

Exemplo Espinhel de Fundo

Gilberto A. Paula

Departamento de Estatística
IME-USP, Brasil
giapaula@ime.usp.br

2^o Semestre 2023

- 1 Espinhel de Fundo
- 2 Análise de Dados Preliminar
- 3 Modelo Gama
- 4 Modelo com Interação
- 5 Conclusões
- 6 Referências

Descrição do estudo

Considere uma amostra piloto referente a um estudo sobre a atividade das frotas pesqueiras de **espinhel de fundo** baseadas em Santos e Ubatuba no litoral paulista (vide Paula e Oshiro, 2001).

O espinhel de fundo é definido como um método de pesca passivo, sendo utilizado em todo o mundo em operações de pesca de diferentes magnitudes, da pesca artesanal a modernas pescarias mecanizadas. A espécie de peixe considerada é o **peixe-batata** pela sua importância comercial e ampla distribuição espacial.

Descrição dos Dados

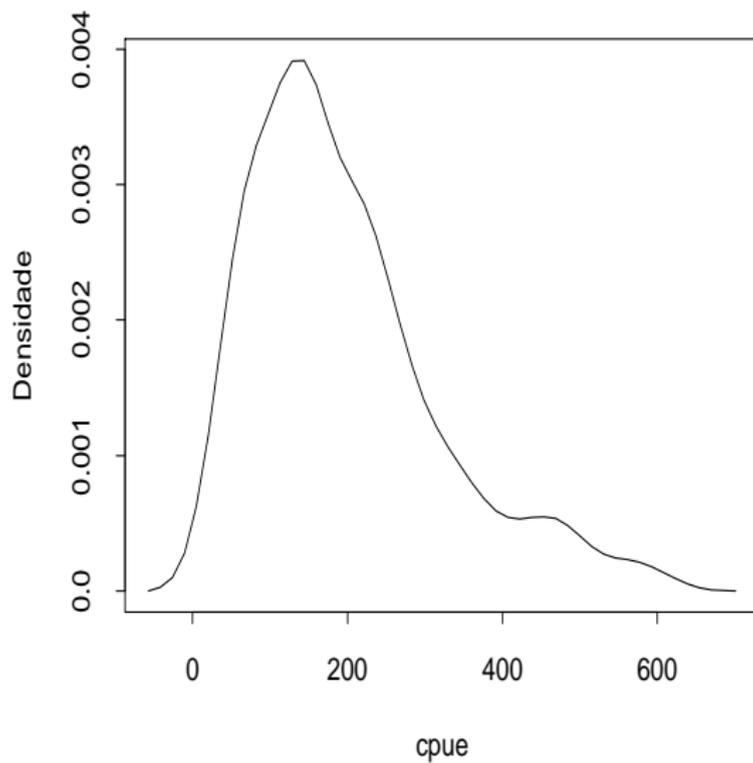
Uma amostra de $n = 156$ embarcações foi analisada no período de 1995 a 1999 sendo 39 da frota de Ubatuba e 117 da frota de Santos. As variáveis consideradas para cada embarcação são as seguintes:

- **Frota** (Santos ou Ubatuba),
- **Ano** (95 a 99),
- **Trimestre** (1 ao 4),
- **Latitude** (de $23,25^{\circ}$ a $28,25^{\circ}$),
- **Longitude**(de $41,25^{\circ}$ a $50,75^{\circ}$),
- **Dias de pesca**,
- **Captura** (quantidade de peixes batata capturados, em kg),
- **CPUE** (captura por unidade de esforço, kg/dias de pesca).

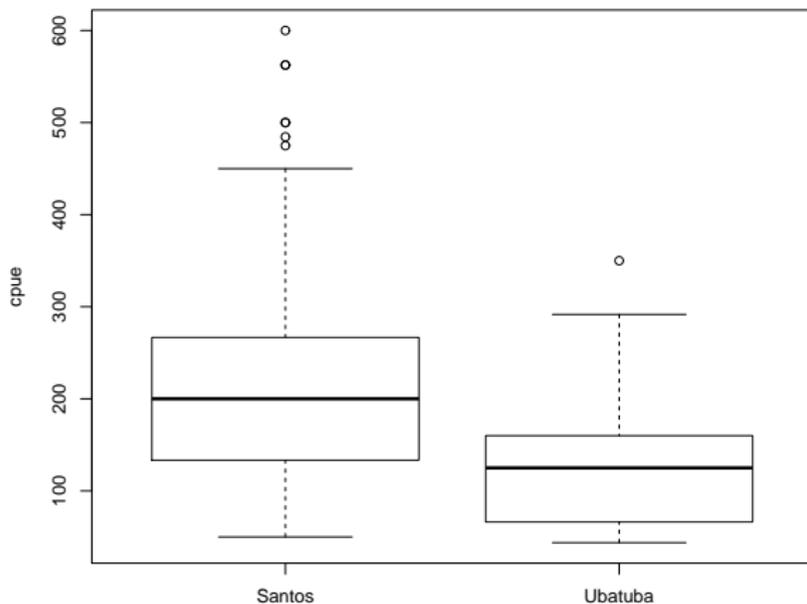
Objetivo do Estudo

Objetivo principal deste estudo é explicar a **cpue média** segundo as variáveis **frota**, **ano**, **trimestre**, **latitude** e **longitude**.

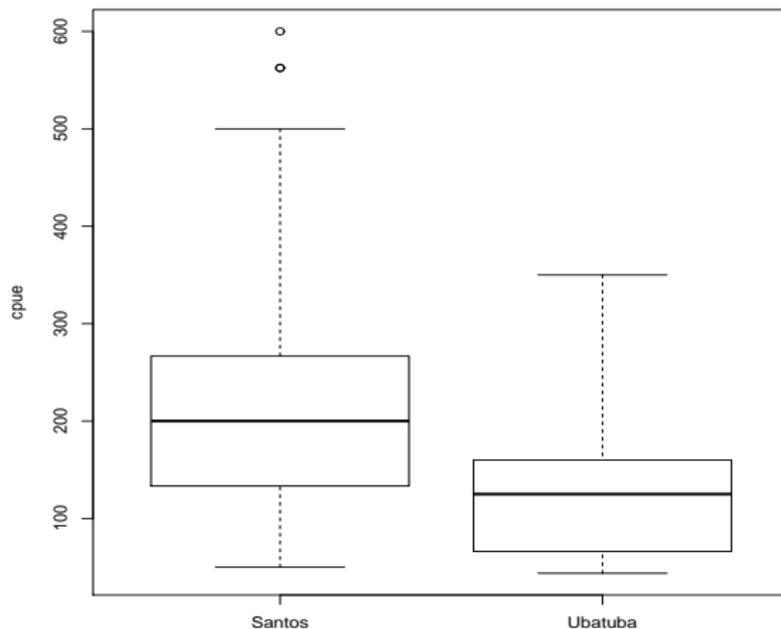
- 1 Espinhel de Fundo
- 2 Análise de Dados Preliminar**
- 3 Modelo Gama
- 4 Modelo com Interação
- 5 Conclusões
- 6 Referências



