this species. Owing to its distinct ecological and behavioral characteristics, Bryde's whale data were not pooled with those of other baleen whales. Given its predominantly coastal distribution (Martins et al., 2026) and exposure to multiple anthropogenic impacts, it is also crucial to estimate their abundance in the region, particularly because baseline abundance information is currently unavailable for evaluating future population trajectories and regional conservation status.

For delphinids, our abundance estimates reveal pronounced differences within the delphinid assemblage: small delphinids were substantially more numerous than large delphinids, a pattern widely reported in ecosystems of the southwestern Atlantic and across other temperate and subtropical coastal–oceanic systems (Di Tullio et al., 2016;

Roberts et al., 2016). This is consistent with the broader ecological tendency for small delphinids to form large schools and exploit a wide range of habitats, resulting in higher local densities (Würsig, 2019). In the Santos Basin, Martins et al. (2026) showed that small delphinids (including both coastal and oceanic species) predominantly use shallow waters, the outer continental shelf, and the upper slope, whereas large delphinids are closely associated with deeper oceanic habitats around the 1000-m isobath, particularly in the northern sector where depths of 2000–3000 m appear especially suitable. Despite this spatial segregation, the two groups occasionally overlap and form mixed-species associations (Lima et al., 2021). Although small delphinids dominated the assemblage, detection bias may accentuate these differences. Large delphinids generally occur in smaller groups, display more cryptic surfacing behavior, and frequently occupy deep offshore waters - traits known to reduce detectability in line-transect surveys (Wedekin et al., 2014; Würsig, 2019; Rankin et al., 2020). Thus, their lower encounter rates likely reflect a combination of genuine ecological patterns and methodological limitations. In addition to these spatial and detectability contrasts, seasonal trends also differed between groups. Large delphinids exhibited stronger fluctuations in abundance across surveys, with a potential increase during fall, although uncertainty remains high due to relatively few detections. In contrast, small delphinids maintained consistently high abundance across campaigns, with broadly overlapping credible intervals, although estimates showed a slight increase during summer. These results reinforce that ecological responses to seasonal drivers may vary within the delphinid community of the Santos Basin and highlight the importance of continued temporal monitoring to better characterize seasonal dynamics and to detect potential ecological changes affecting delphinid population abundance. For delphinids, projections suggest that environmental conditions in the basin may remain generally suitable for the presence of most groups (Martins et al., 2026). Nonetheless, broader ecological changes - such as potential shifts in prey availability or increased interspecific competition - could alter community structure and influence population trajectories. The abundance estimates presented here provide an initial quantitative reference against which such future changes may be evaluated.

In line transect distance sampling, it is not required that animals are uniformly distributed in space. Instead, unbiased inference relies on transects being placed randomly (or systematically with a random start), ensuring that all individuals in the study area have a non-zero probability of inclusion in the sample (Buckland et al., 2001). In our study, transects were placed at random across the study region, providing good spatial coverage and supporting the representativeness of the resulting abundance estimates. Incorporating environmental covariates into the abundance model allows density to vary spatially in response to ecological gradients, rather than assuming a constant density across the study area. This typically results in more ecologically realistic and spatially explicit estimates of abundance (Hedley and Buckland, 2004; Miller et al., 2013).

The Bayesian distance-sampling model implemented in Stan used here is computationally intensive because it estimates a large number of parameters and generates full posterior distributions for both detection and abundance components, which substantially increases model dimensionality and run times under Hamiltonian Monte Carlo sampling. Despite this complexity, the framework still enables the evaluation of alternative covariate sets and their comparison via LOOIC, as long as such comparisons remain informed by ecological reasoning. In this study, the considerable computational cost limited the feasibility of exploring a large number of candidate predictors. Thus, the analysis focused on covariates with established or anticipated ecological relevance, balancing model parsimony with interpretability and computational efficiency.

Because perpendicular distances were modeled in discrete intervals, the detection component effectively follows a multinomial-type formulation, which offers practical advantages over the traditional continuous likelihood. In particular, it reduces sensitivity to distance-measurement error, mitigates heaping, and improves numerical stability when fitting hierarchical Bayesian models for multiple species.

Populations fluctuate over time as a result of both intrinsic processes and external environmental conditions. Mechanisms such as density regulation, demographic variability, and trophic interactions generate internal dynamics, whereas

external influences arise from environmental conditions that can vary across short or long temporal scales. Among these external factors, atmospheric and oceanographic conditions often change most rapidly and can propagate through marine ecosystems in complex ways (Tredennick et al., 2021). Quantifying the relative influence of internal versus external forcing remains a central challenge in ecology. In practice, however, many biologically meaningful internal drivers are difficult to measure directly or are unavailable at appropriate spatial and temporal scales. Accordingly, and consistent with our previous work, we relied primarily on abiotic covariates, which are more readily quantified and offer a pragmatic basis for explaining broad-scale variation despite limited a priori biological information for some species groups.

References

Andriolo, A., Kinas, P. G., Engel, M. H., Martins, C. C. A., Rufino, A. M. (2010). Humpback whales within the Brazilian breeding ground: distribution and population size estimate. Endangered Species Research, 11 (3), 233-243.

Ballance, L. T. (2018). Cetacean ecology. In: Würsig, B., Thewissen, J. G. M., Kovacs, K. M. (eds) Encyclopedia of marine mammals. Academic Press, London.

Bortolotto, G. A., Thomas, L., Hammond, P., Zerbini, A. N. (2021). Alternative method for assessment of southwestern Atlantic humpback whale population status. PLOS ONE. 16 (11), e0259541.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., Thomas, L. (2001). Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford.

Buckland, S., Anderson, D., Burnham, K., Laake, J., Borchers, D., Thomas, L. (2004). Advanced distance sampling: Estimating abundance of biological populations. Oxford University Press, Oxford.

Burnham, K. P., Anderson, D. R., Laake, J. L. (1980). Estimation of density from line transect sampling of biological populations. Wildlife Monographs, (72), 3-202.

