

Modelagem espaço-temporal da captura em pescarias comerciais através de modelos hierárquicos bayesianos com estrutura gaussiana latente

Fernando de Pol Mayer - LEG, UFPR¹

Wagner Hugo Bonat - LEG/DEST, UFPR²

Paulo Justiniano Ribeiro Jr - LEG, UFPR³

Resumo: *A captura de espécies exploradas em pescarias comerciais normalmente é modelada sem qualquer consideração sobre a dependência do espaço e do tempo. Os modelos espaço-temporais representam uma maneira mais acurada para a modelagem de dados de populações biológicas, onde existe dependência no espaço, no tempo, e possivelmente na interação espaço-tempo. Neste trabalho, foram utilizados modelos hierárquicos bayesianos com estrutura gaussiana latente para modelar dados de uma espécie capturada incidentalmente. Os parâmetros foram estimados através de inferência bayesiana, utilizando a integração aproximada aninhada de Laplace (INLA). O modelo demonstrou que não há uma tendência temporal clara, porém, que existe uma forte dependência temporal entre as observações. O efeito espacial apresentou um padrão mais nítido, onde as capturas tendem a aumentar em áreas mais próximas do equador. A inferência realizada através do método INLA demonstrou ser uma alternativa plausível para o uso em grandes bases de dados, como a utilizada neste trabalho.*

Palavras-chave: *Modelos espaço-temporais, Captura, Inferência bayesiana, INLA.*

1 Introdução

Pescarias comerciais são atividades extrativistas, praticadas em todos os oceanos. Há grandes variações quanto às espécies capturadas e ao ambiente em que se encontram, desde pequenos animais bentônicos (moluscos e camarões) até grandes peixes pelágicos (atuns e tubarões). O impacto gerado pela extração de organismos pela pesca vai desde a alteração do ambiente físico (*e.g.* redes de arrasto passando sobre fundos coralinos) até a depleção dos estoques pesqueiros.

Em pescarias comerciais, frequentemente ocorrem capturas de espécies que não são o alvo principal, denominadas de “captura incidental”. No caso da pescaria com espinhel de superfície no oceano Atlântico, onde os alvos são os atuns e o espadarte, outras espécies de peixes são também atraídas pelas iscas nos anzóis e capturadas de forma não intencional.

A análise dos padrões de distribuição espaciais e temporais destas espécies é fundamental para o conhecimento tanto do estado atual dos estoques, quanto para que medidas de manejo

¹fernandomayer@gmail.com

²wagner@leg.ufpr.br

³paulojus@leg.ufpr.br

possam ser realizadas com a finalidade de diminuição das capturas incidentais. A identificação de áreas onde a captura de determinada espécie é muito provável, pode ser útil para se impedir ou diminuir o esforço de pesca sobre esta área. Nesse sentido, os modelos espaço-temporais são ferramentas que podem auxiliar na identificação de padrões que colaborem para o gerenciamento de estoques pesqueiros, que apresentam uma grande variabilidade espaço-temporal.

Normalmente, a modelagem da captura de uma determinada espécie é realizada sem qualquer suposição sobre a estrutura ou a dependência das observações no espaço e no tempo. No entanto, sabe-se por exemplo que as capturas de um ano são influenciadas, entre muitos outros fatores, pelas capturas realizadas no ano anterior. Biologicamente, a abundância de uma população no tempo t é influenciada pela abundância no tempo $t - 1$. Sabe-se ainda, que diferentes espécies tendem a ocupar uma determinada área por diversos motivos, como por exemplo a disponibilidade de comida e a temperatura da água, inclusive dependendo da época do ano. Por esses motivos, a dependência entre espaço, tempo e a interação espaço-tempo devem ser levadas em consideração na modelagem de dados de populações biológicas.

Neste trabalho foram ajustados e comparados modelos espaço-temporais para dados de captura do agulhão-negro (*Makaira nigricans*), uma espécie com população frágil, e que é capturada de forma incidental por barcos de espinhel de superfície no Atlântico sul. Os modelos analisados foram construídos considerando diferentes estruturas para o espaço e para o tempo.

