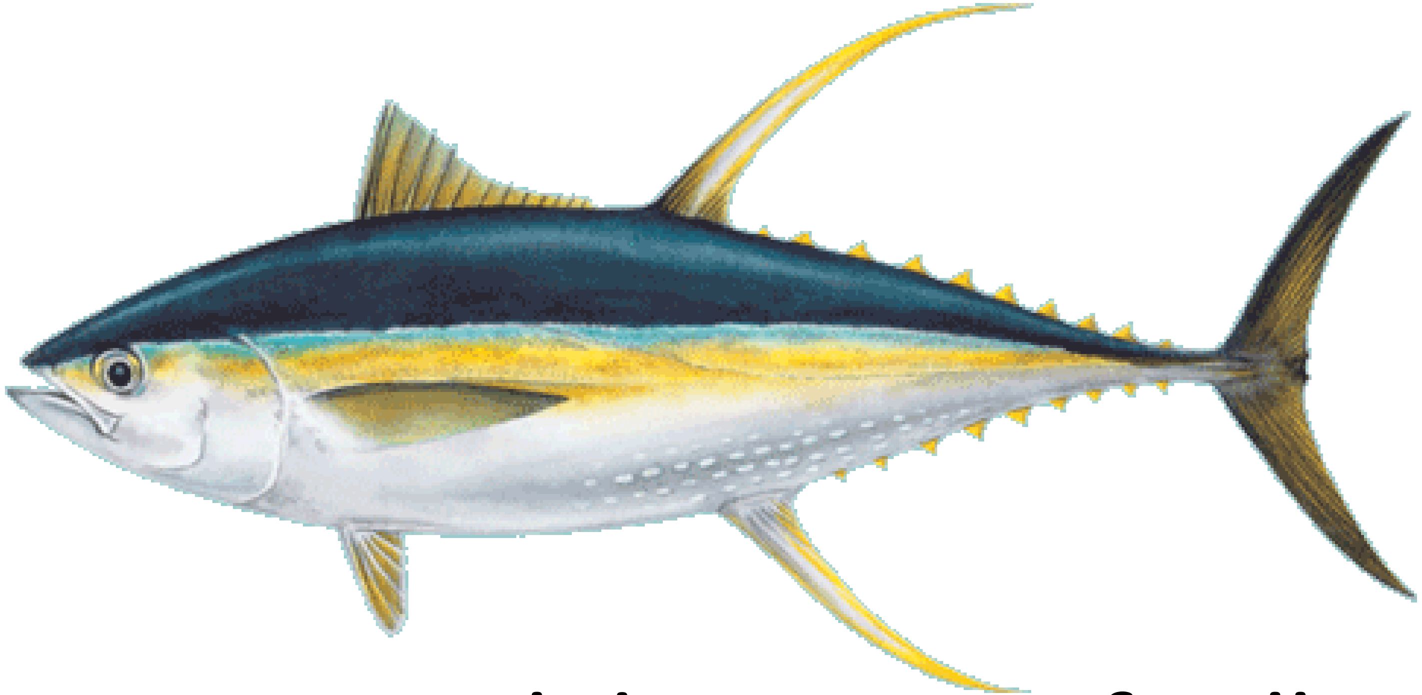


The spatiotemporal dynamics of yellowfin and bigeye tuna in the Eastern Pacific Ocean

Haikun Xu and Peter Kuriyama
Inter-American Tropical Tuna Commission
Scripps Institution of Oceanography

CAPAM Spatiotemporal Workshop
Feb. 28, 2018

Acknowledgements

- Cleridy Lennert-Cody, Mark Maunder, and Carolina Minte-Vera
- Jim Thorson
- Jin Gao

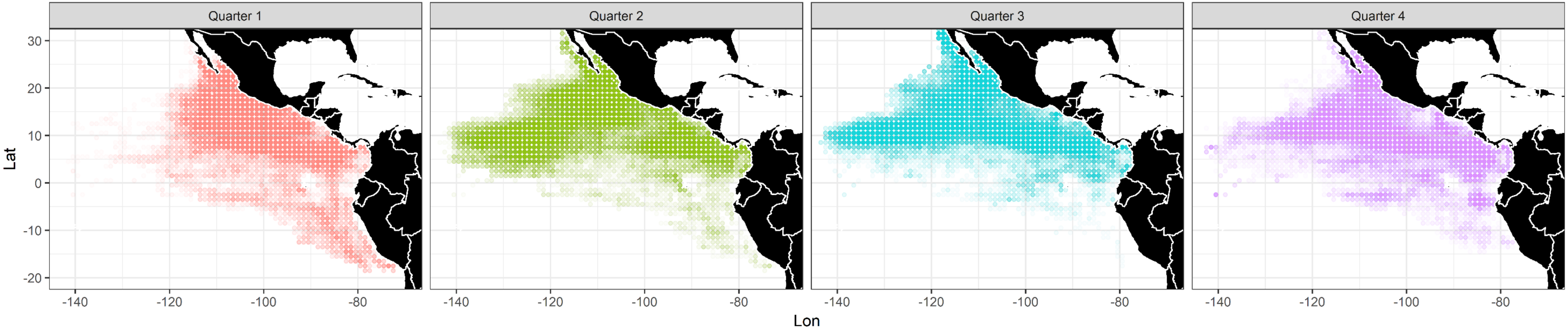
Topics

- Apply VAST to the aggregate catch rate of yellowfin tuna in the EPO between 1975-2016
- Apply VAST to the length composition of yellowfin tuna in the EPO between 2000-2016

1. CPUE: Data

- 1°*1° catch (in biomass) and effort (in days) data from the vessels with >75% dolphin-associated sets
- Nominal CPUE = catch / effort

Distribution of YFT fisheries between 1975-2016



1. CPUE: Spatiotemporal Model

VAST separately models encounter probability (p) and positive catch rate (λ) for sample i:

1. $p_i = \mathrm{logit}^{-1}\big(\beta_1(t_i) + \omega_1(s_i) + \varepsilon_1(s_i, t_i)\big)$
2. $\lambda_i = \exp\big(\beta_2(t_i) + \omega_2(s_i) + \varepsilon_2(s_i, t_i)\big)$

$\beta_1(t_i)$ and $\beta_2(t_i)$: intercept in year t_i

$\omega_1(s_i)$ and $\omega_2(s_i)$: spatial variation at location s_i

$\varepsilon_1(s_i, t_i)$ and $\varepsilon_2(s_i, t_i)$: spatiotemporal variation at location s_i in year t_i

1. CPUE: Spatiotemporal Model

Autocorrelated spatial and spatiotemporal residuals:

$\omega_1 \sim \mathrm{MVN}(\mathbf{0}, \sigma_{\omega 1}^2 \mathbf{R}_1)$

$\omega_2 \sim \mathrm{MVN}(\mathbf{0}, \sigma_{\omega 2}^2 \mathbf{R}_2)$

$\varepsilon_1(,t) \sim \mathrm{MVN}(\mathbf{0}, \sigma_{\varepsilon 1}^2 \mathbf{R}_1)$

$\varepsilon_2(,t) \sim \mathrm{MVN}(\mathbf{0}, \sigma_{\varepsilon 2}^2 \mathbf{R}_2)$

where

$$\mathrm{R}_1(s,s') = \frac{1}{2^{\nu-1}\Gamma(n)} \times (\kappa_1|\mathbf{H}(s-s')|)^{\nu} \times K_{\nu}(\kappa_1|\mathbf{H}(s-s')|)$$

$$\mathrm{R}_2(s,s') = \frac{1}{2^{\nu-1}\Gamma(n)} \times (\kappa_2|\mathbf{H}(s-s')|)^{\nu} \times K_{\nu}(\kappa_2|\mathbf{H}(s-s')|)$$

decorrelation distance

geometric anisotropy

Matern smoothness (=1)

1. CPUE: Spatiotemporal Model

The probability of catch data c for sample i:

$$\Pr(c_i = C) = \begin{cases} 1 - p_i & \text{if } C = 0 \\ p_i \times lognormal(c_i | \lambda_i, \sigma_m^2) & \text{if } C > 0 \end{cases}$$

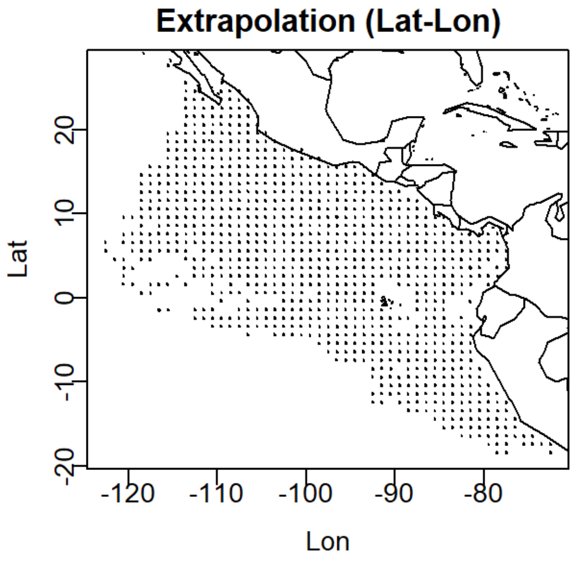
Joint Log-Likelihood = $\sum_i \log(\Pr(c_i = C))$