Carvalho, R. R., Quintana, C. E. P., dos Santos, C. V. C., et al. (2025). Abundance of two Guiana dolphin (Sotalia guianensis) populations inhabiting adjacent areas in southeastern Brazil. Marine Mammal Science, 41(2), e13188.

Clapham P. J., Baker C. S. (2018). Whaling, modern. In: Würsig, B., Thewissen, J. G. M., Kovacs, K. M. (eds) Encyclopedia of marine mammals. Academic Press, London.

Cremer, M. J., & Simões-Lopes, P. C. (2008). Distribution, abundance and density estimates of franciscanas, Pontoporia blainvillei (Cetacea: Pontoporiidae), in Babitonga bay, southern Brazil. Revista Brasileira de Zoologia, 25(3), 397-402.

Dalpaz, L., Cypriano-Souza, A. L., Lodi, L., Wedekin, L., Daura-Jorge, F. (2023). Bryde's whales in South Brazil Bight: Evidence of low genetic diversity and seasonal habitat use. Marine Biology 170 (8), 94.

De Almeida, I. G., Percequillo, A. R., & Rollo, M. M. (2023). Assessing Guiana dolphin abundance and density in the Southwestern Atlantic: insights from conservation areas. Journal of the Marine Biological Association of the United Kingdom, 103, e57.

Di Tullio, J. C., Gandra, T. B. R., Zerbini, A. N., Secchi, E. R. (2016). Diversity and distribution patterns of cetaceans in the subtropical Southwestern Atlantic outer continental shelf and slope. PLOS ONE, 11 (5), e0155841.

ESRI, 2018. ArcGIS Desktop: Release 10.6. Redlands, CA: Environmental Systems Research Institute. URL https://www.esri.com.

Flanders Marine Institute (2019). Maritime Boundaries Geodatabase: Maritime Boundaries and Exclusive Economic Zones (200NM), version 11. URL https://www.marineregions.org/.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B. (2013). Bayesian Data Analysis. 3rd ed. Chapman & Hall/CRC, Boca Raton.

Hazen, E. L., Abrahms, B., Brodie, S., Carroll, G., Jacox, M. G., Savoca, M. S., Scales, K. L., Sydeman, W. J., Bograd, S. J. (2019). Marine top predators as climate and ecosystem sentinels. Frontiers in Ecology and the Environment, 17, 565–574.

Hedley, S. L., & Buckland, S. T. (2004). Spatial models for line transect sampling. Journal of Agricultural, Biological, and Environmental Statistics, 9(2), 181–199.

Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio). (2011). Plano de Ação Nacional para a Conservação dos Mamíferos Aquáticos – Pequenos Cetáceos. Brasília, ICMBio/MMA.

International Whaling Commission (IWC). Status of whale populations. Available at: https://iwc.int/about-whales/population-status. Accessed: October 24, 2025.

Kellner, K. F., Fowler, N. L., Petroelje, T. R., Kautz, T. M., Beyer Jr, D. E., Belant, J. L. (2022). ubms: an R package for fitting hierarchical occupancy and N-mixture abundance models in a Bayesian framework. Methods in Ecology and Evolution, 13(3), 577-584.

Kéry, M. (2010). Introduction to WinBUGS for Ecologists: A Bayesian approach to regression, modelling and hierarchical Bayes. Academic Press, London.

Kéry, M. & Royle, J. A. (2016). Applied hierarchical modeling in ecology: Modeling distribution, abundance and species richness using R and BUGS. Volume 1: Prelude and Static Models. Academic Press, London.

Kinas, P. G., & Bethlem, C. B. P. (1998). Empirical Bayes Abundance Estimation of a Closed Population using Mark-Recapture Data, with Application to Humpback Whales, *Megaptera novaeangliae*, in Abrolhos, Brazil. Report of the International Whaling Commission (IWC) 48, 447–450.

Kinzey, D., & Gerrodette, T. (2001). Conversion factors for binocular reticles. Marine Mammal Science, 17(2), 353-361.

LeDuc, R. (2018). Biogeography. In: Würsig B., Thewissen J.G.M., Kovacs K.M. Encyclopedia of Marine Mammals. 3rd ed. Academic Press, London.

Lerczak, J. A., & Hobbs, R. C. (1998). Calculating sighting distances from angular readings during shipboard, aerial, and shore-based marine mammal surveys. Marine Mammal Science, 14(3), 590-599.

Lima, R. C., Di Tullio, J. C., Secchi, E. R., Castro, F. R., Troina, G. C. (2021). Delphinid Mixed-Species Associations in the Oceanic Waters of the Western South Atlantic. Aquatic Mammals, 47(1), 2021.

Martins, A. C., Kinas, P. G., Wedekin, L. L., Dalla Rosa, L. (2023). Spatial and seasonal patterns of cetacean species richness: A Bayesian approach. Deep Sea Research Part I: Oceanographic Research Papers, 196, 104046.

Martins, A. C., Kinas, P. G., Wedekin, L. L., Dalla Rosa, L. (2026). Habitat-based BART models for cetaceans in the western South Atlantic: current and future distribution under climate change scenarios. Marine Environmental Research, 217, 107964.

Miller, D. L., Burt, M. L., Rexstad, E., Thomas, L. (2013). Spatial models for distance sampling data: recent developments and future directions. Methods in Ecology and Evolution, 4(11), 1001–1010.

Pavanato, H. J., Wedekin, L. L., Guilherme-Silveira, F. R., Engel, M. H., Kinas, P. G. (2017). Estimating humpback whale abundance using hierarchical distance sampling. Ecological Modelling, 358, 10-18.

Rankin, S., Oedekoven, C., Archer, F. (2020). Mark recapture distance sampling: using acoustics to estimate the fraction of dolphins missed by observers during shipboard line-transect surveys. Environmental and Ecological Statistics, 27(2), 233-251.

Roberts, J. J., Best, B D., Mannocci, L., Fujioka, E. I., Halpin, P. N., Palka, D. L., Garrison, L. P., Mullin, K. D., Cole, T. V., Khan, C. B., McLellan, W. A., Pabst, D. A., Lockhart, G. G. (2016). Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico. Scientific Reports, 6 (1), 22615.