2 Metodologia

Os dados de captura em número do agulhão-negro são provenientes do banco de dados *Task 2* da *International Commission for the Conservation of Atlantic Tunas* (ICCAT), entidade internacional responsável pelo gerenciamento das pescarias realizadas pelos diversos países que capturam atuns e espécies afins no Atlântico. Esta base de dados está disponível no endereço www.iccat.int. Foram utilizados dados da frota do Japão para os anos 1970–2009. Cada linha da base de dados contém informações como a posição geográfica, o trimestre, o esforço de pesca (número de anzóis) e a captura. As posições geográficas são agregadas em quadrados de $5^\circ \times 5^\circ$, e frequentemente contém mais de uma informação por ano e trimestre. Por isso foi criado um *grid* regular através da soma da captura em cada quadrado por ano e trimestre. Este *grid* corresponde à todos os quadrados de $5^\circ \times 5^\circ$ do oceano Atlântico abaixo da latitude 5°N , que corresponde à área ocupada pelo estoque sul do agulhão-negro. As combinações de ano/trimestre/área onde não haviam informações, foram criadas para que predições pudessem ser realizadas. Ao todo, foram utilizadas 28160 linhas de dados, sendo que 6557 (23%) continham informações.

Para analisar os padrões no espaço e no tempo das capturas do agulhão-negro no Atlântico sul, foram construídos modelos hierárquicos bayesianos com estrutura gaussiana latente para representar os efeitos espaciais, temporais e possíveis interações. Nesta abordagem, assume-se que a variável resposta “captura” (em forma de contagem do número de peixes), Y_{it} , com $i = 1, \dots, n$ áreas e $t = 1, \dots, T$ unidades de tempo (combinação de ano e trimestre), segue uma distribuição de Poisson com parâmetro λ_{it} . A contagem do número de peixes capturados foi realizada por número de anzóis (esforço de pesca), um valor conhecido porém não necessariamente constante. Por isso foi utilizado um *offset*, f_{it} , com o número de anzóis para considerar estes intervalos de observação desiguais. A ligação entre a variável resposta e o preditor linear η_{it} é dada por

$g(\lambda_{it}) = f_{it} \exp(\eta_{it})$ [7]. Como a função de ligação $g(\cdot)$ canônica é a logaritmica, o modelo latente completo é dado por $\log(\lambda_{it}) = \log(f_{it}) + \eta_{it}$, onde

$$\eta_{it} = \mu + \nu_i + \psi_i + \gamma_t + \phi_t + \delta_{it} + \rho_{it} \quad (1)$$

O parâmetro μ representa o nível médio (intercepto), ν_i é o componente espacial não estruturado i.i.d. normal com média zero e precisão desconhecida. Para considerar que contagens em um quadrado são relacionadas com as de seus vizinhos geograficamente próximos, define-se ψ_i como um componente espacial estruturado de acordo com um modelo condicional autoregressivo intrínscio (ICAR) [1]. Os componentes temporais são γ_t e ϕ_t , que são, respectivamente, não estruturado e estruturado através de um modelo dinâmico do tipo *random walk* de primeira ordem, que considera os efeitos dos tempos $t - 1$ e $t + 1$ em torno de t [2]. O parâmetro δ_{it} representa a interação não estruturada (i.i.d.) entre espaço e tempo, na forma de um efeito diferencial para cada observação, assumindo que observações vizinhas tanto no espaço como no tempo são independentes. O termo ρ_{it} captura o efeito de interação espaço-tempo estruturado, assumindo que vizinhos tanto espaciais como temporais não são independentes. Este efeito é representado no modelo por uma distribuição normal multivariada com matriz de precisão dada pelo produto Kronecker da matriz de precisão do efeito espacial e temporal estruturada [2].

As estimativas dos parâmetros do modelo foram realizadas através de inferência bayesiana, utilizando a integração aproximada aninhada de Laplace (INLA) descrita por Rue *et al.* (2009) [5], que representa uma nova abordagem para inferência em campos aleatórios gaussianos markovianos (GMRF) [6]. Para a comparação entre os modelos foi utilizado o critério de informação da deviança (DIC). Os ajustes foram feitos através do pacote INLA [4], para o ambiente R [3].

3 Resultados

A comparação entre os modelos ajustados está demonstrada na Tabela 1. O modelo 3, com covariáveis estruturadas apresentou uma diminuição pequena no DIC em relação ao modelo 2, sem estrutura espaço-temporal. O menor valor de DIC foi observado no modelo 4, que contém o componente espacial e temporal estruturados, e interação entre espaço e tempo independente (não estruturada). O modelo 5, com estrutura em todos os termos, inclusive na interação, apresentou o DIC extremamente elevado em comparação com os demais modelos.