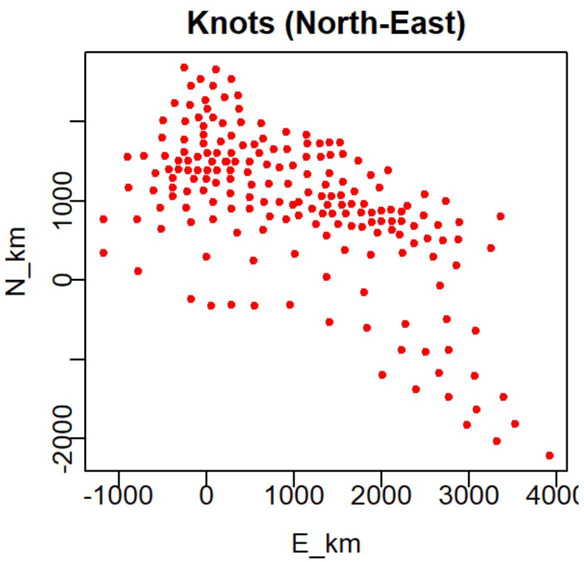
TMB + R

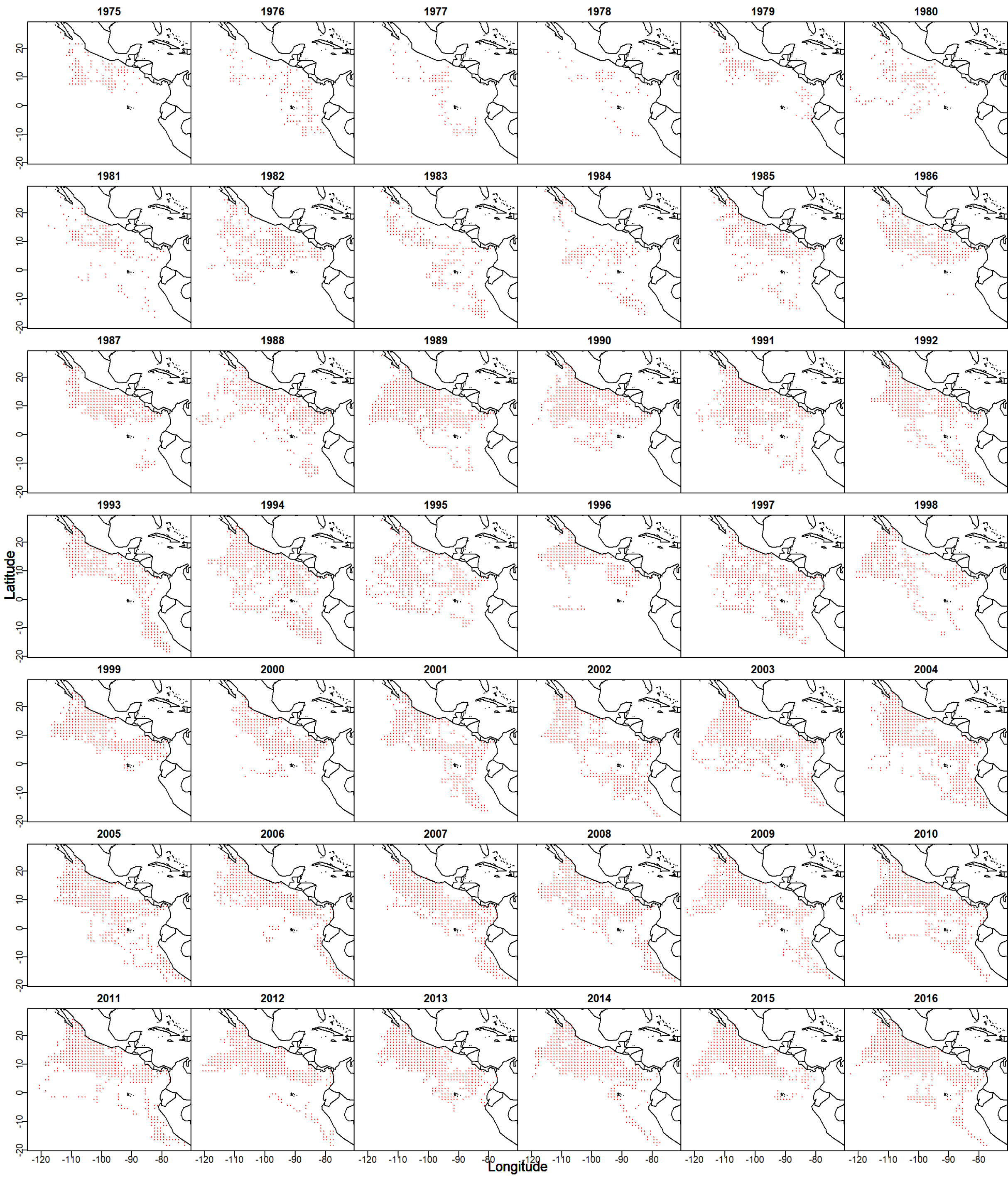
The spatiotemporal dynamics of yellowfin tuna in quarter 1

Large inter-annual variation in fishery locations

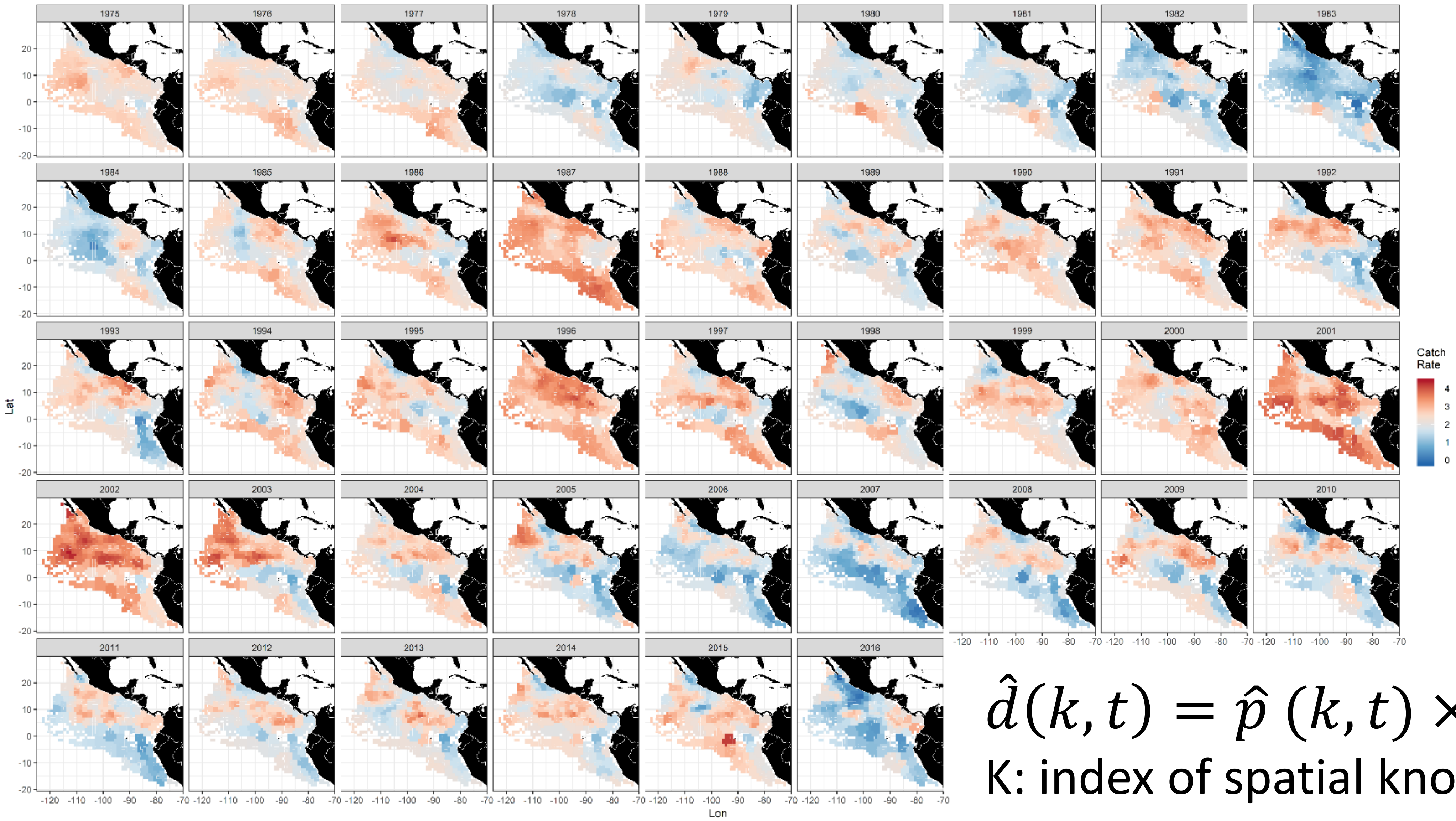
use knots to approximate spatial and spatiotemporal residuals