Roman, J., Estes, J. A., Morissette, L., Smith, C., Costa, D., McCarthy, J., Nation, J., Nicol, S., Pershing, A., Smetacek, V. (2014). Whales as marine ecosystem engineers. Frontiers in Ecology and the Environment, 12 (7), 377–385.

Royle, J. A. & Dorazio, R. M. (2008). Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities. Academic Press, San Diego.

Seyboth, E., Meynecke, J. O., De Bie, J., Roychoudhury, A., Findlay, K. (2023). A review of post-whaling abundance, trends, changes in distribution and migration patterns, and supplementary feeding of Southern Hemisphere humpback whales. Frontiers in Marine Science, 10, 997491.

Sollmann, R., Gardner, B., Williams, K. A., Gilbert, A. T., Veit, R. R. (2016). A hierarchical distance sampling model to estimate abundance and covariate associations of species and communities. Methods in Ecology and Evolution, 7(5), 529-537.

Tredennick, A.T., Hooker, G., Ellner, S.P., Adler, P.B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. Ecology, 102 (6), e03336.

van Compernolle, M., Morris, J., Calich, H. J., et al. (2025). Vulnerability of marine megafauna to global at-sea anthropogenic threats. Conservation Biology, 40, e70147.

van de Pol, M., Bailey, L. D., McLean, N., Rijsdijk, L., Lawson, C. R., Brouwer, L. (2016). Identifying the best climatic predictors in ecology and evolution. Methods in Ecology and Evolution, 7(10), 1246-1257.

Vehtari, A., Gelman, A., Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing, 27(5), 1413-1432.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., Bürkner, P. C. (2021). Rank-normalization, folding, and localization: An improved R ˆ for assessing convergence of MCMC (with discussion). Bayesian Analysis, 16(2), 667-718.

von Hammerstein, H., Setter, R.O., van Aswegen, M., Currie, J.J., Stack, S.H. (2022). High-resolution projections of global sea surface temperatures reveal critical warming in humpback whale breeding grounds. Frontiers in Marine Science, 9, 837772.

Wedekin, L.L., Neves, M.C., Marcondes, M.C.C., Baracho, C., Rossi-Santos, M.R., Engel, M.H., Simões-Lopes, P.C. (2010). Site fidelity and movements of humpback whales (*Megaptera novaeangliae*) on the Brazilian breeding ground, southwestern Atlantic. Marine Mammal Science, 26 (4), 787-802.

Wedekin, L. L., Rossi-Santos, M. R., Baracho, C., Cypriano-Souza, A. L., Simões-Lopes, P. C. (2014). Cetacean records along a coastal-offshore gradient in the Vitória-Trindade Chain, western South Atlantic Ocean. Brazilian Journal of Biology, 74(1), 137-144.

Würsig, B. (2019). Ethology and Behavioral Ecology of Odontocetes. Springer, Cham.

Zerbini, A. N., Adams, G., Best, J., Clapham, P. J., Jackson, J. A., Punt, A. E. (2019). Assessing the recovery of an Antarctic predator from historical exploitation. Royal Society Open Science, 6, 190368.

Supplementary material

Supplementary table 1. Survey periods and summaries for cetacean sightings and search effort by cruise.

cruise	year	period	N sightings	N species	N species (cumulative)	effort (h)	effort (km)
1	2015 - 2016	20 Nov - 5 Jan	31	10	10	389.9	4,937.1
2	2016	10 May - 20 Jun	57	12	16	410.9	5,462.5
3	2016	27 Sep - 6 Nov	74	14	18	420.1	5,500.2
4	2017	8 Mar - 18 Apr	38	8	18	434.9	5,541.9
5	2017	25 Aug - 7 Oct	94	11	18	435.7	5,583.8
6	2018	12 Jan - 20 Feb	44	11	21	444.3	5,863.4
7	2018	1 Sep - 20 Oct	110	14	21	477.2	5,930.5
8	2019	13 Mar - 26 Apr	58	17	24	425.3	5,502.3
Total	-	-	506	-	24	3,438.3	44,321.7

Supplementary table 3. Posterior summaries of the parameters describing the abundance (state) process.

Survey	Parameter	Estimate	SD	2.5%	97.5%	n_eff	Rhat
5	(Intercept)	-4.0281	0.186	-4.419	-3.6715	598	1
5	species3	-0.7778	0.357	-1.504	-0.075	685	1
5	species4	-2.2545	0.847	-4.134	-0.8039	760	1
5	sst	-0.1947	0.113	-0.402	0.0374	1062	1

5	species3:sst	0.0742	0.219	-0.482	0.3814	1319	1
5	species4:sst	1.0099	0.704	-0.233	2.4814	1252	1
6	(Intercept)	-6.7089	0.813	-8.455	-5.279	503	1
6	species3	2.4854	0.837	0.992	4.264	485	1
6	species4	-0.8522	1.002	-2.748	1.173	711	1
6	sst	-0.2272	0.565	-1.339	0.888	876	1
6	species3:sst	0.0691	0.591	-1.23	1.086	910	1
6	species4:sst	-0.1453	0.774	-1.665	1.366	942	1
7	(Intercept)	-3.8076	0.162	-4.127	-3.4983	6666	1
7	species3	-0.7193	0.373	-1.48	-0.0103	7443	1
7	species4	-1.3291	0.447	-2.25	-0.4987	7140	1
7	sst	0.0325	0.126	-0.206	0.2836	8402	1
7	dist_iso1000	-0.6422	0.103	-0.846	-0.4431	7026	1
7	species3:sst	0.3882	0.246	-0.85	0.1219	8142	1
7	species4:sst	0.0498	0.382	-0.645	0.8583	7869	1
7	species3iso1000	0.4711	0.282	-0.041	1.0557	7477	1
7	species4:iso1000	0.4931	0.355	-0.196	1.2046	8960	1
8	(Intercept)	-5.6923	0.583	-6.9391	-4.6458	2922	1
8	species3	1.1628	0.622	0.0204	2.4712	3051	1
8	species4	-1.1327	0.763	-2.6159	0.4117	3684	1
8	sst	-0.5488	0.317	-1.1483	0.0708	3885	1
8	elevation	0.5393	0.527	-0.3912	1.6959	2972	1
8	species3:sst	0.6525	0.357	-0.0434	1.3516	4071	1
8	species4:sst	1.3117	0.576	0.2408	2.5234	4549	1
8	species3:elevation	-0.0519	0.56	-1.2498	0.965	3085	1
8	species4:elevation	-0.9704	0.651	-2.3164	0.2209	3368	1

Supplementary table 4. Posterior summaries of the parameters describing the detection (observation) process.