Tabela 1: Deviança residual, número de parâmetros (NP), critério de informação da deviança (DIC), e ordenada preditiva condicional (CPO) para os diferentes modelos ajustados.

	Modelo	Deviança	NP	DIC	CPO
1	$Y \sim 1$	235742.65	3.52	235746.17	18.00
2	$Y \sim \nu_i + \gamma_t$	62356.00	285.61	62690.34	5.06
3	$Y \sim \psi_i + \phi_t$	62310.02	261.15	62576.70	4.95
4	$Y \sim \psi_i + \phi_t + \delta_{it}$	21586.32	4181.00	26473.79	0.01
5	$Y \sim \psi_i + \phi_t + \rho_{it}$	1.42E+47	3123.79	2.83E+47	0.95

Não foi detectado padrão marcante na tendência da captura do agulhão-negro ao longo do tempo. Inicialmente há um declínio de 1970 até o final de 1976. Após esse período, o efeito

tende a ser muito próximo de zero (*i.e.* nulo), com uma leve tendência de aumento até o ano 2000. A partir de então há um pequeno declínio e uma nova tendência de aumento das capturas a partir do final de 2004 até 2009 (Figura 1).

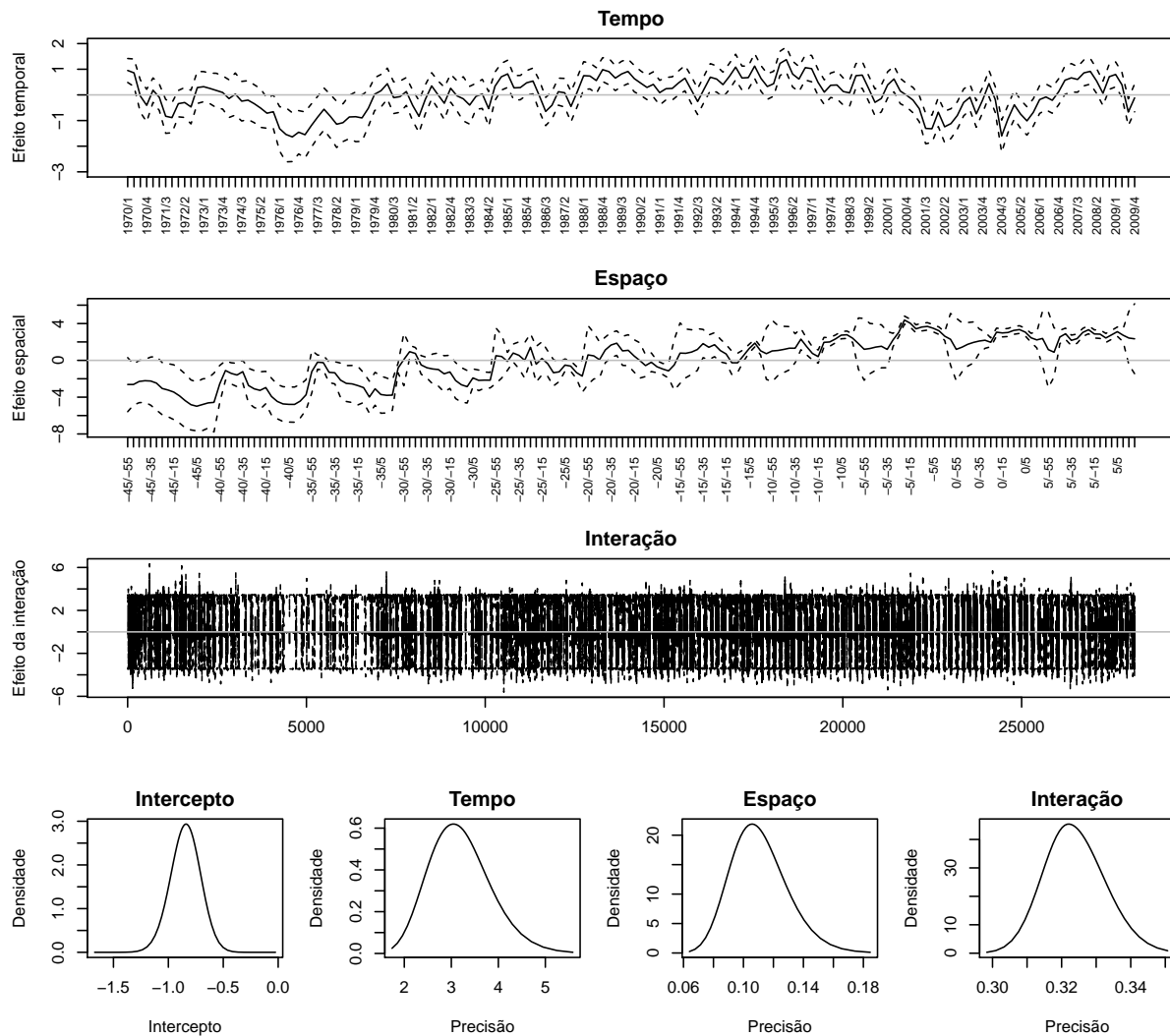


Figura 1: Painéis superiores (de cima para baixo): efeito temporal (ano/trimestre), espacial (latitude/longitude), e da interação espaço-temporal sobre as capturas do agulhão-negro. Painel inferior (da esquerda para a direita): distribuição posterior do intercepto, e precisão dos componentes do tempo, espaço, e da interação espaço-tempo.

As capturas apresentam um claro padrão espacial. Como a malha utilizada na estrutura do modelo tem início no ponto extremo oeste/sul e segue até o ponto extremo leste/norte, há uma tendência de aumento na captura em direção às latitudes mais próximas do equador. A latitude 20°S mostra-se como um ponto de divisão, com capturas maiores em latitudes superiores, e menores em latitudes inferiores (Figura 1).

Inferências sobre os parâmetros de precisão dos componentes do modelo mostram que o efeito do tempo parece ser dominante, com capturas dependentes temporalmente. A precisão do termo espacial foi a menor observada entre todos os componentes. A interação espaço-temporal teve a precisão um pouco maior do que a do espaço (Figura 1).

4 Discussão

Inferências sobre os efeitos do modelo analisado permitem concluir que não há uma clara tendência temporal no estoque do agulhão-negro no Atlântico sul, porém existe uma forte dependência temporal entre as observações. Nota-se que as capturas tendem a ser maiores em latitudes mais próximas do equador, provavelmente devido ao fato de que esta espécie se concentra em águas tropicais em períodos de reprodução, ficando assim mais vulneráveis à captura.