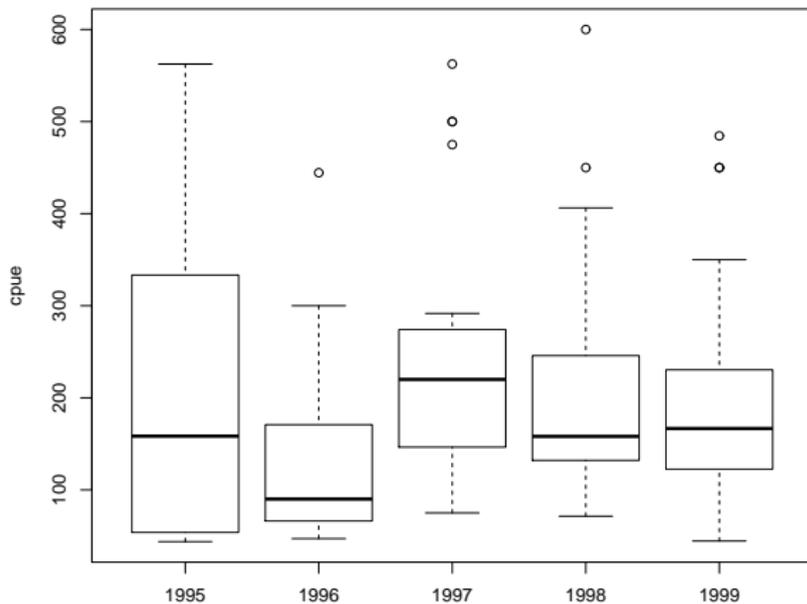
Boxplots da cpue segundo a Frota



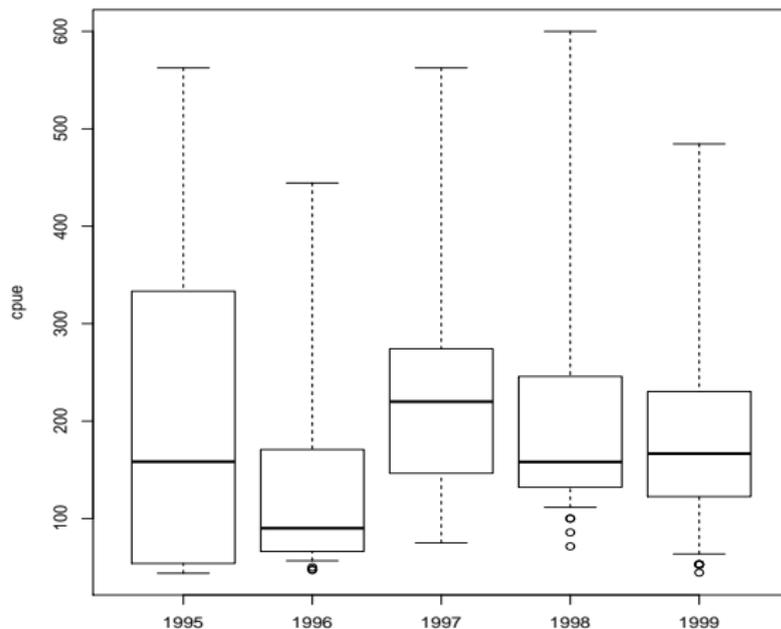
Boxplots robusto da cpue segundo a Frota