Survey	Parameter	Estimate	SD	2.5%	97.5%	n_eff	Rhat
5	(Intercept)	0.0571	0.223	-0.246	0.556	266	1.01
5	species3	-0.521	0.304	-1.142	0.036	369	1.01
5	species4	-1.1317	0.821	-1.993	1.129	587	1.01
6	(Intercept)	-0.0603	0.934	-0.909	2.83	617	1
6	species3	0.3279	0.944	-2.587	1.26	614	1
6	species4	1.9662	1.575	-0.554	5.83	998	1
7	(Intercept)	0.796	0.116	0.595	1.049	6474	1
7	species3	-0.96	0.242	-1.43	-0.476	7539	1
7	species4	-0.941	0.308	-1.497	-0.292	8218	1
8	(Intercept)	-0.424	0.36	-1.008	0.32	3797	1
8	species3	1.007	0.531	0.147	1.88	1787	1
8	species4	2.437	1.359	0.617	5.75	3912	1
8	beaufort3	-0.3564	0.251	-0.8451	0.1234	1006	1.0
8	beaufort4	-0.3691	0.311	-0.9214	0.2898	1117	1.0
8	beaufort5	-0.7527	0.329	-1.3644	-0.0559	819	1.0
8	beaufort6	-0.2337	0.542	-1.0954	1.0075	1045	1.01

CAPÍTULO V

Conclusões gerais e perspectivas futuras

Conclusões gerais e perspectivas futuras

Este estudo fornece a primeira avaliação abrangente da diversidade, distribuição e abundância de cetáceos na Bacia de Santos, combinando modelagem ecológica, amostragem sistemática e abordagens estatísticas avançadas. De forma geral, os resultados destacam a importância ecológica singular da região, que funciona como corredor migratório, área de alimentação e potencial local de reprodução para uma comunidade diversa de cetáceos.

Os resultados obtidos ao longo da tese demonstram que diferentes grupos da comunidade de cetáceos da Bacia de Santos respondem de forma consistente a gradientes fisiográficos, oceanográficos e sazonais em escala regional. Embora a elevada diversidade de cetáceos já fosse conhecida para a região, este estudo permitiu identificar padrões espaciais e temporais robustos relacionados à riqueza, distribuição e abundância de diferentes grupos ecológicos, evidenciando a influência conjunta de gradientes batimétricos, características geomorfológicas e condições oceanográficas sobre a estruturação da comunidade.

A riqueza de espécies foi maior nas proximidades da isóbata de 1000 m, em áreas com topografia acentuada (alta rugosidade e declividade) e na porção norte da bacia, padrão consistentemente observado ao longo de todo o período amostrado. A temperatura surgiu como um preditor sazonal importante, com um aumento geral na riqueza, particularmente no sul, durante o período frio. Esse aumento sazonal da riqueza pode ser interpretado como resultado da coexistência temporária de espécies residentes e migratórias na região. Em particular, pode refletir a modulação das condições oceanográficas associada à maior influência de massas de água subantárticas e ao aporte de nutrientes provenientes de sistemas continentais adjacentes. Em contraste, durante o período quente, a dominância de massas de água mais oligotróficas tende a favorecer a persistência de espécies tipicamente tropicais, resultando em uma comunidade menos diversa em termos absolutos. A inclusão de covariáveis ambientais e espaço-temporais permitiu capturar padrões ecológicos robustos. A análise dos resíduos não indicou a

presença de autocorrelação espacial remanescente, sugerindo que a estrutura espacial foi adequadamente representada pelos preditores incluídos no modelo.

Os padrões atuais de uso de habitat indicam que os diferentes grupos taxonômicos ocupam o espaço de forma estruturada ao longo de gradientes ambientais. Baleias migratórias estão associadas a regiões do talude e a faixas térmicas mais restritas, e as baleias-de-Bryde estão associadas principalmente a ambientes costeiros e de plataforma externa. Pequenos delfinídeos são amplamente distribuídos na plataforma continental e suas porções externas, enquanto grandes delfinídeos, cachalotes e baleias-bicudas estão associados a ambientes de talude e regiões oceânicas mais profundas, com diferenciação espacial ao longo da bacia, sendo os grandes delfinídeos mais associados à porção norte e os cachalotes e baleias-bicudas à porção sul. Esse arranjo espacial constitui a base observacional a partir da qual se projetam respostas diferenciais às mudanças ambientais futuras.

As projeções de mudança climática sugerem respostas diferenciadas entre os grupos taxonômicos: enquanto baleias migratórias podem sofrer contração e deslocamento de habitat em direção à costa, pequenos delfinídeos e baleias-de-Bryde podem expandir sua distribuição, e grandes delfinídeos, cachalotes e baleias-bicudas devem manter relativa estabilidade. Isso ressalta a necessidade de estratégias de conservação adaptativas, capazes de lidar com dinâmicas futuras distintas. De maneira geral, as projeções não indicam uma perda generalizada de adequabilidade ambiental para os grupos de cetáceos na Bacia de Santos, mas sim possíveis reorganizações na distribuição e alterações relativas na probabilidade de ocorrência entre grupos. Embora estudos conduzidos em escalas mais amplas frequentemente projetem impactos negativos e deslocamentos distributivos para diversas espécies de cetáceos, os resultados obtidos neste estudo sugerem que a Bacia de Santos poderá continuar oferecendo condições ambientais propícias para a ocorrência dos grupos atualmente registrados, ainda que alguns deles apresentem alterações na extensão ou na configuração espacial de seus habitats potencialmente mais favoráveis..