A análise de dados de pescarias comerciais normalmente envolve a padronização da captura como a primeira etapa para análises subsequentes. Normalmente são utilizados modelos lineares generalizados (MLGs) para essa tarefa, onde, no entanto, a estrutura espacial e temporal raramente é levada em consideração através de estruturas adequadas. A utilização de modelos espaço-temporais representa uma alternativa mais acurada para a modelagem deste tipo de dados, uma vez que as capturas em pescarias comerciais são dependentes no tempo e no espaço.

A realização de inferência bayesiana através do método INLA representa uma alternativa viável e computacionalmente mais eficiente do que os tradicionais métodos de Monte Carlo via cadeias de Markov (MCMC). Especialmente em dados de pescarias comerciais, como os analisados neste trabalho, onde há uma grande quantidade de informação, e possivelmente também um grande número de covariáveis, este método pode ser uma opção mais plausível para os cálculos da inferência bayesiana na análise e seleção de modelos para bases de dados com várias espécies e grande número de observações.

Referências

- [1] BESAG, J.; YORK, J; MOLLIE, A., Bayesian image restoration with two applications in spatial statistics, *Statistics in Medicine*, v. 14, p. 2433–2443, 1991.
- [2] KNORR-HELD, L., Bayesian modelling of inseparable space-time variation in disease risk, *Statistics in Medicine*, v. 19, p. 2555–2567, 2000.
- [3] R DEVELOPMENT CORE TEAM, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2011, ISBN 3-900051-07-0, URL www.R-project.org.
- [4] RUE, H.; MARTINO, S., INLA: Functions which allow to perform a full Bayesian analysis of structured additive models using Integrated Nested Laplace Approximation. R package version 0.0, 2009
- [5] RUE, H.; MARTINO, S.; CHOPIN, N., Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the Royal Statistical Society B*, v. 71, p. 319–392, 2009.
- [6] RUE, H.; HELD, L., *Gaussian Markov Random Fields*, Boca Raton: Chapman & Hall/CRC, 2005.
- [7] SCHRÖDLE, B.; HELD, L., Spatio-temporal disease mapping using INLA, *Environmetrics*, 2010.