Predicted log catch rate for quarter 1



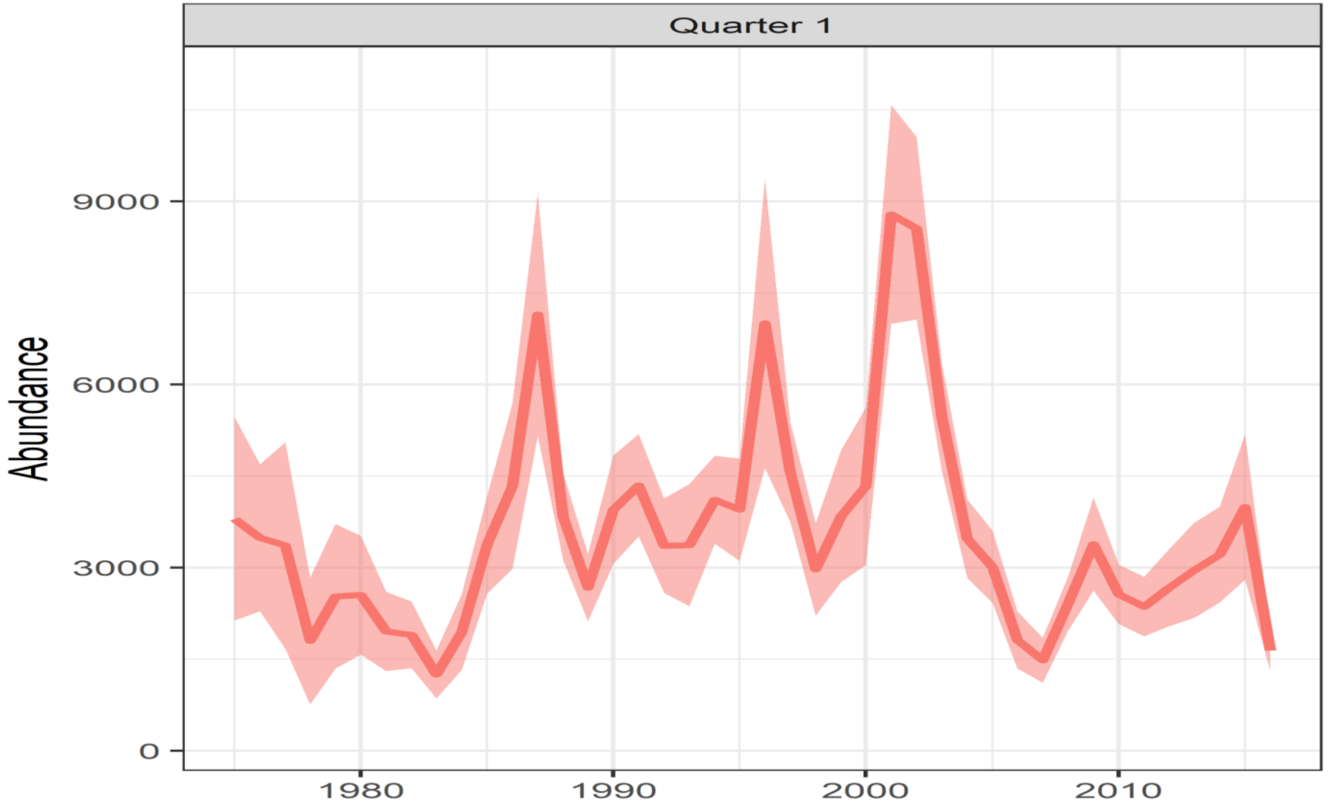
$$\hat{d}(k,t) = \hat{p}\,(k,t) \times \hat{\lambda}(k,t)$$

K: index of spatial knot

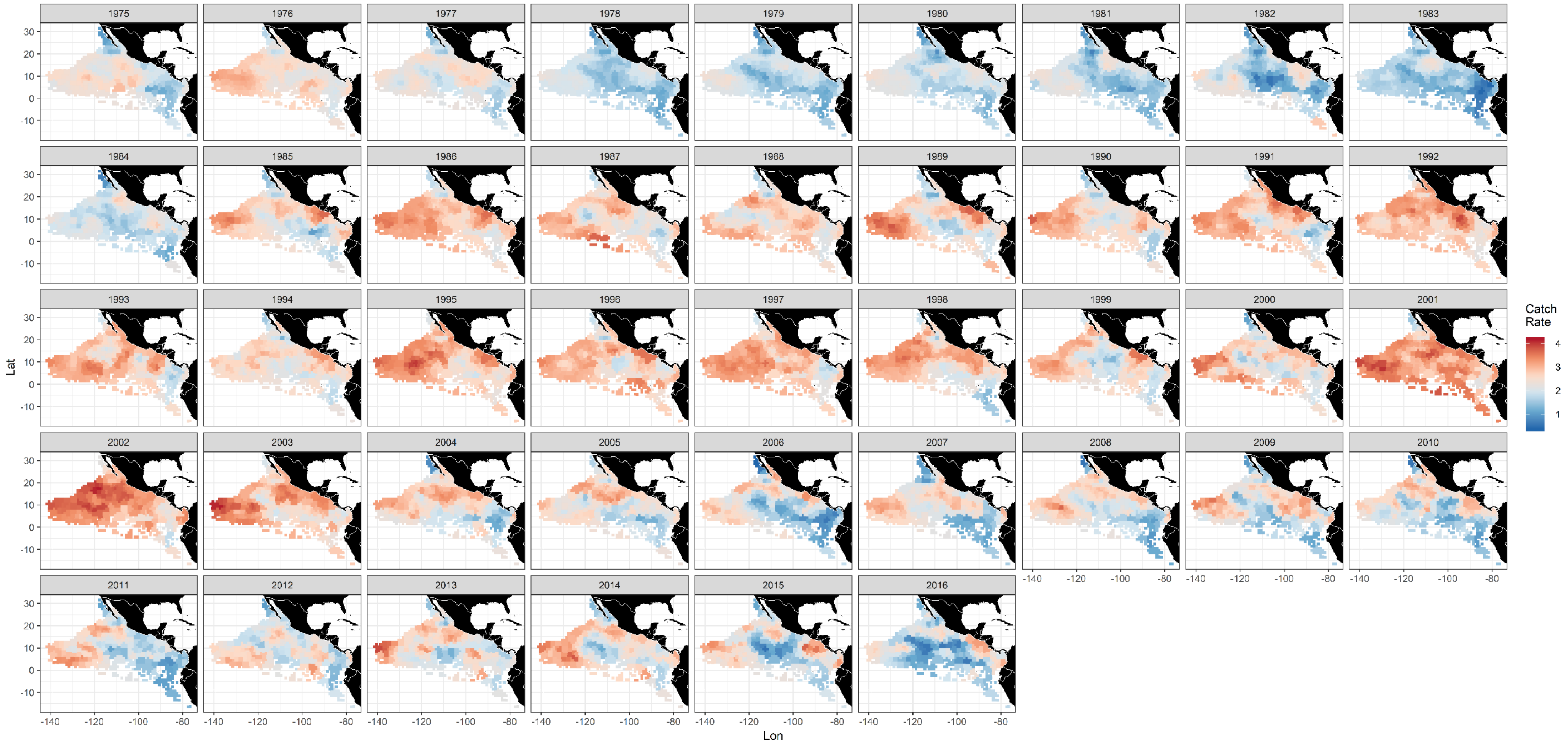
1. CPUE: Spatiotemporal Model

Total abundance for the entire domain

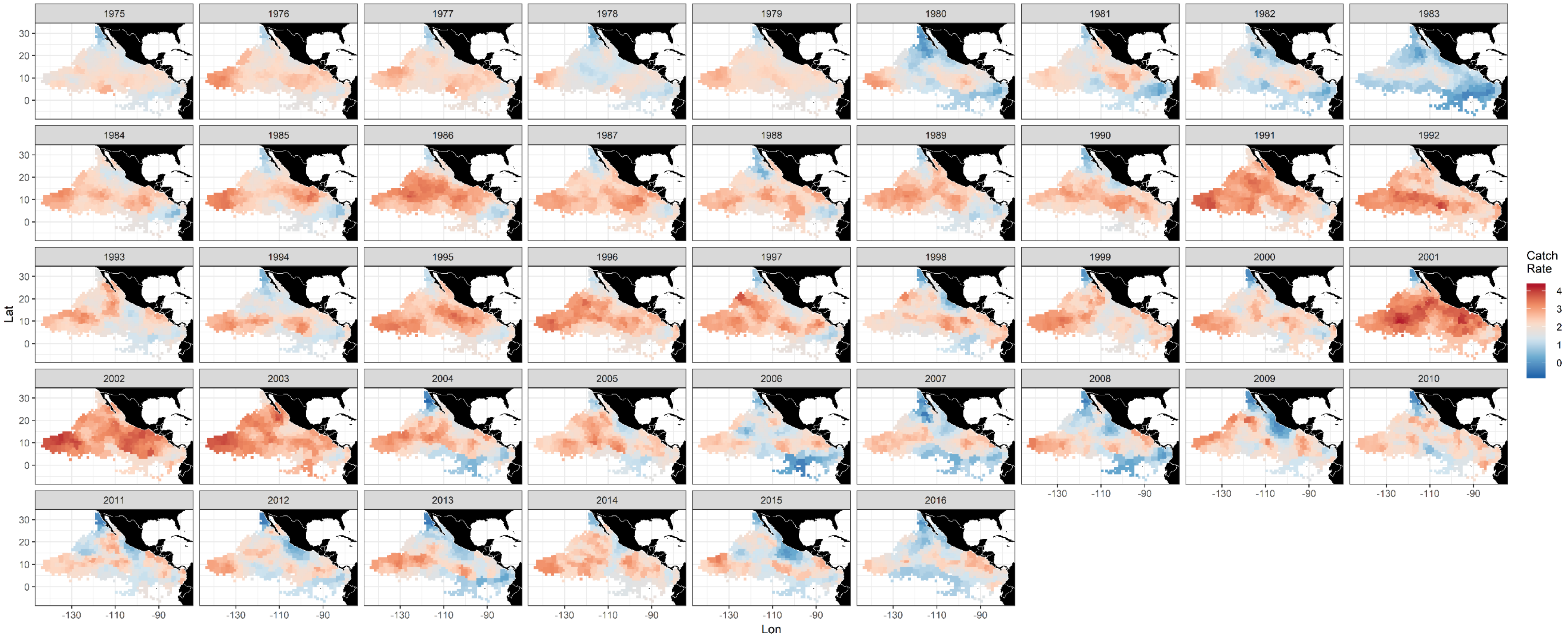
$I(t) = \sum_{k=1}^{n_k} \left(a(k) \times \hat{d}(k,t) \right)$ -- k: spatial knot