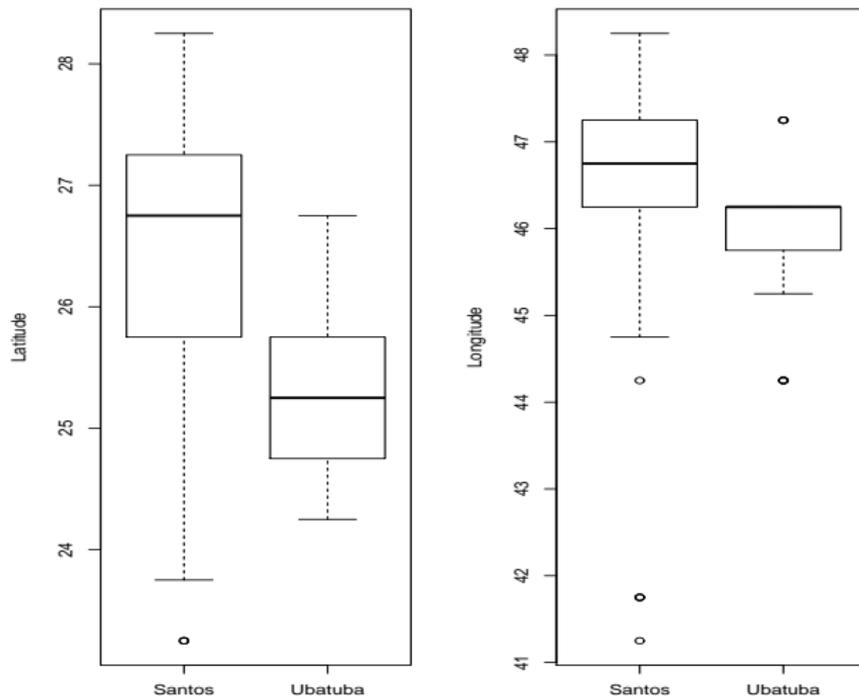
Boxplots da cpue segundo o Ano



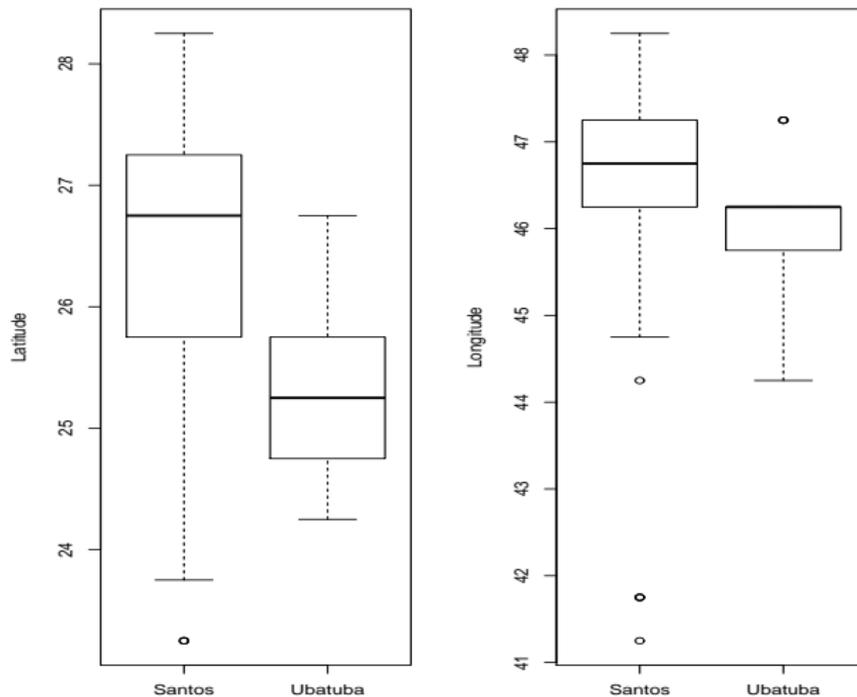
Boxplots robusto da cpue segundo o Ano



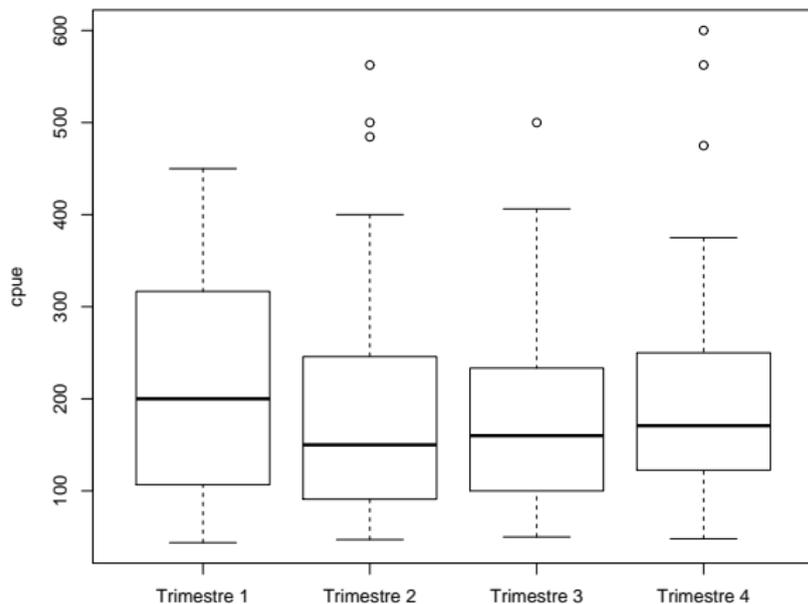
Boxplots da Latitude e Longitude segundo a Frota