As estimativas de abundância estabelecem uma linha de base demográfica essencial para a região. A abundância de pequenos delfinídeos foi substancialmente

maior do que a de grandes delfinídeos, bem como apresentou menor sazonalidade, um padrão consistente com a ecologia desses grupos e com relatos em outras regiões. Foram estimados mais de 3.000 indivíduos de misticetos utilizando a região sazonalmente. A bacia serve tanto como rota migratória quanto como habitat temporário para grandes baleias migratórias, especialmente para a baleia-jubarte, que se concentra na porção norte e ao longo do talude continental, com pico de abundância durante o período frio e início da primavera.

Esses resultados são consistentes com os padrões de riqueza observados no Capítulo 1, reforçando a importância da entrada sazonal de misticetos na estruturação da comunidade durante o período frio. Adicionalmente, as estimativas de abundância sugerem um aumento no número de grandes delfinídeos no outono, com maior variabilidade entre campanhas, embora com elevada incerteza associada ao menor número de registros.

O estudo também ampliou o conhecimento sobre a ocorrência de espécies na Bacia de Santos, elevando para 24 o número de espécies registradas em levantamentos de grande escala, em comparação com levantamentos sistemáticos anteriores realizados na região. Incluem-se novos registros de *Balaenoptera borealis*, *Delphinus delphis*, *Lagenodelphis hosei*, *Peponocephala electra* e *Kogia sima*. As áreas mais ricas em espécies coincidem com zonas de importância ecológica crítica, como a quebra de plataforma e a porção norte da bacia, que também são regiões de elevada pressão antrópica,caracterizando áreas de sobreposição entre alta diversidade e uso intensivo do espaço marinho por atividades humanas . A região costeira, embora com menor riqueza, abriga espécies de especial interesse para a conservação, como *Sotalia guianensis* e *Pontoporia blainvillei*, esta última criticamente ameaçada.

Do ponto de vista metodológico, este estudo demonstra a aplicabilidade de abordagens estatísticas avançadas, como modelos bayesianos GLM, BART e distance sampling hierárquico, para quantificar padrões de distribuição e abundância de cetáceos em larga escala. A escolha cuidadosa de covariáveis ambientais e a consideração de efeitos espaciais e sazonais permitiram a construção de modelos robustos, que capturam as principais relações espécie-habitat.

É recomendável considerar mais dados de campanhas de amostragem em diferentes estações e anos para validar padrões sazonais e interanuais, refinando as estimativas de abundância e distribuição numa escala temporal mais fina. Pode-se considerar a possibilidade de integrar dados de campanhas de avistagem embarcada com dados de campanhas de telemetria e campanhas de monitoramento aéreo. Além disso, o acompanhamento contínuo de longo prazo permitirá detectar mudanças na presença de espécies, deslocamentos de habitat e efeitos de variabilidade climática ou antrópica.

Uma abordagem necessária é a sobreposição espacial dos mapas de presença de espécies com camadas de impacto humano (tráfego de embarcações, exploração de óleo e gás, pesca, poluição acústica e química). Esta análise de risco espacial é fundamental para identificar áreas de conflito emergente, orientar medidas de mitigação (como rotas de navegação ou restrições sazonais) e otimizar a criação e gestão de Áreas Marinhas Protegidas (AMPs), com definição potencial de áreas prioritárias ao longo da quebra de plataforma e do setor norte da bacia, especialmente em períodos de maior concentração de cetáceos.

A inclusão de variáveis bióticas, como a disponibilidade de presas e a interação entre espécies, pode complementar e melhorar o poder preditivo dos modelos, uma vez que as mudanças climáticas podem intensificar a competição interespecífica e promover alterações em nível de ecossistema, gerando novos desafios à conservação. Tais informações permitiriam capturar relações tróficas e ecológicas, reduzindo a dependência exclusiva de proxies abióticos, já que alterações ambientais podem modificar a distribuição, o recrutamento e a disponibilidade local das espécies forrageiras, potencialmente influenciando o uso de habitat por cetáceos por meio de processos *bottom-up*. De forma complementar, a aplicação de modelos mecanicistas, que incorporam processos ecológicos fundamentais como migração, reprodução, competição e dinâmica populacional, pode fornecer previsões mais realistas sobre a distribuição e abundância futura de cetáceos sob diferentes cenários ambientais e climáticos, desde que as relações causais sejam claramente entendidas e corretamente modeladas.

Adicionalmente, para uma visão mais holística da conservação marinha, futuros estudos podem integrar os modelos de cetáceos com modelos de distribuição de outros componentes-chave do ecossistema, como aves marinhas, tartarugas, peixes e seus recursos tróficos. Esta abordagem ecossistêmica permitiria identificar *hotspots* de biodiversidade e produtividade multifuncionais.

As distribuições posteriores geradas neste estudo podem constituir prioris informativos valiosos para futuras análises Bayesianas na região, promovendo uma acumulação eficiente de conhecimento. Assim, permite incorporar novos dados de forma progressivamente mais precisa, fortalecendo a base científica para subsidiar decisões de gestão.