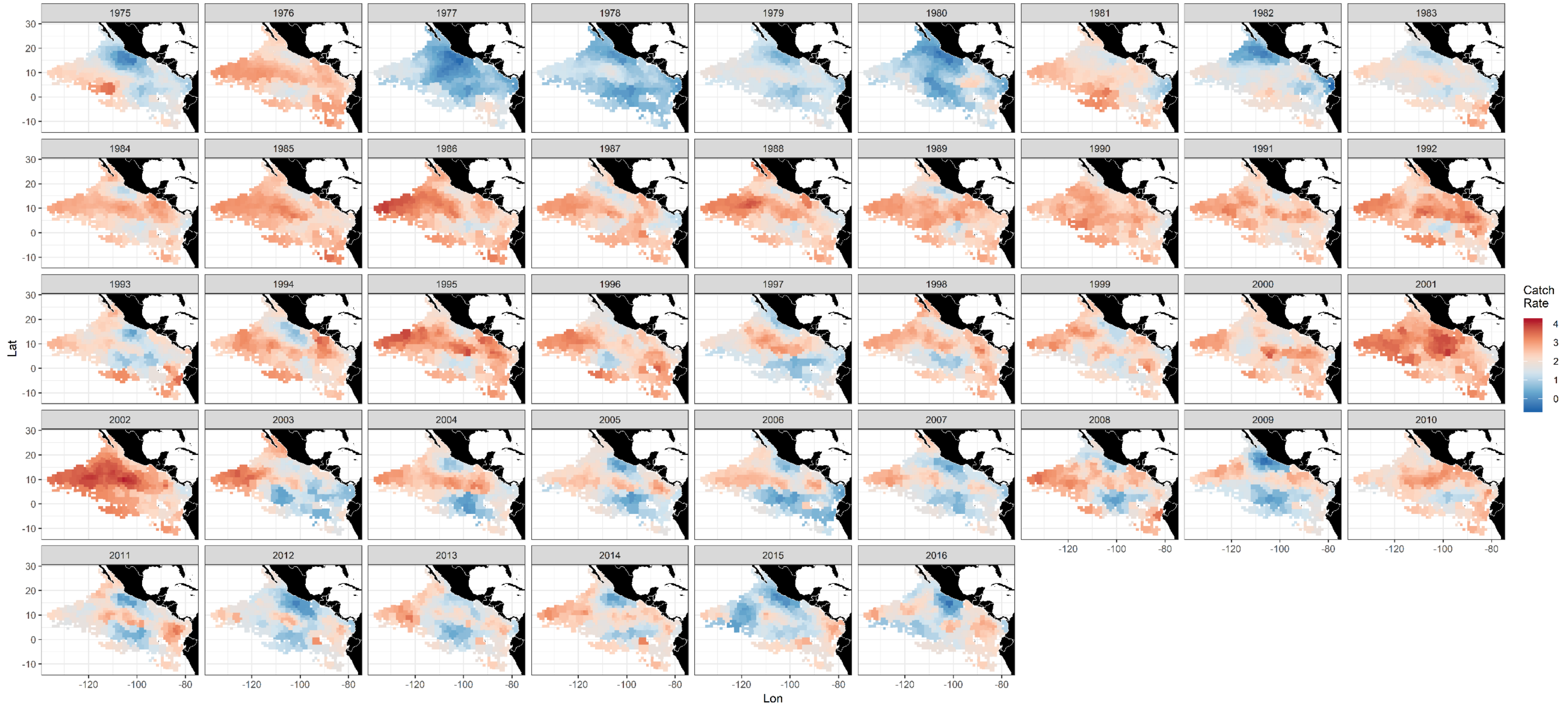
Predicted log catch rate for quarter 2



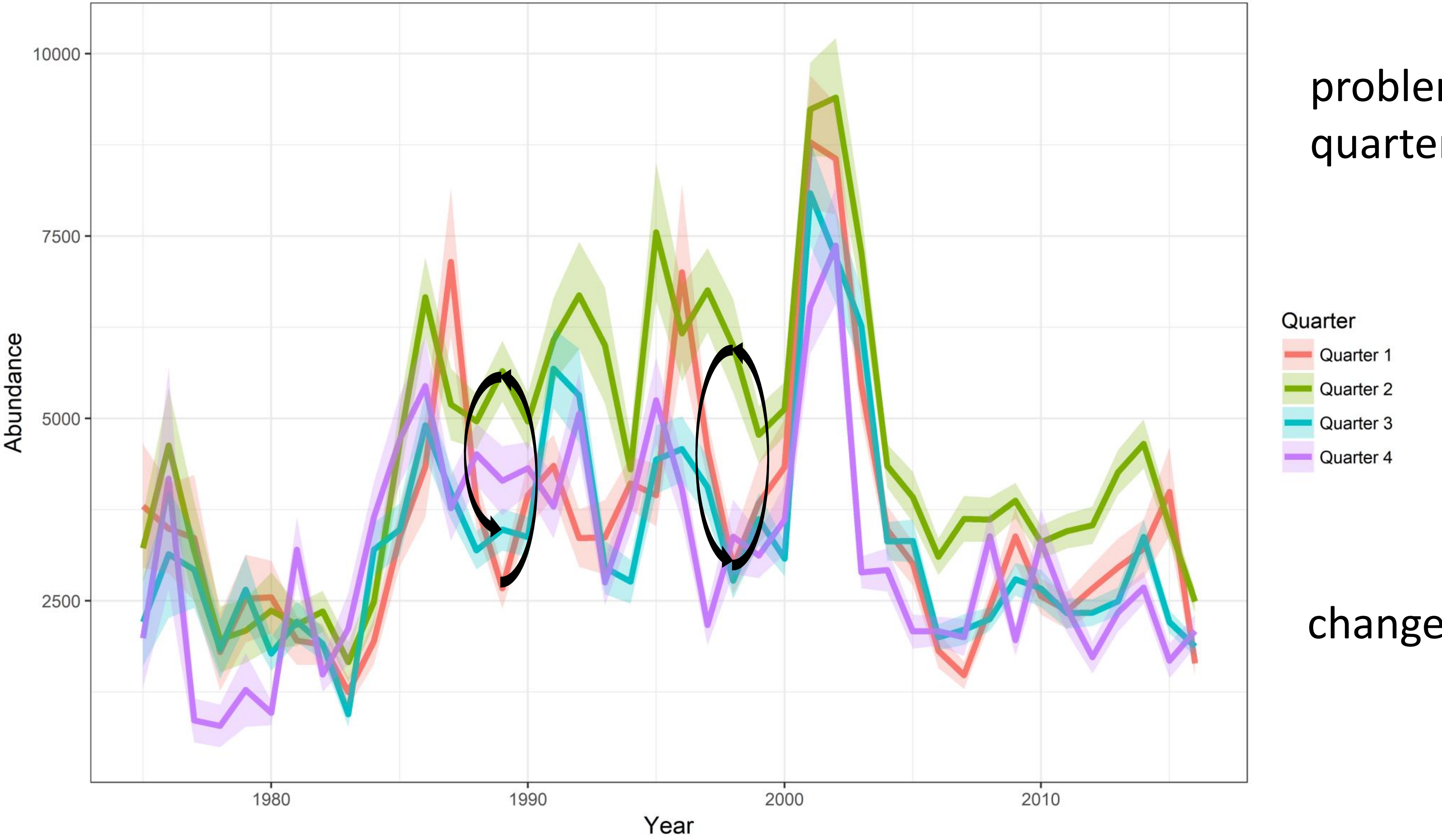
Predicted log catch rate for quarter 3



Predicted log catch rate for quarter 4

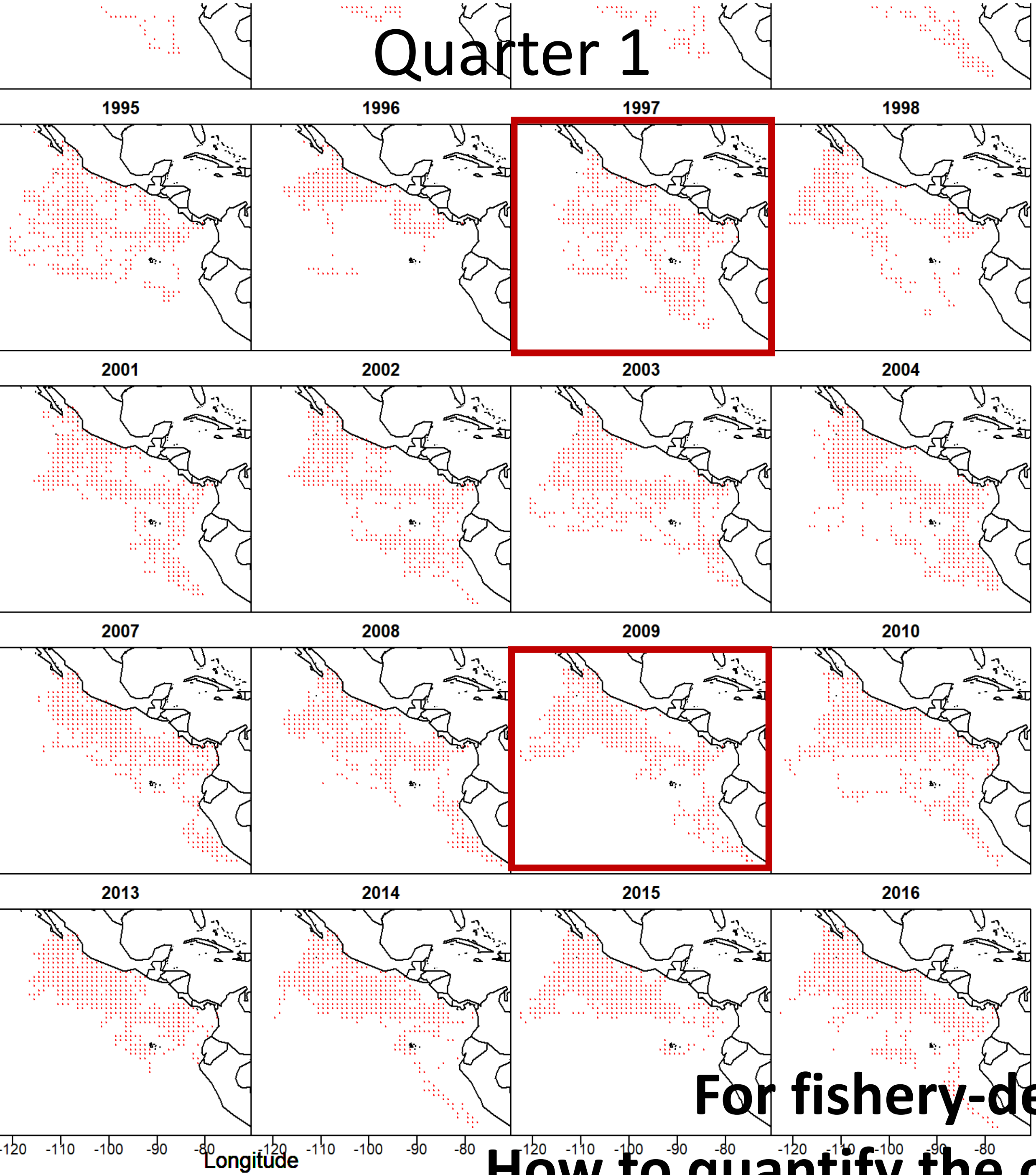


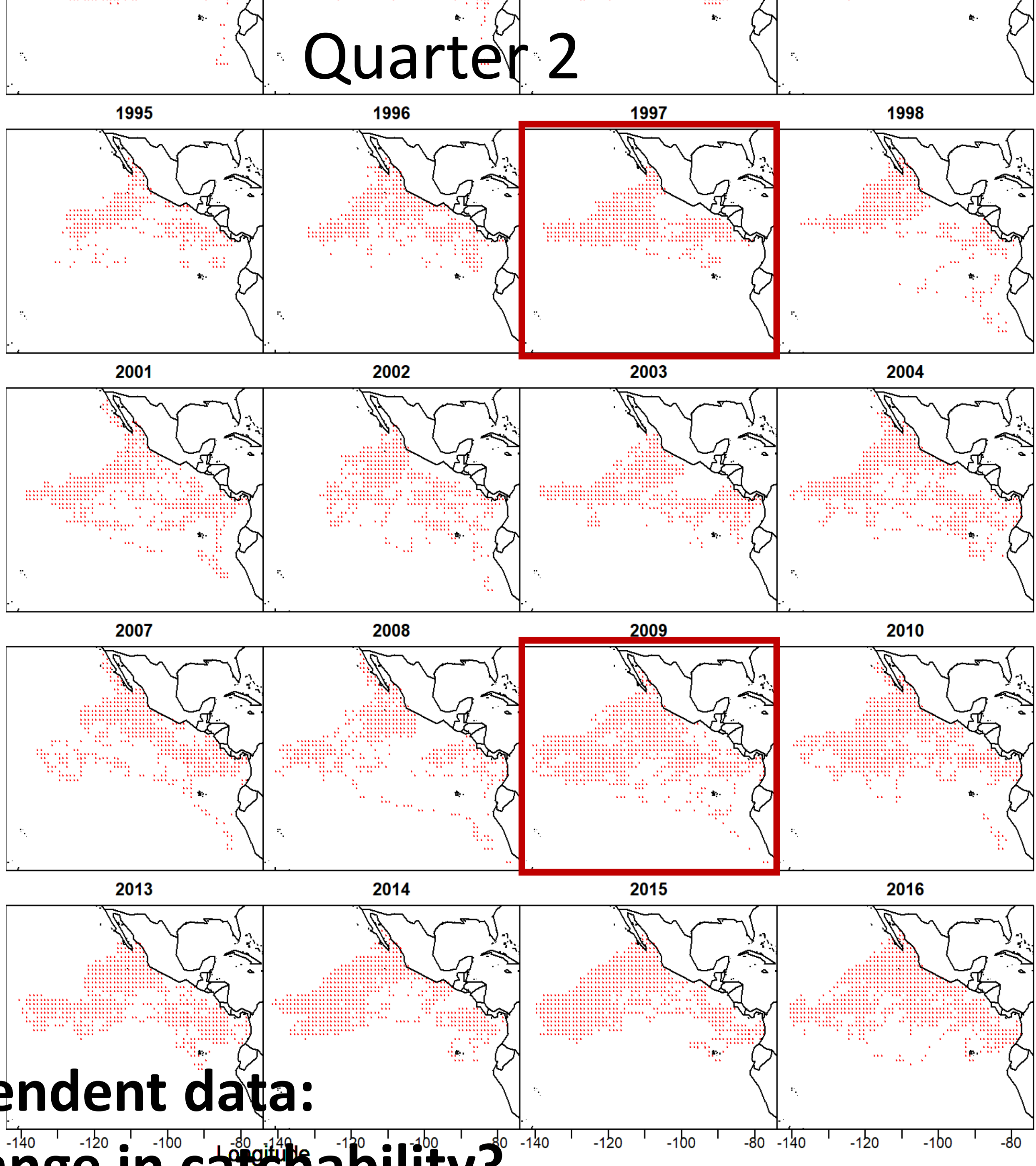
Comparison the 4 quarterly indices of abundance



problematic for the quarter as year approach

change in catchability?

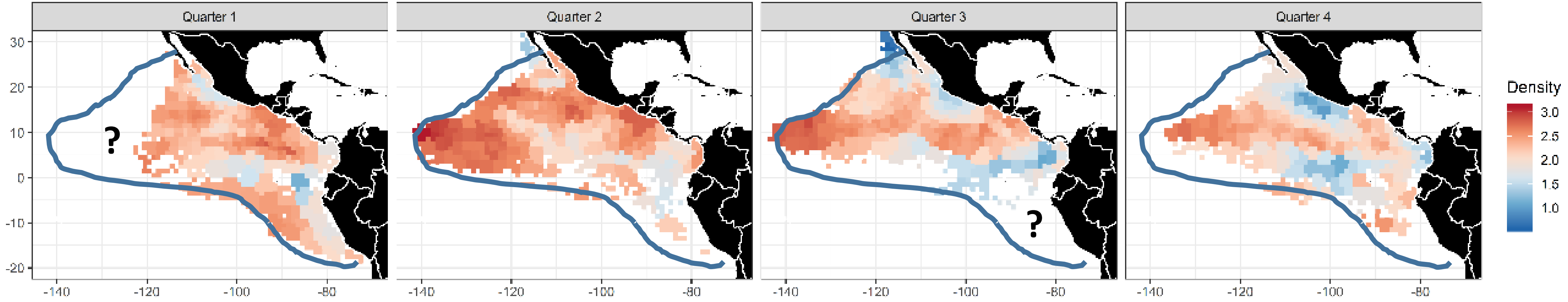




For fishery-dependent data:

How to quantify the change in catchability?

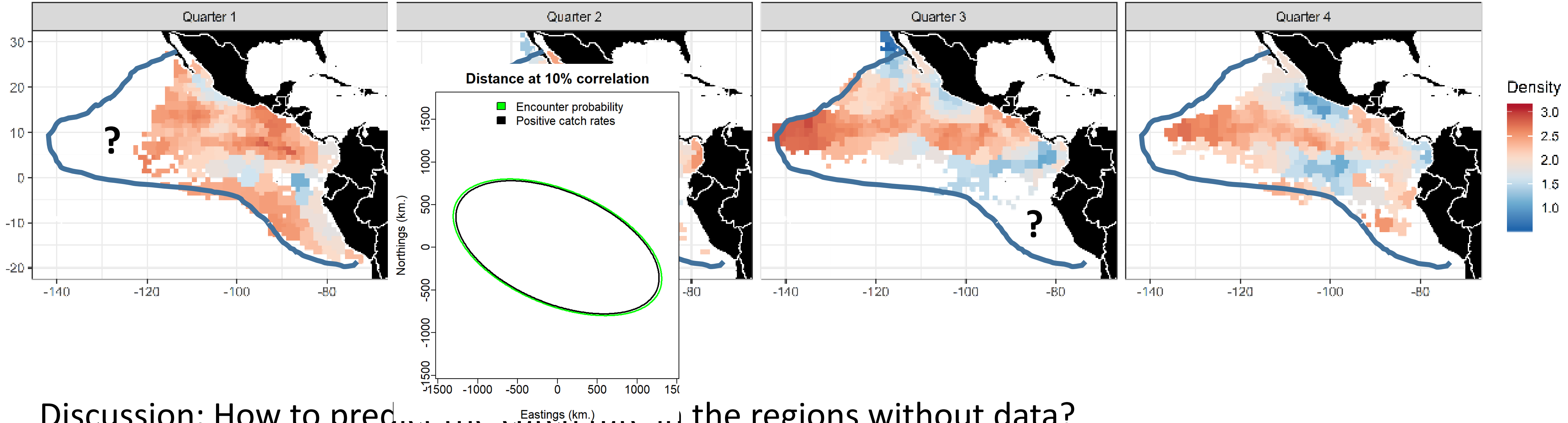
Long-term mean log catch rate



The locations of fisheries activities change from quarter to quarter, but the quarterly indices of abundance should be calculated based on the same domain in the EPO.

Discussion: How to predict the catch rate in the regions without data?

Long-term mean log catch rate

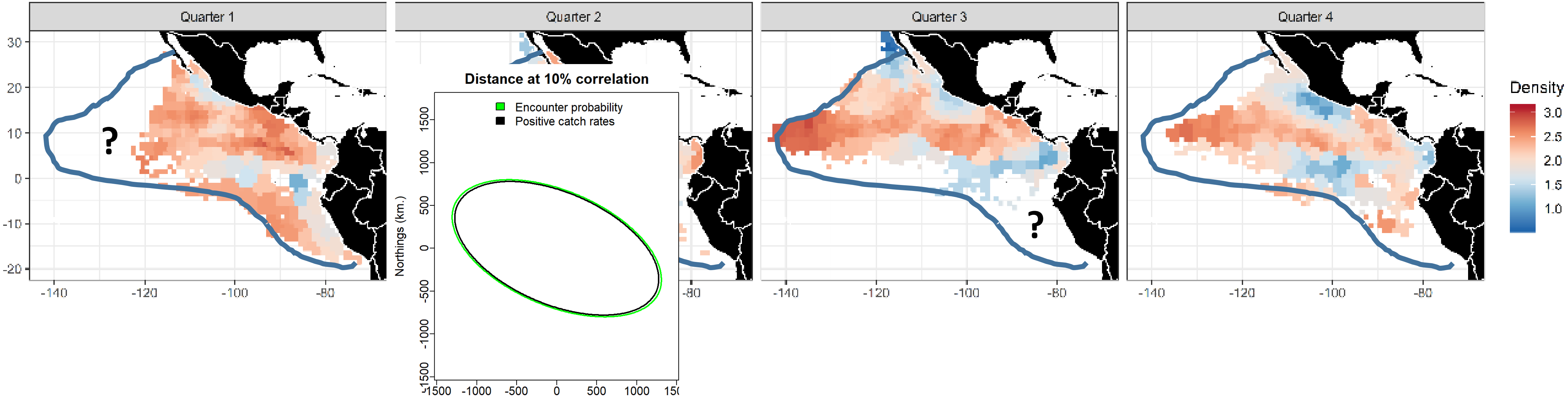


Discussion: How to predict the catch rate in the regions without data?