Boxplots robusto da Latitude e Longitude segundo a Frota



Boxplots da cpue segundo o Trimestre



Boxplots robusto da cpue segundo o Trimestre

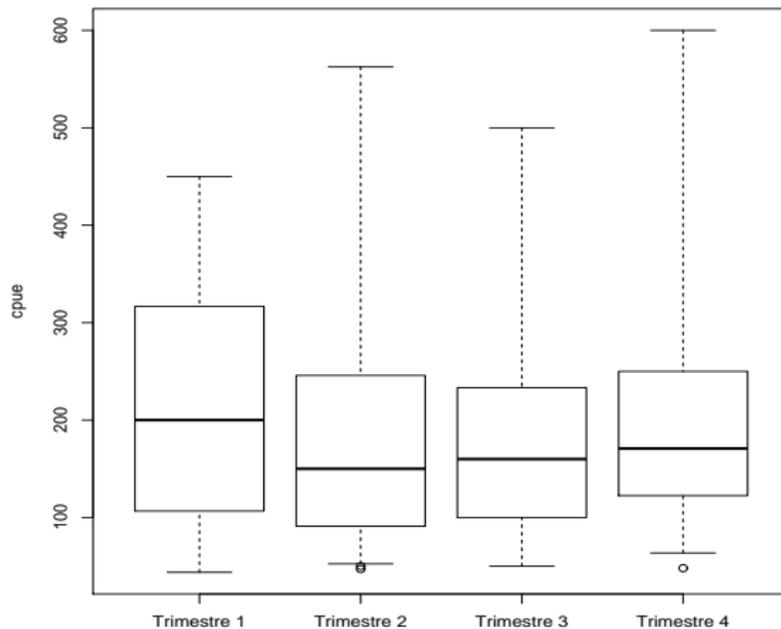


Diagrama de dispersão da cpue contra a Latitude

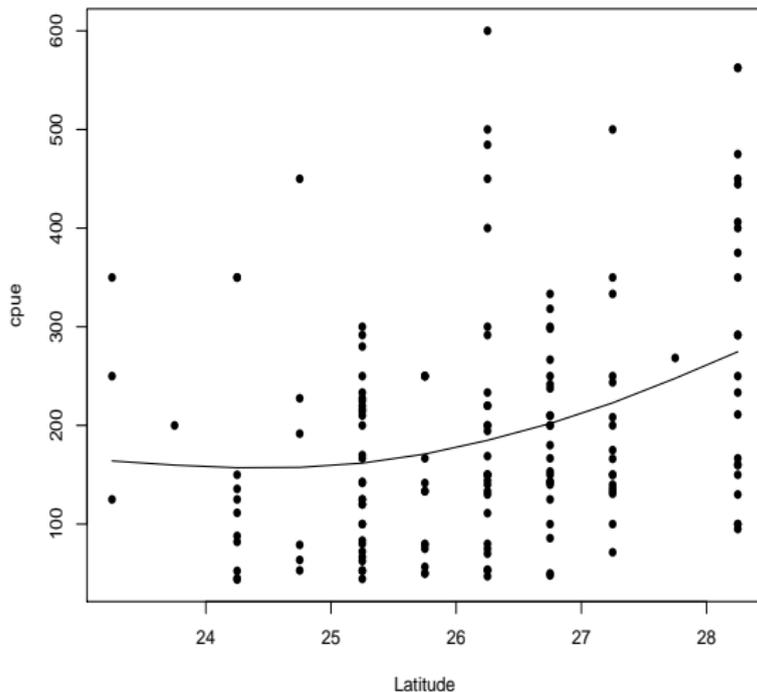


Diagrama de dispersão da cpue contra a Longitude

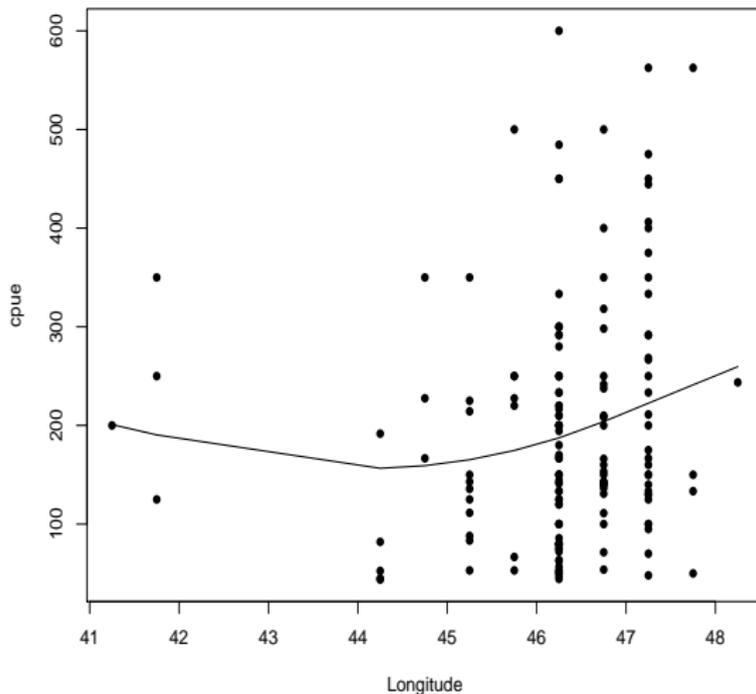
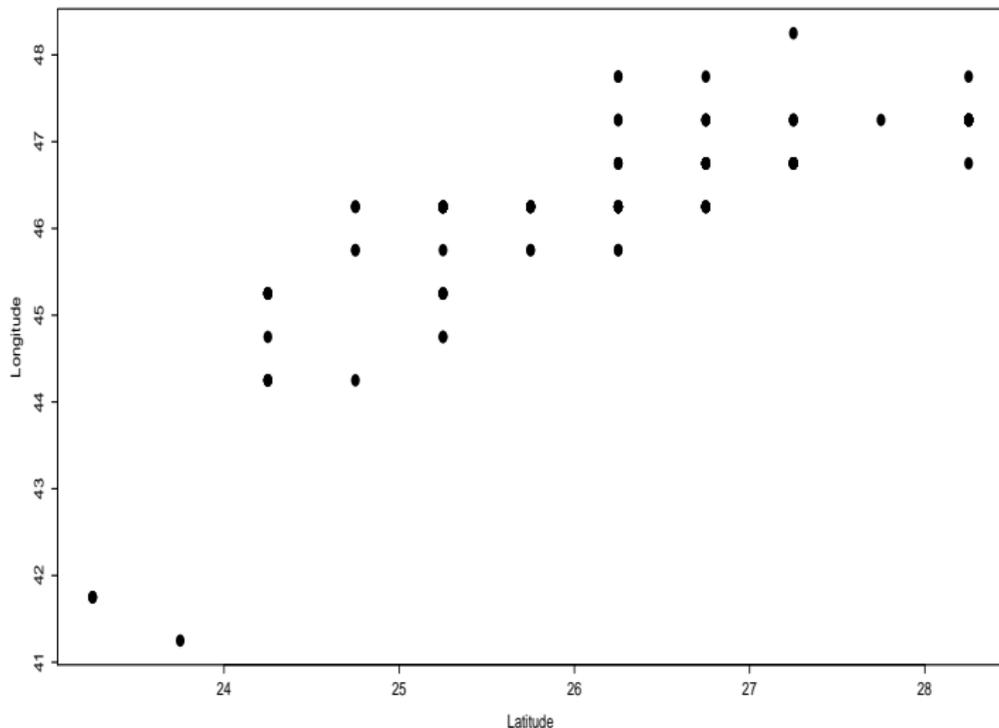


Diagrama de dispersão entre Latitude e Longitude¹



¹1 grau de latitude ou longitude \cong 111 km

Comentários

- Aplicando-se o **boxplot robusto (Hubert e Vandervieren, 2008)**, que faz correção do boxplot levando-se em conta a assimetria, nota-se uma redução no número de observações aberrantes na parte superior dos gráficos e um aumento na parte inferior desses gráficos.
- Nota-se também uma tendência crescente da cpue com o aumento da latitude, porém com relação à longitude há inicialmente uma tendência decrescente e depois crescente da cpue com o aumento da longitude.
- Seguindo estudos preliminares da época do experimento realizados na costa mexicana e na costa mediterrânea, será aplicado um **modelo log-linear gama**. Ou seja, um MLG com resposta gama e ligação logarítmica.

- 1 Espinhel de Fundo
- 2 Análise de Dados Preliminar
- 3 Modelo Gama**
- 4 Modelo com Interação
- 5 Conclusões
- 6 Referências

Descrição

Seja $y_{ijk\ell}$ a cpue observada para a i -ésima embarcação da j -ésima frota (Santos ($j = 1$); Ubatuba($j = 2$)), no k -ésimo ano e ℓ -ésimo trimestre, ($k = 1, 2, 3, 4, 5$) e ($\ell = 1, 2, 3, 4$). Considere o seguinte modelo casela de referência:

- $Y_{ijk\ell} | \mathbf{x}_{ijk\ell} \stackrel{\text{ind}}{\sim} G(\mu_{ijk\ell}, \phi)$,
- $\log(\mu_{ijk\ell}) = \alpha + \beta_j + \gamma_k + \theta_\ell + \delta_1 \times \text{Latitude}_{ijk\ell} + \delta_2 \times \text{Longitude}_{ijk\ell}$,

em que β_j denota o efeito de frota ($\beta_1 = 0$ (Santos), β_2 : Ubatuba), γ_k denota o efeito ano ($\gamma_1 = 0$ (95), γ_2 : 96, γ_3 : 97, γ_4 : 98, γ_5 : 99) e θ_ℓ denota o efeito trimestre ($\theta_1 = 0$ (trimestre 1), θ_2 : trimestre 2, θ_3 : trimestre 3, θ_4 : trimestre 4).

Resultado

Aplicando-se o método de Akaike retiramos o fator Trimestre, permanecendo no modelo os fatores **Frota** e **Ano** além das variáveis quantitativas **Latitude** e **Longitude**.

Descrição

Efeito	Estimativa	E. Padrão	Valor-z
Constante	5,994	2,257	2,66
Latitude	0,166	0,072	2,30
Longitude	-0,110	0,076	-1,45
Frota-Ubatuba	-0,276	0,132	-2,09
Ano96	-0,153	0,188	-0,81
Ano97	0,330	0,174	1,90
Ano98	0,116	0,161	0,72
Ano99	0,070	0,147	0,48
ϕ	3,383	0,366	9,24