APÊNDICE

Códigos utilizados no R

```
##########################################################################
#                                                                        #
#                    Modelagem de Riqueza de espécies                    #
#                                                                        #
##########################################################################
### por Ana Carolina Martins e Paul Kinas
### maio de 2021
### ana.carolbm@hotmail.com

####  Parte I: stepwise regression ####
# carregar o arquivo de dados
load (file="data.Rda") # planilha de dados contendo informações extraídas de 244 células de grid (122 período quente, 122 período frio)

# scale
data$z.elev <- scale(data$elev)
….
data$z.long <- scale(data$long)

# omitir linhas com algum NA
dat0 <- na.omit(dat)

# Exploratória
names(dat0)[47:54]
plot(dat[,47:54]) # pairplot relações entre variáveis
```

```
# analisar relações entre variáveis que pareçam relacionadas
plot(dat0$z.chl, dat0$z.sss)
…

# boxplot variáveis
boxplot(elev)
…
boxplot(dist_front)

# analisar cada variável em função da variável resposta
attach(data)
plot(elev, n_sp)
…
plot(dist_front, n_sp)

# parece haver uma relação quadrática para 'z.sst' em função da var. resposta
 # construímos os polinômios ortogonais de segunda ordem:
 t.poly.sst <- poly(as.numeric(dat0$z.sst), 2)
 dat0$z.sst.l <- t.poly.sst[,1]
 dat0$z.sst.q <- t.poly.sst[,2]

# stepwise
fullmod <- glm(n_sp~z.elev+z.lslope+z.r3+z.ncoast+z.n200+
```

```
                z.lat+z.long+ z.sst.l+z.sst.q+z.chl.l+z.chl.q+
            z.ssh+z.sla+z.vel_u+z.vel_v+z.dist_front+
                log(esf), family=poisson)
summary(fullmod)

nothing <- glm(n_sp ~ 1,family=poisson) # null
summary(nothing)

backwards <- step(fullmod) # Backwards selection is the default

forwards <- step(nothing,
                        scope=list(lower=formula(nothing),upper=formula(fullmod)),
direction="forward")
…
# D  rodar step para sst (l,q) + log(chl) + iso1000 (sqrt) + log (slope)
 fullmodD <- glm(n_sp~z.elev+z.lslope+z.r3+z.s1000+
                z.lat+z.long+
                z.sst.l+z.sst.q+z.chl+
                z.ssh+z.sla+z.vel_u+z.vel_v+z.sdist_front+
                log(esf),family=poisson)
summary(fullmodD)

nothing <- glm(n_sp ~ 1,family=poisson)
summary(nothing)
```

```
backwardsD <- step(fullmodD) # Backwards selection is the default


forwardsD <- step(nothing,
                          scope=list(lower=formula(nothing),upper=formula(fullmodD)),
direction="forward")


####  Parte II: Bayesian inference  ####


library(arm)
# inference
n.sim <- 2000
poisim <- sim(mod, n.sim=n.sim) # mod: best model selecionado em Parte I


# criar matriz com coeficientes posteriores simulados
Mcoef <- poisim@coef


# Graficos
nomes <- names(coef(mod))


# separar os termos polinomias 'sst.l' e 'sst.q'
bayesci(Mcoef[,c(2,3,4,7)],labels=nomes[c(2,3,4,7)],
      main="Posterior 95% Credibility Intervals", xlab="coefficients", ylab="estimates")
abline(h=0,lty=21)
# os outros coeficientes:
par(mfrow=c(3,1))
```

```
bayeshist(Mcoef[,1],main="Marginal Posterior",xlab="Intercept", nclass=25)
abline(v=0, lty=21)
bayeshist(Mcoef[,5],main="Posterior SST linear",xlab="z.sst.l", nclass=25)
abline(v=0,lty=21)
bayeshist(Mcoef[,6],main="Posterior SST quadratic",xlab="z.sst.q", nclass=25)
abline(v=0,lty=21)
par(mfrow=c(1,1))

# fazer um gráfico para indicar o efeito de 'sst' sobre o número esperado de espécies
newdat <- data.frame(
  esf = mean(dat0$esf),
  z.r3 = 0,
  z.s1000 = 0,
  z.sst = seq(min(dat0$z.sst), max(dat0$z.sst), length.out=100),
  z.lat = 0,
  z.lslopea = 0)

newdat$z.sst.l<- predict(t.poly.sst, newdat$z.sst)[,1]
newdat$z.sst.q <- predict(t.poly.sst, newdat$z.sst)[,2]

Xmat <- model.matrix(~ log(esf) + z.r3 + z.s1000 + z.sst.l + z.sst.q + z.lat + z.lslopea,
data=newdat)

fitmat <- exp(Xmat %*% t(Mcoef))   # usando a escala natural
```

```
#fitmat <- matrix(nrow=nrow(newdat), ncol=n.sim)
#for(i in 1:n.sim) fitmat[,i] <- exp(Xmat %*% bsim@fixef[i,])
newdat$mean <- apply(fitmat,1,mean)
newdat$lwr <- apply(fitmat, 1, quantile, prob=0.025)
newdat$upr <- apply(fitmat, 1, quantile, prob=0.975)

# para usar a escala de SST em graus Celsius, recuperar a média e dp usadas na padronização
m.sst <- attr(dat$z.sst,"scaled:center")
sd.sst <- attr(dat$z.sst,"scaled:scale")
newdat$sst <- newdat$z.sst*sd.sst + m.sst
# e finalmente o gráfico:
plot(n_sp ~ mean_sst, data=dat0,type="n", main="Nr of Species by SST",
    xlab="Degrees Celsius", ylab="Species Count",xlim=c(20,28))
polygon(c(newdat$sst, rev(newdat$sst)),
       c(newdat$lwr, rev(newdat$upr)),
       col=grey(0.7), border=NA)
lines(newdat$sst, newdat$mean, lwd=2) # population mean
# (a) sobrepor os pontos observados (versao simples)
#points(jitter(n_sp) ~ jitter(mean_sst), data=dat0)   # versão com jitter
# OU:
# (b) sobrepor os pontos observados
# linhas de 1 a114 são 'per = q' e linhas de 115 a 224 sao período 'f'
points(jitter(n_sp) ~ jitter(mean_sst),
      col=c(rep("red",114),rep("blue",112)),
```

```
      pch=as.character(c(rep("q",114),rep("f",112))),data=dat0)   # versão com jitter

## Espacializacao dos dados
plot(dat$long[1:122],dat$lat[1:122],pch=19)

medias_sst <- tapply(dat$z.sst,dat$per,mean,na.rm=T)
m_sst_q <- medias_sst[1]
m_sst_f <- medias_sst[2]