Solution 1: catch rate = 0

Problem: Predicted catch rate is high in some adjacent regions with data

Long-term mean log catch rate



Discussion: How to predict the catch rate in the regions without data?

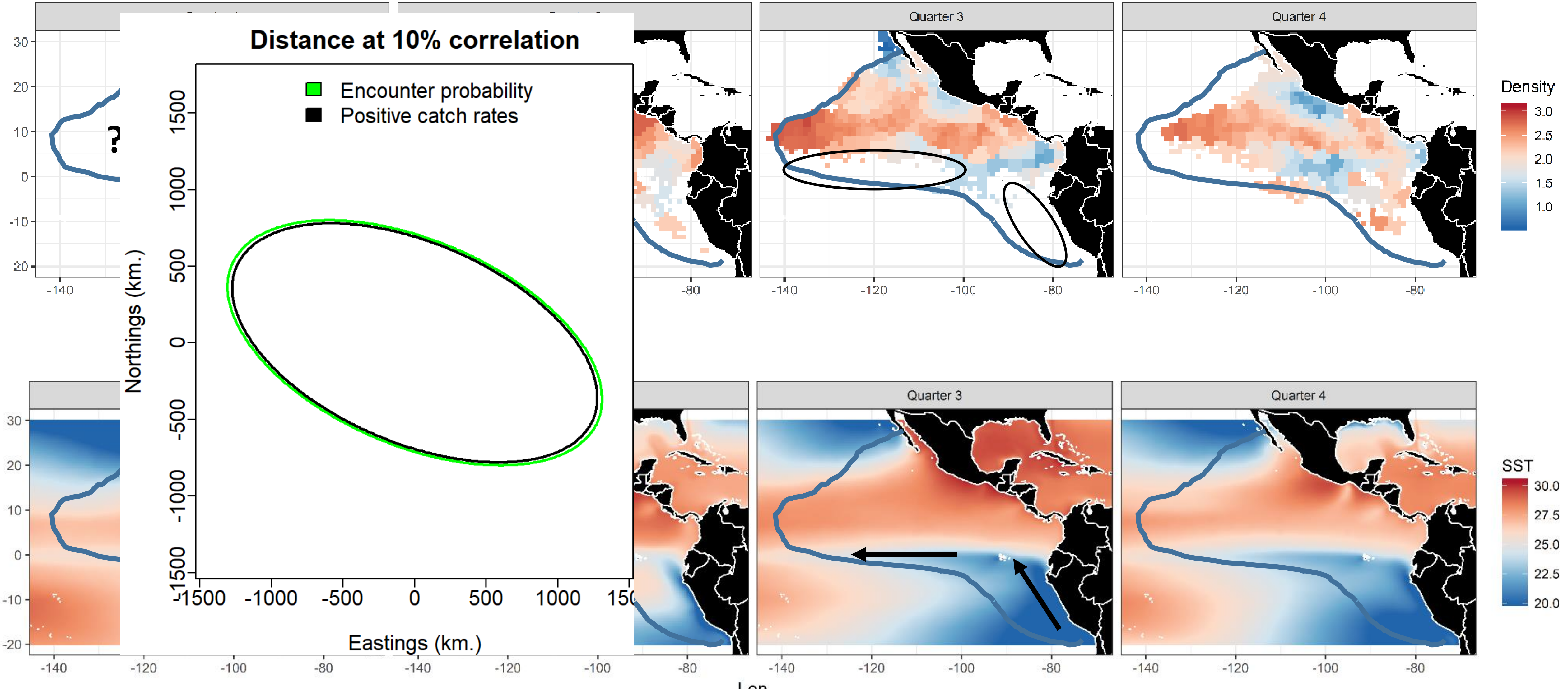
Solution 1: catch rate = 0

Problem: Predicted catch rate is high in adjacent regions with data

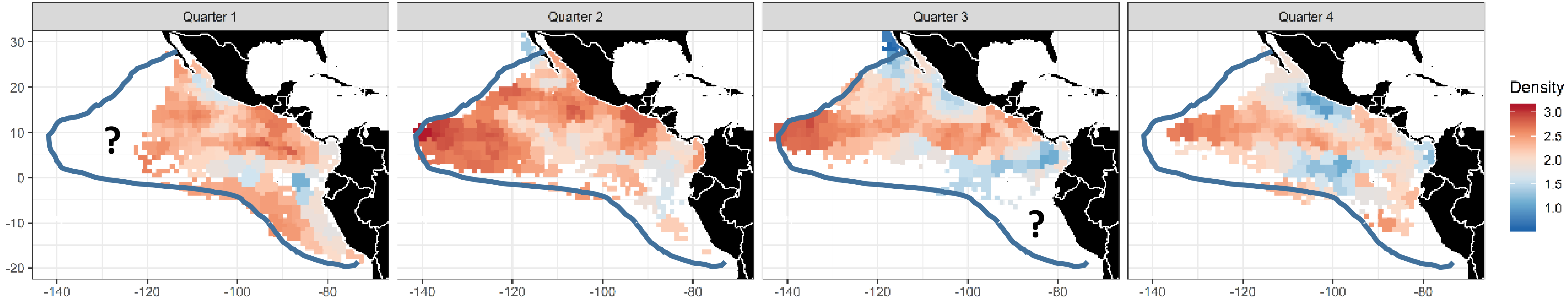
Solution 2: interpolate catch rate based on spatial autocorrelation in residuals

Problem: spatial correlation in residuals may vary over space

The fishery follows the movement of warm waters



Long-term mean log catch rate



Discussion: How to predict the catch rate in the regions without data?

Solution 1: catch rate = 0

Problem: Predicted catch rate is high in adjacent regions with data

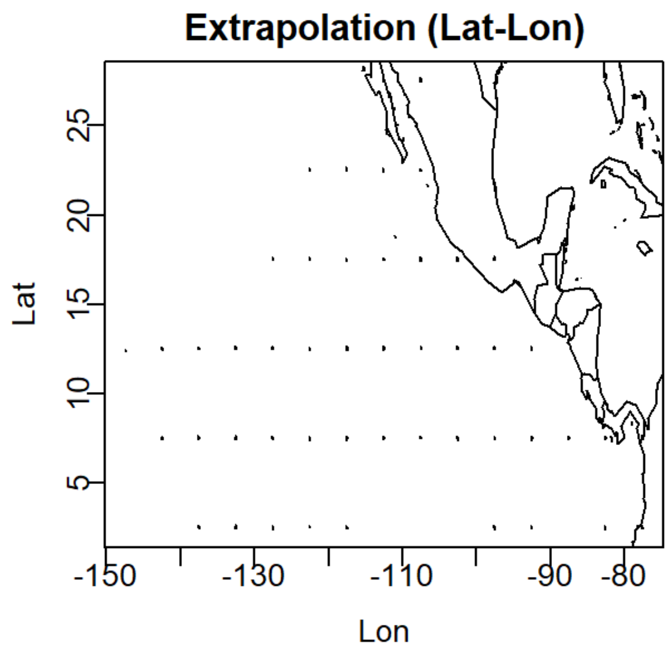
Solution 2: interpolate catch rate based on spatial autocorrelation in residuals

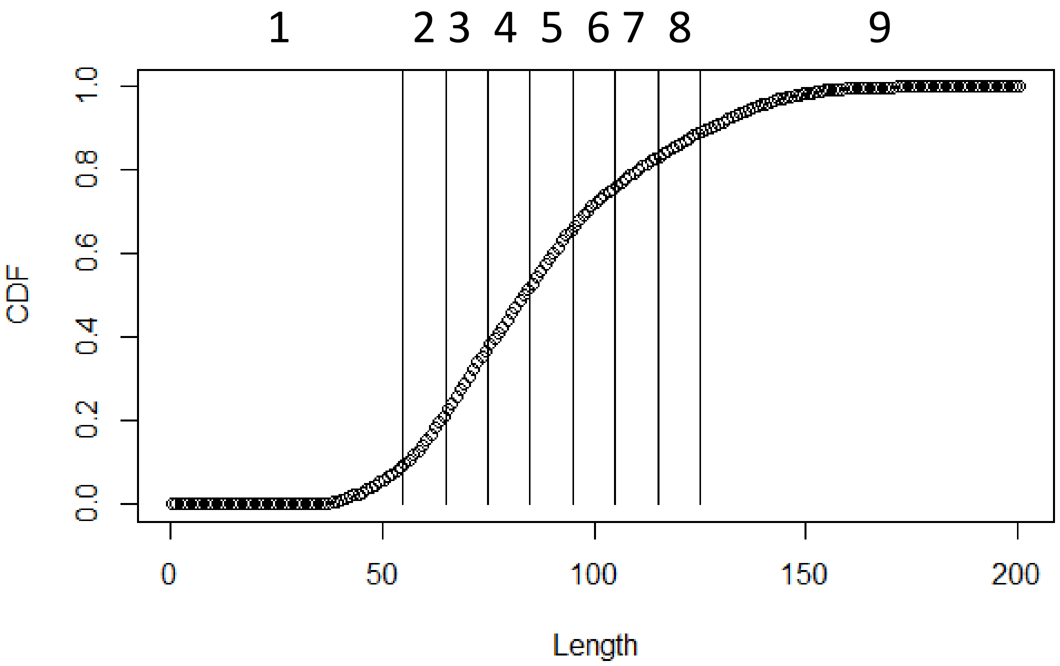
Problem: spatial correlation in residuals may vary over space

Any suggestions?