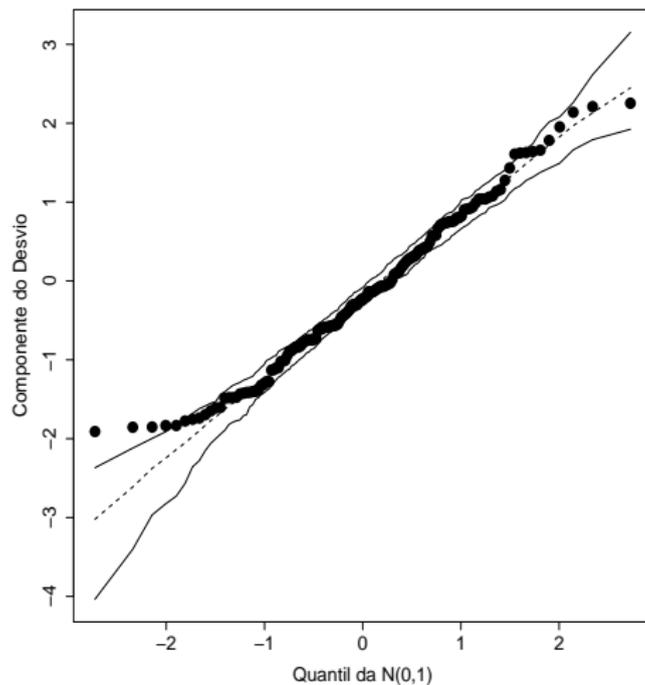
Qualidade do Ajuste

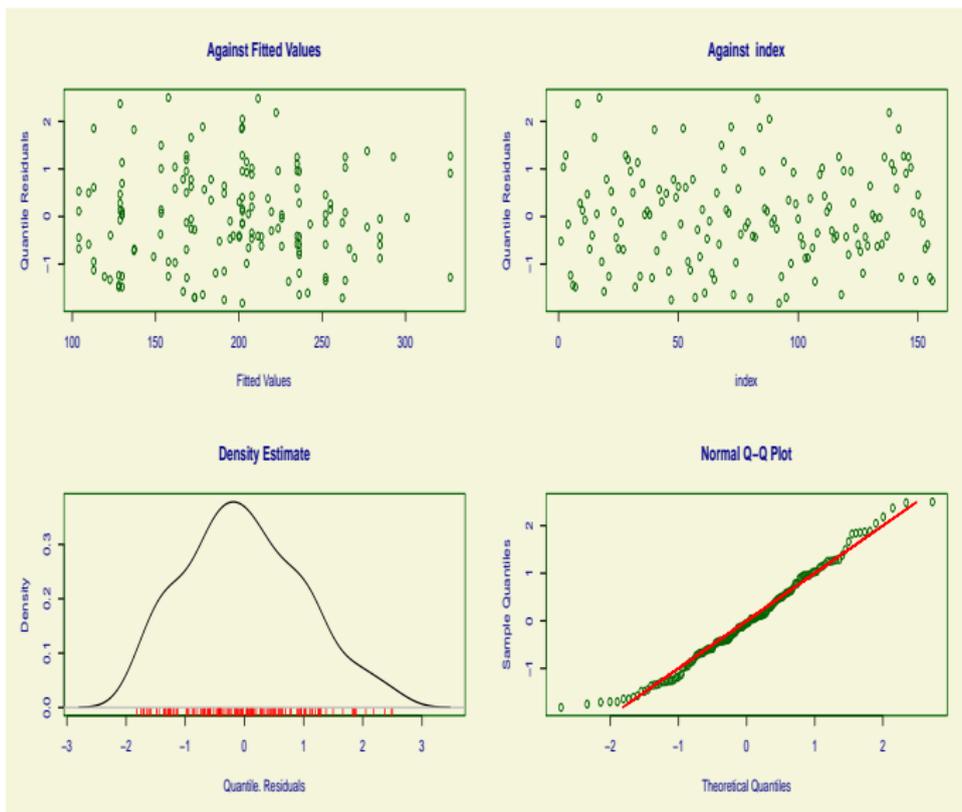
Obtemos $\hat{\phi} = 3,383(0,366)$ confirmando a assimetria à direita para a distribuição da cpue.

O desvio do modelo foi de $D^*(\mathbf{y}; \hat{\boldsymbol{\mu}}) = 163,62$ para 148 graus de liberdade com $P = 0,18$, indicando um modelo bem ajustado.

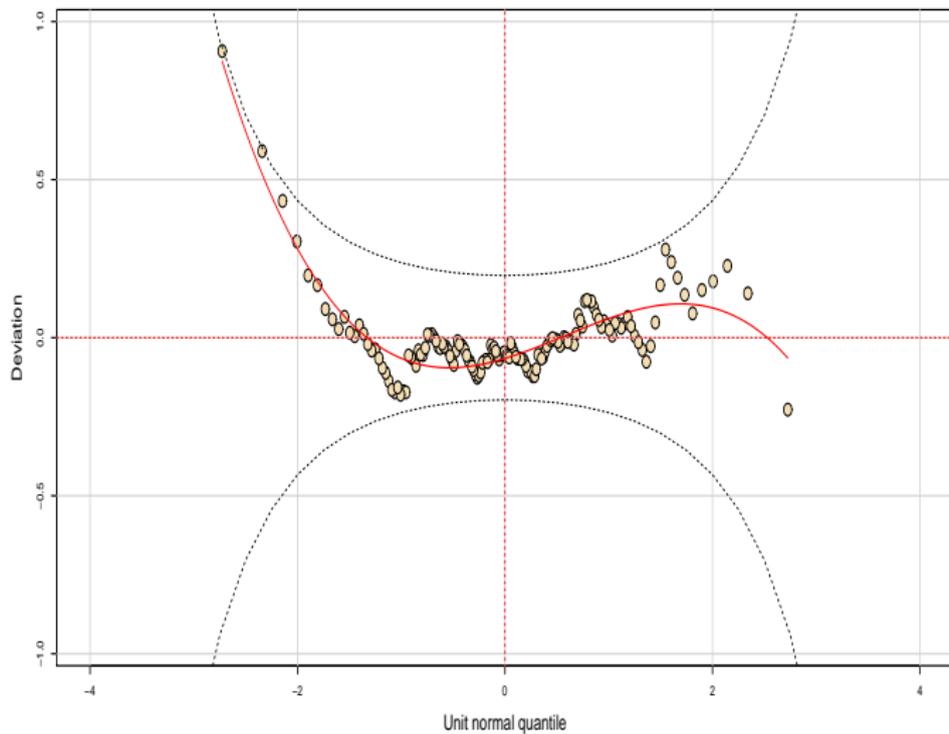
Tem-se que $R^2 = 1 - \frac{D(\mathbf{y}; \hat{\boldsymbol{\mu}})}{D(\mathbf{y}; \bar{\mathbf{y}})} = \frac{50,924}{59,362} = 0,1421$.

Resíduos Modelo Ajustado





Worm Plot GAMLSS



- 1 Espinhel de Fundo
- 2 Análise de Dados Preliminar
- 3 Modelo Gama
- 4 Modelo com Interação**
- 5 Conclusões
- 6 Referências

Interação

O teste da razão de verossimilhanças para incluir a interação entre os dois fatores que permaneceram no modelo, **Frota** e **Ano**, foi de $\xi_{RV} = 14,26$ para 4 graus de liberdade obtendo-se $P = 0,006$. Portanto, a interação será incluída no modelo.

Descrição

Portanto, o seguinte modelo foi selecionado:

- $Y_{ijk} | \mathbf{x}_{ijk} \stackrel{\text{ind}}{\sim} \mathbf{G}(\mu_{ijk}, \phi),$
- $\log(\mu_{ijk}) = \alpha + \beta_j + \gamma_k + \delta_1 \times \text{Latitude}_{ijk} + \delta_2 \times \text{Longitude}_{ijk} + \theta_{jk},$

em que y_{ijk} denota a cpue observada para a i -ésima embarcação da j -ésima frota e no k -ésimo ano, $j = 1, 2$ e $k = 1, 2, 3, 4, 5$.