# predicao espacial para periodo 'quente'
new_q <- data.frame(
  esf = mean(dat0$esf),
  z.r3 = dat$z.r3[1:122],
  z.s1000a = z.s1000a[1:122],
  z.sst = rep(m_sst_q,122),
  z.lat = dat$z.lat[1:122],
  z.lslopea = z.lslopea[1:122])
#
new_q$z.sst.l<- predict(t.poly.sst, new_q$z.sst)[,1]
new_q$z.sst.q <- predict(t.poly.sst, new_q$z.sst)[,2]

Xq <- model.matrix(~ log(esf) + z.r3 + z.s1000a + z.sst.l + z.sst.q + z.lat + z.lslopea,
data=new_q)

# cada linha de 'fitq' contem as 2000 simulacoes posteriores de 'mu'
```

```
fitq <- exp(Xq %*% t(Mcoef))   # usando a escala natural

mu_q <- apply(fitq,1,mean)
summary(mu_q)

# predicao espacial para periodo 'frio'
new_f <- data.frame(
  esf = mean(dat0$esf),
  z.r3 = dat$z.r3[123:244],
  z.s1000a = z.s1000a[123:244],
  z.sst = rep(m_sst_f,122),
  z.lat = dat$z.lat[123:244],
  z.lslopea = z.lslopea[123:244]
  )
#
new_f$z.sst.l<- predict(t.poly.sst, new_f$z.sst)[,1]
new_f$z.sst.q <- predict(t.poly.sst, new_f$z.sst)[,2]

Xf <- model.matrix(~ log(esf) + z.r3 + z.s1000a + z.sst.l + z.sst.q + z.lat + z.lslopea,
data=new_f)

# cada linha de 'fitq' contem as 2000 simulacoes posteriores de 'mu'
fitf <- exp(Xf %*% t(Mcoef))   # usando a escala natural

mu_f <- apply(fitf,1,mean)
```

```
summary(mu_f)

# obs: processamento de variáveis ambientais e criação do grid realizados no ArcGIS
## as simulações posteriores em 'mu_x' foram transferidas para o grid no ArcGIS, para visualização e criação de mapas

#--------- suplementar
###  bubble plot residuals
dat0$Resid <- rstandard(mod)
library(sp)
coordinates(dat0) <- ~lat + long  #effectively convert the data into a spatial df
b <- bubble(dat0, "Resid", main="")

###  variogram
library(geoR)
# Coordinates
coords <- cbind(dat0$long, dat0$lat)

# Deviance residuals from the selected Poisson GLM
res <- residuals(mod, type = "deviance")

# Optional: subset if needed
idx <- 123:222
coords <- coords[idx, ]
res <- res[idx]
```

```
# Create geodata object
geo_res <- as.geodata(cbind(coords, res))

# Empirical variogram
vg <- variog(geo_res, bin.cloud = TRUE)

# Monte Carlo envelopes
vg_env <- variog.mc.env(geo_res, obj.var = vg)

# Plot
plot(vg, envelope = vg_env)

### proporção de zeros
nsim <- 2000
bsim <- sim(poi8, n.sim=nsim)

fitmat <- matrix(ncol=nsim, nrow=nrow(new_q))
yrep <- matrix(ncol=nsim, nrow=nrow(new_q))
Xmat <-  model.matrix(~ log(esf) + z.r3 + z.s1000a + z.sst.l + z.sst.q + z.lat + z.lslopea,
data=new_q)
for(i in 1:nsim) {
  fitmat[,i] <- exp(Xmat %*% bsim@coef[i,])
  yrep[,i] <- rpois(nrow(new_q), lambda=fitmat[,i])
}
```

```
propzeros <- function(x) sum(x==0)/length(x)
propzeros(data_1$n_sp)

pzerossim <- apply(yrep, 2, propzeros) # prop. zero values in yrep
quantile(pzerossim, prob=c(0.01, 0.99))
sum(pzerossim>=propzeros(data_1$n_sp))/nsim

### ecdf
load("dat0.Rda") # dados sem NA
n_sp <- dat0$n_sp

#ecdf plots
plot(ecdf(yrep),xlim=c(0,quantile(yrep,0.98)))
plot(ecdf(yrep))

# seleciona apenas 100 réplicas
nam <- 100
amostra <- yrep[ ,sample(1:ncol(yrep),nam)] # mudei o original 1:nrow para 1:ncol
for(i in 1:nam) {
  plot(ecdf(amostra[,i]), col="grey",add=T)
}
plot(ecdf(n_sp),add=T, col="orange")

####---------------------- F I M -------------------------####
```

```
##########################################################################
#                                                                        #
#                    Bayesian Additive Regression Trees                  #
#                                                                        #
#                                                                        #
##########################################################################
### por Ana Carolina Martins
### abril de 2024

library(dbarts)
library(embarcadero)
library(reshape)

#importar dados
load(file="data.Rda") # planilha de dados contendo informações extraídas de 4080 segmentos amostrados
names(data)

xvars<- xvars[c(9:15)] # colunas com informação das variáveis ambientais

#### etapa 1 avaliacao e construção do classificador para grupo '1' baleen ####
# bart.step: funcao automatizada para escolha do conjunto de variaveis
my <- bart.step(x.data = data[,xvars], # variaveis ambientais em cada segmento (linha)
                y.data = data,"g1_oc"], # coluna c dados de ocorrencia (0,1) do grupo 1 em cada segment
```

```
        full = T,
        quiet = F)

# partial plots
partial(my,
    c('lat','long', 'dist_coast','dist_iso200', 'sst'),
    trace = FALSE,
    ci = TRUE,
    panel = TRUE,
    smooth = 3)
# o mesmo para todos os grupos

#### etapa 2 - predicao  ####
# para fazer camada de predicao: usar rasters #
library(raster)
elevation <- raster('elevation.tif')
…
sla <- raster('sla.tif')

# projecoes
proj4string(elevation) <- CRS("+init=EPSG:4326")
# ...
dim(elevation)
```

```
# if rasters have different extent
lat <- resample(lat,elevation)
…

# criar um objeto stack com todas as camadas raster ambientais
# obs: todas devem ter a mesma dimensao
stk <- stack(lat,long,dist_coast,dist_iso200,sst) # climatologia sst

# predict baleen current scenario
baleen_cur <- predict(object = my, # modelo
                      x.layers = stk, # stack
                  quantiles = c(0.025, 0.975),
                      splitby = 20,
                      quiet = TRUE)

writeRaster(baleen_cur,  file="baleen_cur.tif")