2. Length Comps: Data

- 5°*5° number at length (1cm bin) and effort (in days) data from the vessels with >75% dolphin-associated sets
- Case study: Northern hemisphere; Quarter 2





2. Length Comp: Spatiotemporal Model

$$p_i = \text{logit}^{-1}\left(\beta_1(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_1}(l_i, f)\omega_1(s_i, f) + \sum_{f=1}^{n_f} L_{\varepsilon_1}(l_i, f)\varepsilon_1(s_i, f, t_i)\right)$$

$$\lambda_i = \exp\left(\beta_2(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_2}(l_i, f)\omega_2(s_i, f) + \sum_{f=1}^{n_f} L_{\varepsilon_2}(l_i, f)\varepsilon_2(s_i, f, t_i)\right)$$

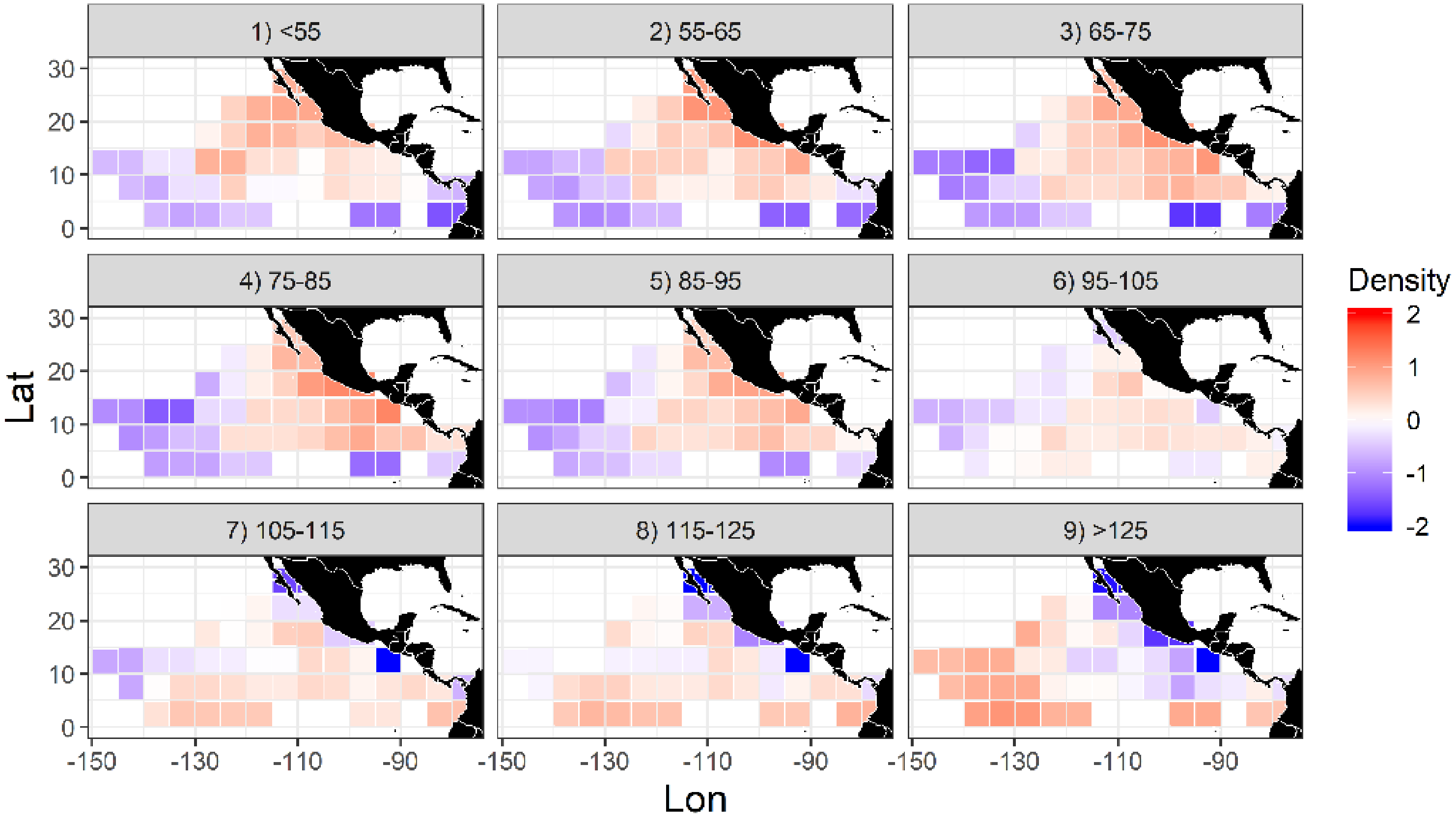
l_i: length bin (1-9);

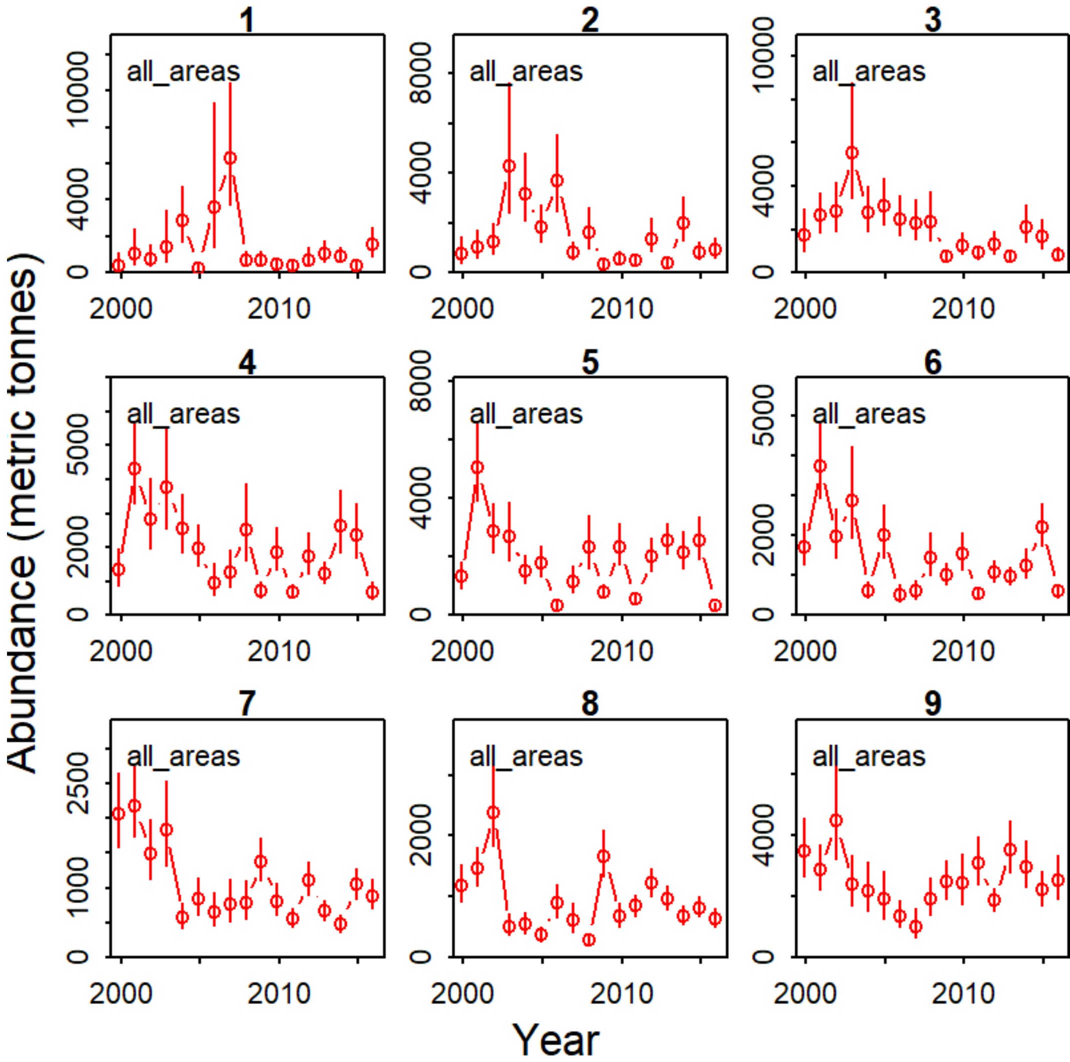
L_{ω_1}, L_{ε_1}, $L_{\omega 2}$, L_{ε_2}: generate the loading matrixes ($\boldsymbol{L}^T \boldsymbol{L}$) for the spatial and spatiotemporal covariance among l_i