Os parâmetros θ_{jk} denotam a interação entre frota e ano com a seguinte especificação: $\theta_{22} : \text{FrotaUb}^* \text{Ano96}$, $\theta_{23} : \text{FrotaUb}^* \text{Ano97}$, $\theta_{24} : \text{FrotaUb}^* \text{Ano98}$ e $\theta_{25} : \text{FrotaUb}^* \text{Ano99}$. Os demais parâmetros são nulos.

Descrição

Efeito	Estimativa	E. Padrão	Valor-z
Constante	6,897	2,302	3,00
Latitude	0,204	0,072	2,83
Longitude	-0,150	0,076	-1,97
Frota-Ubatuba	-1,359	0,369	-3,68
Ano96	-0,064	0,242	-0,26
Ano97	0,141	0,189	0,75
Ano98	-0,043	0,169	-0,25
Ano99	-0,009	0,156	-0,06
FrotaUb*Ano96	0,806	0,455	1,77
FrotaUb*Ano97	1,452	0,454	3,20
FrotaUb*Ano98	1,503	0,453	3,32
FrotaUb*Ano99	1,113	0,403	2,76
ϕ	3,678	0,399	9,22

Qualidade do Ajuste

Obtemos $\hat{\phi} = 3,678(0,399)$ confirmando a assimetria à direita para a distribuição da cpue.

O desvio do modelo foi de $D^*(\mathbf{y}; \hat{\boldsymbol{\mu}}) = 162,66$ para 144 graus de liberdade com $P = 0,14$, indicando um modelo bem ajustado.

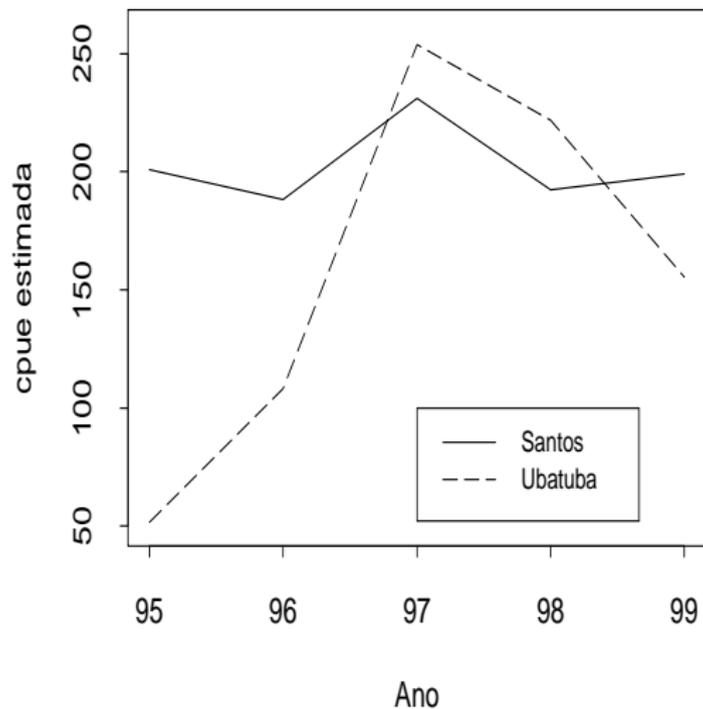
Tem-se que $R^2 = 1 - \frac{D(\mathbf{y}; \hat{\boldsymbol{\mu}})}{D(\mathbf{y}; \bar{\mathbf{y}})} = \frac{49,464}{59,362} = 0,1667$.

Efeitos principais e interação

Nota-se que à medida que aumenta a latitude espera-se um aumento da cpue, ocorrendo tendência contrária à medida que aumenta a longitude. Logo, para latitudes altas e longitudes baixas (dentro dos limites amostrais), espera-se valores maiores para a cpue.

A interação entre frota e ano é significativa, isto é, a diferença entre as cpue's médias das duas frotas não é constante ao longo dos anos.

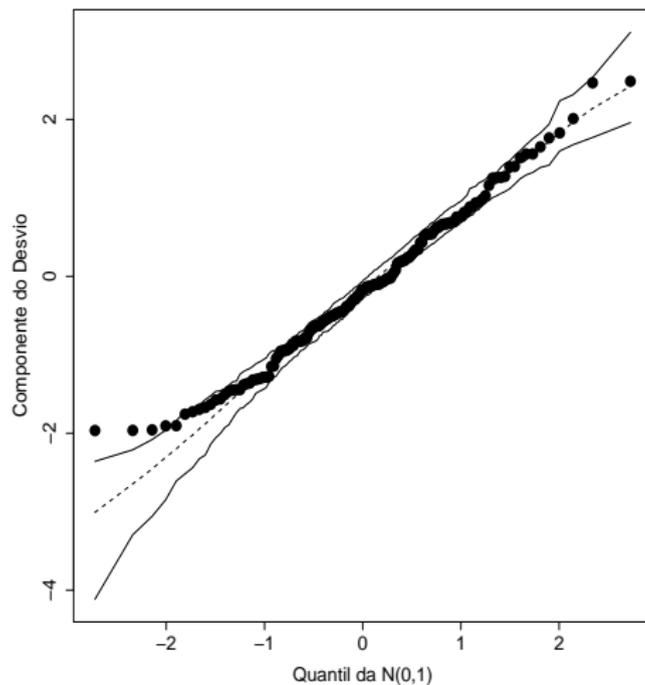
Cpue Estimada fixando Latitude e Longitude