# o mesmo para todos os grupos
# para projecoes futuras, criar stack com novos rasters
# obs: predições foram visualizadas e mapas finais criados no ArcGIS 10.6.1

####---------------------- F I M -------------------------####
```

```
##########################################################################
#                                                                        #
#                        Hierarchical Distance Sampling                  #
#                                                                        #
##########################################################################
### por Paul Kinas e Ana Carolina Martins
### iniciado setembro de 2022, retomado em setembro de 2025

#importar o arquivo de cada campanha
load(file="datac5i.Rda")
c5 <- datac5i

## variaveis preditoras
covar <- c5[,c(3,5,6,19:29)] # seleciona colunas onde estao variaveis
head(covar)
str(covar)
summary(covar)

## para facilitar futuro acesso e antes de padronizar as variaveis,
## guardar media e dp das variaveis preditoras (colunas)
media_pred <- apply(preditor,2,mean)
dp_pred <- apply(preditor,2,sd)

## segue a matriz das covariaveis padronizadas
```

```
pred_z <- data.frame(length_segm=preditor[,1],
                     apply(preditor[,-1],2,scale))
summary(pred_z)

## sample distance data
load(file="dd5new.Rda") # carrega distancias perpendiculares para cada avistagem

summary(dd5new)
# grafico em dist_perp_by_group_c5.R

## maxima distancia perpendicular nos dados
## (distancias colocadas em km)
m_width <- max(dd5new_trunc[,,1:5],na.rm=TRUE)*1.01/1000
## comprimentos dos segmentos medidos nos dados (km) no vetor 'line'
line <- c5$length_segm
# numero de segmentos
nsites <- nrow(c5)
# numero de especies (grupos)
nsp <- dim(dd5new)[3]

#
# colocar distancias perpendiculares em classes de distance:
# 1. converter distancias medidas de m para km
dd5new_trunc <- dd5new_trunc[,,1:5]/1000
```

```
max(dd5new_trunc[,,1:5],na.rm=TRUE)


hist(dd5new, main="Histogram of perpendicular distances", xlab="Distance (km)")


# 2. define 4 binning distances
nbin <- 4
bin <- seq(0,m_width,length.out=(nbin+1))
bin
## my binning function
bin_func <- function(x){
  y <- as.numeric(table(cut(x,breaks=bin)))
  y
}
## binnen distance frequencies matriz for each species
db <- array(rep(NA,nsites*nsp*nbin),dim=c(nsp,nsites,nbin))
for(i in 1:nsp){
  db[i,,] <- t(apply(dd5new_trunc[,,i],1,bin_func))
}


#########
## Stacking:
##    only for selected species (here, just groups 1 3 4)
db_stack <- data.frame(rbind(db[1,,], db[3,,], db[4,,]))
db_stack$id <- rep(1:nsites, 3)
```

```
db_stack$site <- as.factor(rep(1:nsites, 3))
db_stack$species <- as.factor(rep(c(1,3,4),rep(nsites,3)))

# stacking standardized predictive variables
covar <- pred_z[,c(4,10,15)]
dx_stack <- data.frame(rbind(covar,covar,covar))
dbx_stack <- cbind(db_stack,dx_stack)

str(dbx_stack)
dim(db_stack)

line_stack <- rep(line,3)
################################

#################################################
##                Preparing Data within UBMS                ##
#################################################
library(ubms)

## Alternative data for selected species
## Data preparation for only groups 1 and 3 (g1 and g3)
data_13 <- unmarkedFrameDS(y=as.matrix(dbx_stack[,1:4]),
                           dist.breaks=bin,
                           siteCovs = dbx_stack[,5:10],
```

```
                    unitsIn="kmsq",
                    tlength = line_stack,
                    survey="line")
head(data_13)
str(data_13)

# exemplo de modelo testado
mod_c5_com_beau <- stan_distsamp(
  ~ 1 + species + beaufort  # fórmula da detecção
  ~ 1 + offset(log(line_stack)) + species +  sst + sst:species,  # fórmula da abundância
  data_13,
  keyfun = "halfnorm", #### ATENCAO
  output = "abund",         # ou "density"
  unitsOut = "kmsq",        # ou "ha"
  chains = 3,
  cores = 3,
  iter = 4000,
  warmup = 1000,
  # Prioris
  prior_intercept_state = normal(0, 1.5),
  prior_coef_state = normal(0, 1),
  prior_intercept_det = normal(0, 1),
  prior_coef_det = normal(0, 0.7)
)
```

```
## 3. Calcular média de sigma para cada grupo
# sigma_vec_correct calculado a partir da curva de deteccao
sigma1_mean <- sigma_vec_correct["g1"]
sigma3_mean <- sigma_vec_correct["g3"]
sigma4_mean <- sigma_vec_correct["g4"]

## 4. Calcular área efetiva percorrida por segmento para cada grupo
area_stack <- c(2 * sigma1_mean * line,  # grupo 1
                2 * sigma3_mean * line,  # grupo 3
                2 * sigma4_mean * line)  # grupo 4

# corrigir para área efetiva percorrida
data_13_corr <- unmarkedFrameDS(
  y = as.matrix(dbx_stack[,1:4]),
  dist.breaks = bin,
  siteCovs = dbx_stack[,5:10],
  unitsIn = "kmsq",
  tlength = area_stack,
  survey = "line"
)

## 5. Ajustar modelo final com offset correto em log(area_stack)
mod_c5_corr <- stan_distsamp(
  ~ 1 + species + beaufort  # fórmula da detecção
```

```
  ~ 1 + offset(log(line_stack)) + species +  sst + sst:species,  # fórmula da abundância
  data_13_corr,
  keyfun = "halfnorm",
  output = "abund",
  unitsOut = "kmsq",
  chains = 3,
  cores = 3,
  iter = 4000,
  warmup = 1000,
  # Prioris
  prior_intercept_state = normal(0, 1.5),
  prior_coef_state = normal(0, 1),
  prior_intercept_det = normal(0, 1),
  prior_coef_det = normal(0, 0.7)
)

####---------------------- F I M ---------------------------####
```