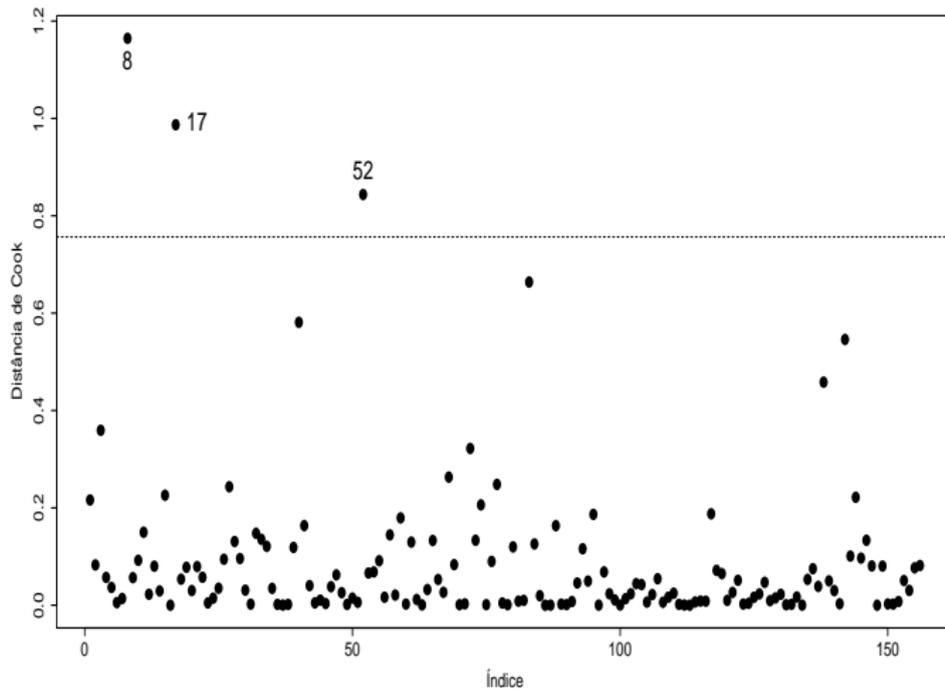
Cpue Estimada

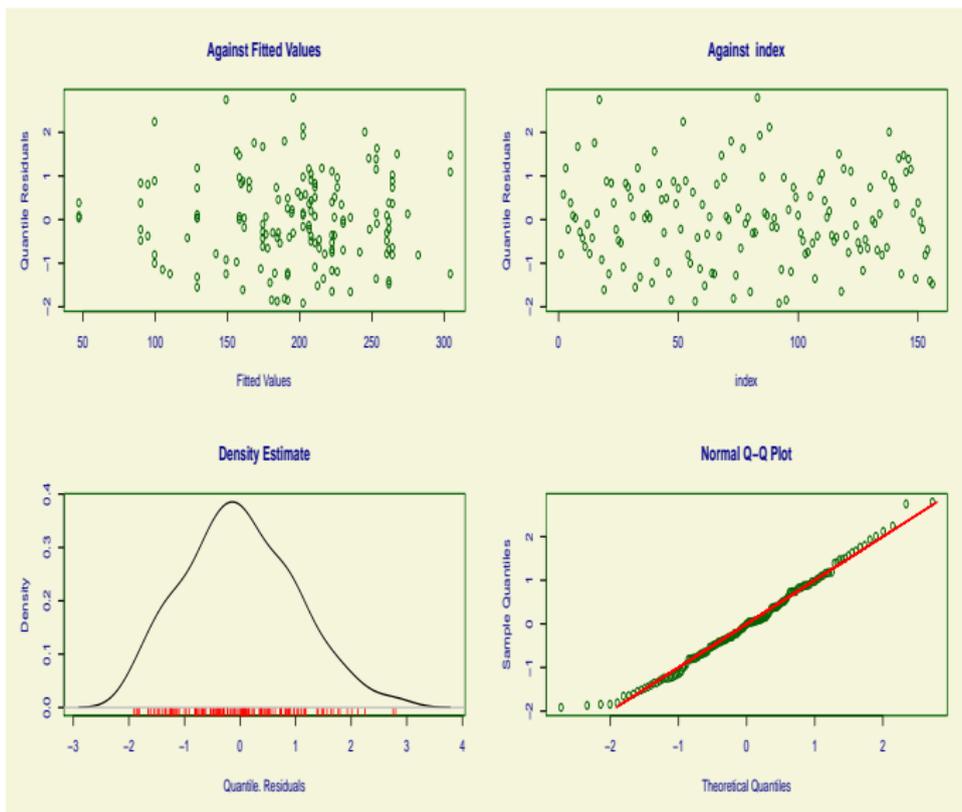
As cpues médias foram estimadas fixando latitude e longitude nos valores 26° e 46° , respectivamente. Pode-se observar o seguinte:

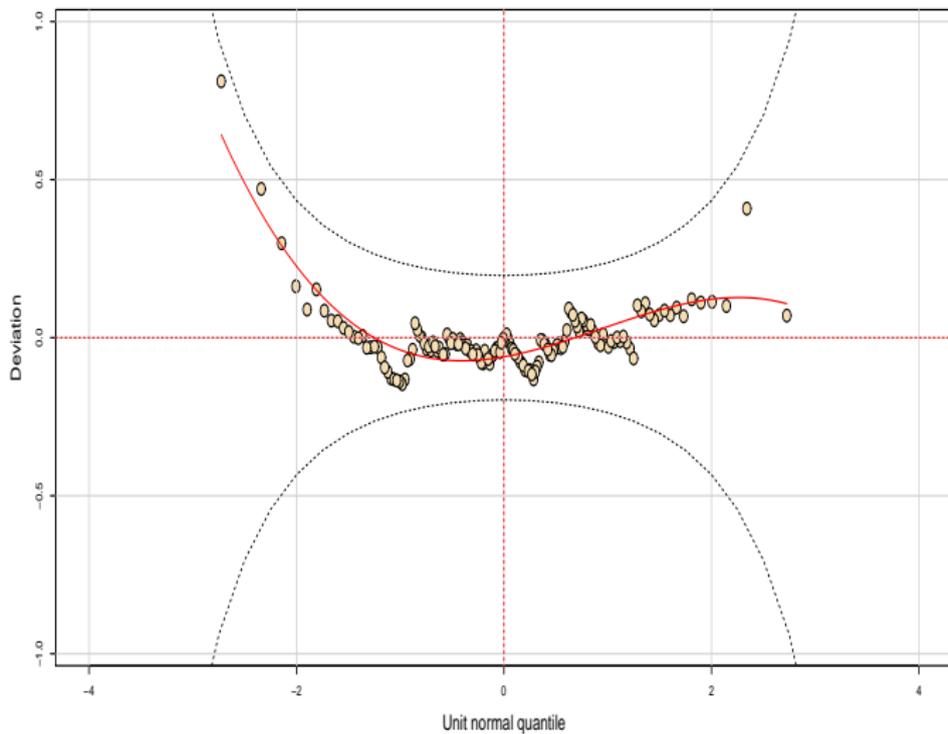
- até 96 os valores preditos para a frota de Ubatuba são bem menores do que os valores preditos para a frota de Santos,
- a partir de 97 as diferenças entre os valores médios preditos para as duas frotas diminuem,
- os valores preditos para a frota de Santos variam pouco no período 95-99, diferentemente dos valores preditos para a frota de Ubatuba.



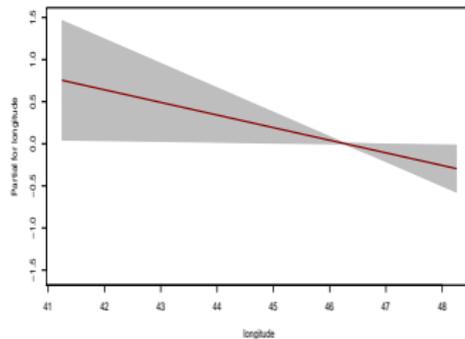
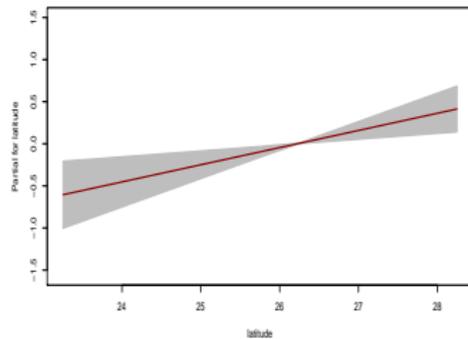
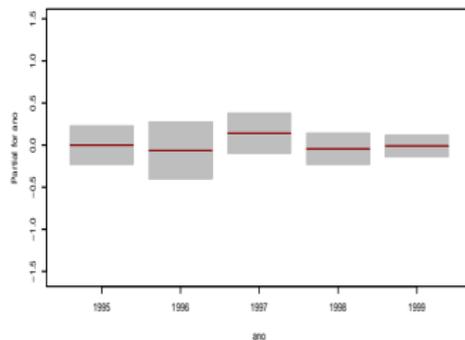
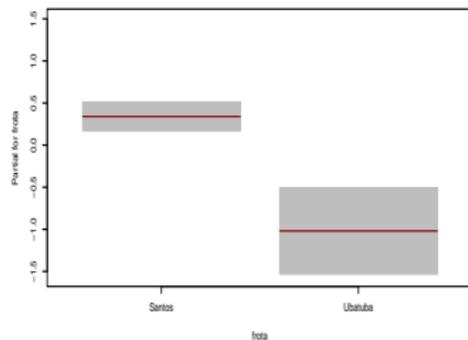
Distância de Cook Modelo Final







Term Plot GAMLSS

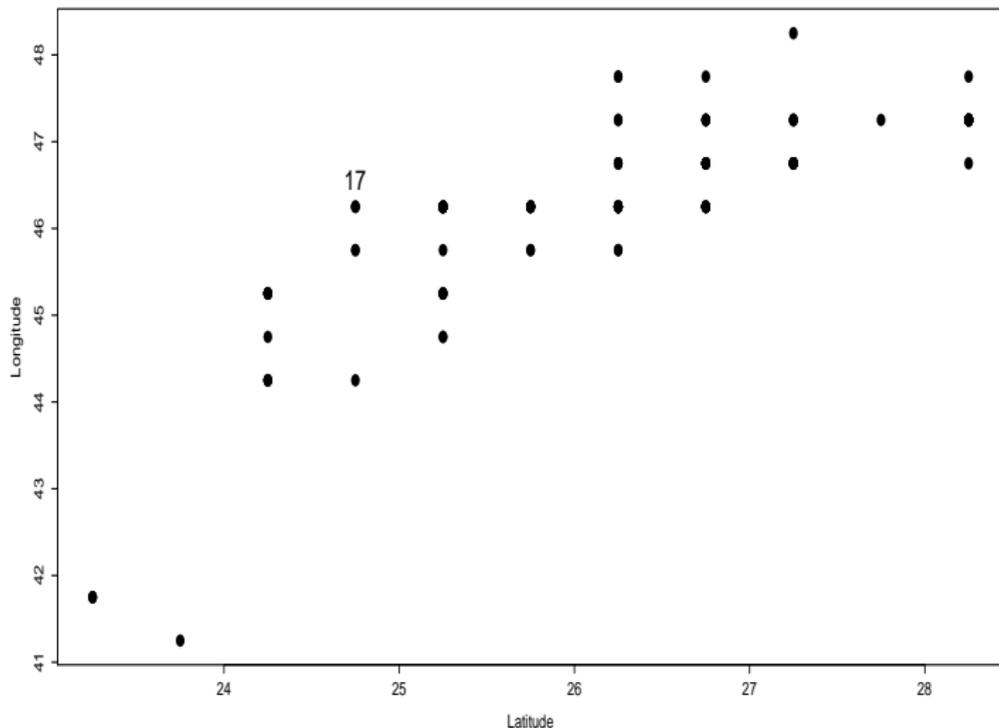


Interpretações

Podemos observar o seguinte:

- Três observações aparecem como possivelmente influentes, **as embarcações #8, #17 e #52**. A retirada de cada embarcação individualmente não muda a inferência. Porém a retirada da observação #17 aumenta a significância da latitude e longitude.
- A embarcação #17 é da frota de Santos, obteve uma cpue de **450** (valor médio 195,5) numa latitude de **$24,75^\circ$** (valor médio $26,22^\circ$) e longitude de **$46,25^\circ$** (valor médio $46,26^\circ$) no ano de 99.
- Esperava-se para essa embarcação um valor menor para a cpue levando-se em conta os valores da latitude e longitude. Trata-se portanto de uma embarcação atípica.

Diagrama de dispersão entre Latitude e Longitude²



²1 grau de latitude ou longitude \cong 111 km

Descrição

Efeito	Estimativa	E. Padrão	Valor-z
Constante	7,558	2,239	3,38
Latitude	0,259	0,072	3,61
Longitude	-0,196	0,074	-2,63
Frota-Ubatuba	-1,316	0,359	-3,67
Ano96	-0,038	0,235	-0,16
Ano97	0,140	0,183	0,77
Ano98	-0,034	0,164	-0,21
Ano99	-0,028	0,152	-0,18
FrotaUb*Ano96	0,801	0,442	1,81
FrotaUb*Ano97	1,490	0,441	3,38
FrotaUb*Ano98	1,539	0,440	3,50
FrotaUb*Ano99	1,146	0,391	2,93
ϕ	3,829	0,417	9,18

- 1 Espinhel de Fundo
- 2 Análise de Dados Preliminar
- 3 Modelo Gama
- 4 Modelo com Interação
- 5 Conclusões**
- 6 Referências

Considerações finais

- As cpue's médias nos anos de 95 a 99 para as embarcações de Santos e Ubatuba são bem ajustadas por um modelo gama com ligação logarítmica e dependem da localização da embarcação.
- Esses resultados confirmam estudos anteriores desenvolvidos com embarcações baseadas nas costas mediterrânea e mexicana.
- No ajuste pelo GAMLSS é disponibilizada a estimativa do coeficiente de variação ao invés da estimativa da precisão que é disponibilizada pelo R.

- 1 Espinhel de Fundo
- 2 Análise de Dados Preliminar
- 3 Modelo Gama
- 4 Modelo com Interação
- 5 Conclusões
- 6 Referências**

Referências

- Goñi, R., Alvarez, F. e Adlerstein, S. (1999). Application of generalized linear modeling to catch rate analysis of western mediterranean fisheries: the Castellón trawl fleet as a case study. *Fisheries Research* **42**, 291-302.
- Hubert, M. e Vandervierin, E. (2008). An adjusted boxplot for skewed distributions. *Computational Statistics and Data Analysis* **32**, 5186-5201.
- Paula, G. A. e Oshiro, C. H. (2001). Relatório de Análise Estatística sobre o Projeto: *Análise de Captura por Unidade de Esforço do Peixe-Batata na Frota Paulista*. RAE-CEA0102, IME-USP.
- Wakida-Kusunoki, A. T. e Solana-Sansores, R. (1997). Modelación del esfuerzo pesquero de la flota camaronera de Ciudad del Carmen, Campeche, México. *Oceánides* **12**, 107-116.