n_f: 2

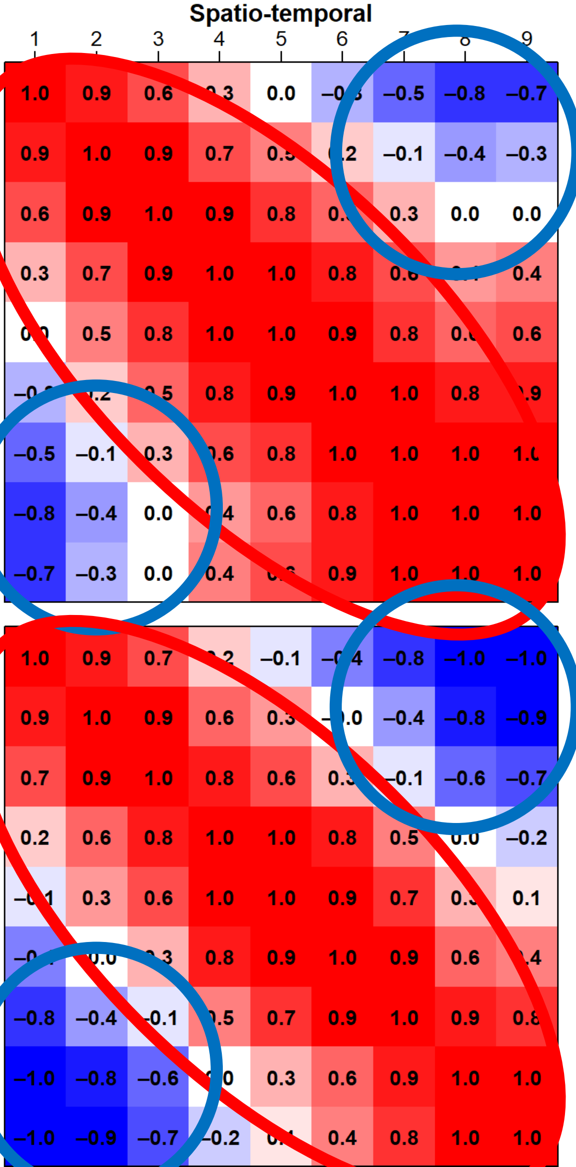
Spatial residuals

Mean predicted log catch rate between 2000-2016





Correlation matrix for spatiotemporal residuals



Encounter

Positive

Size segregation in spatiotemporal distribution too:

positive correlations between close length bins

negative correlations between distant length bins

AR1 in length does not work ***for YFT***

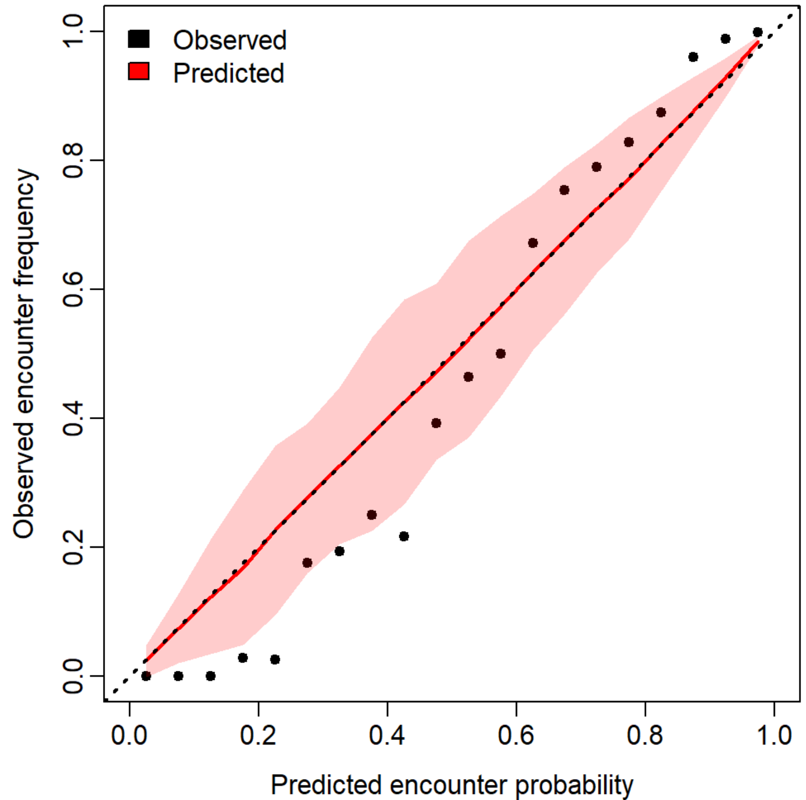
Hypothesis

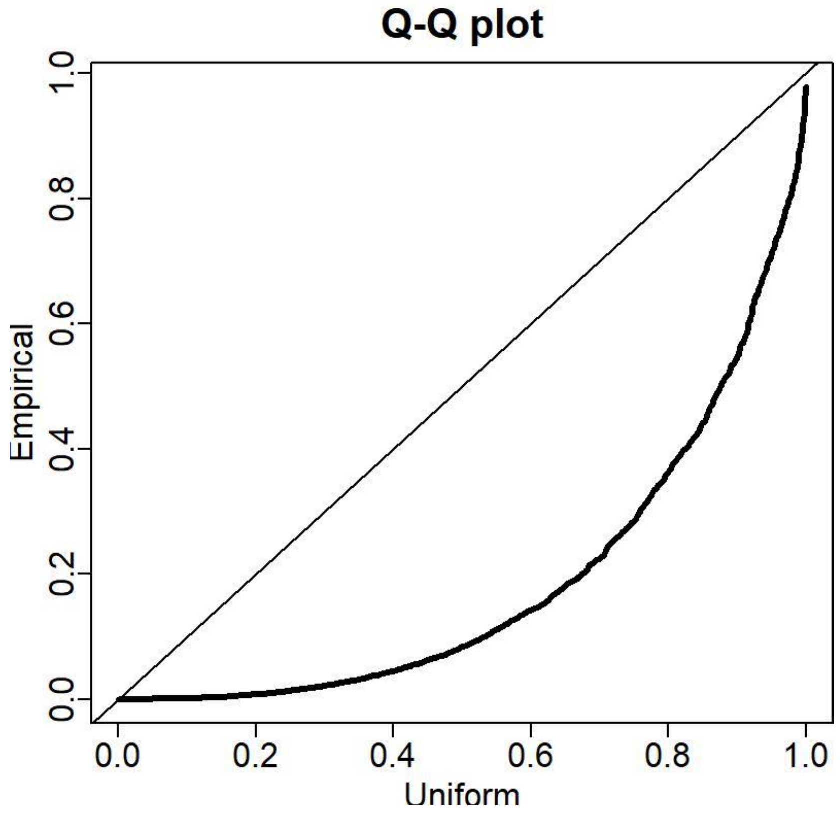
Differing habitat preferences of the smallest and largest YFT are driven by env conditions

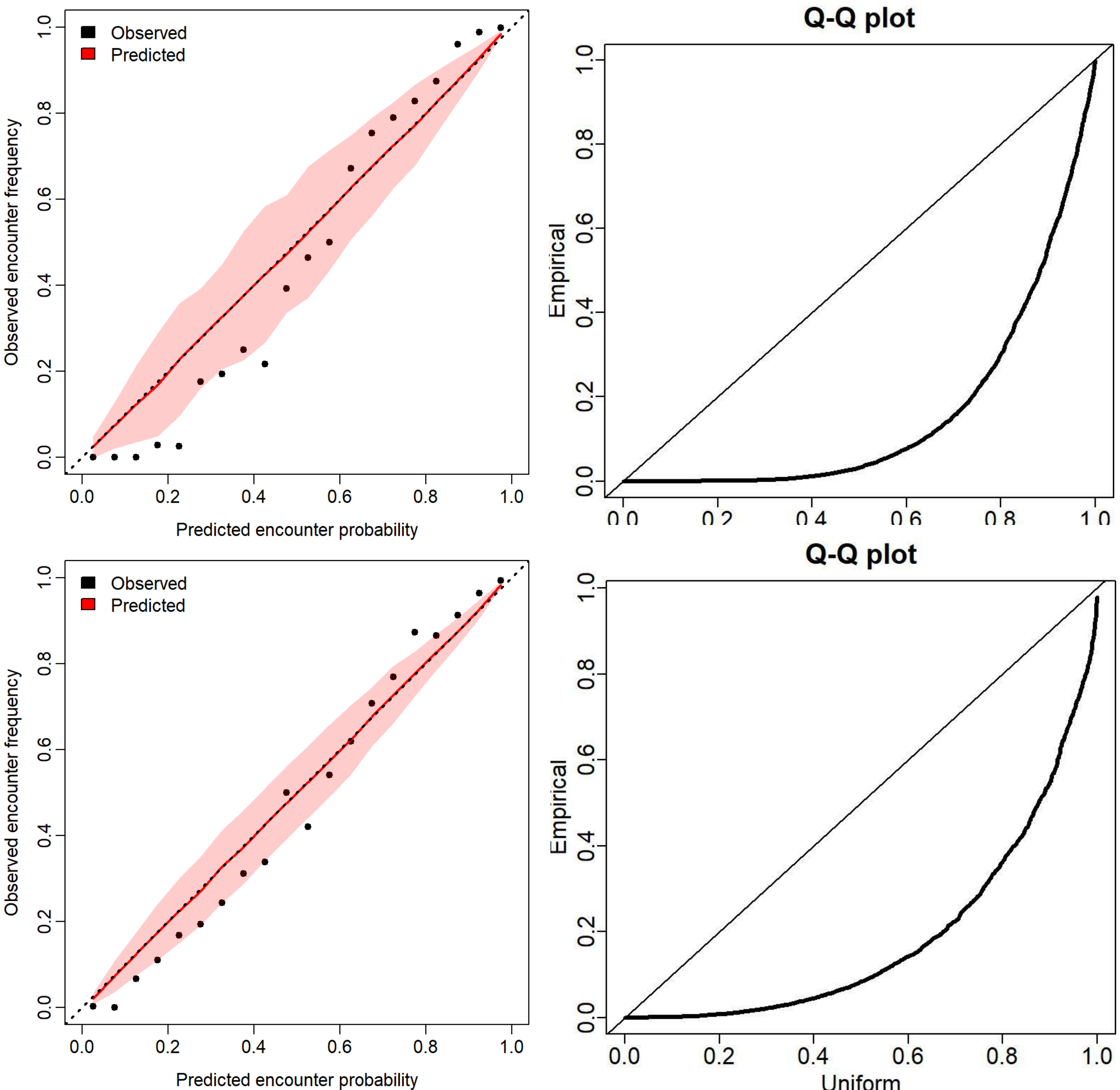
Problems in this length comp modelling

- Spatial resolution is low (5°*5°) -> 100% observed encounter rate occurs in some grid cells

My solution: RhoConfig = c("Beta1"=2, "Beta2"=0) -> random walk







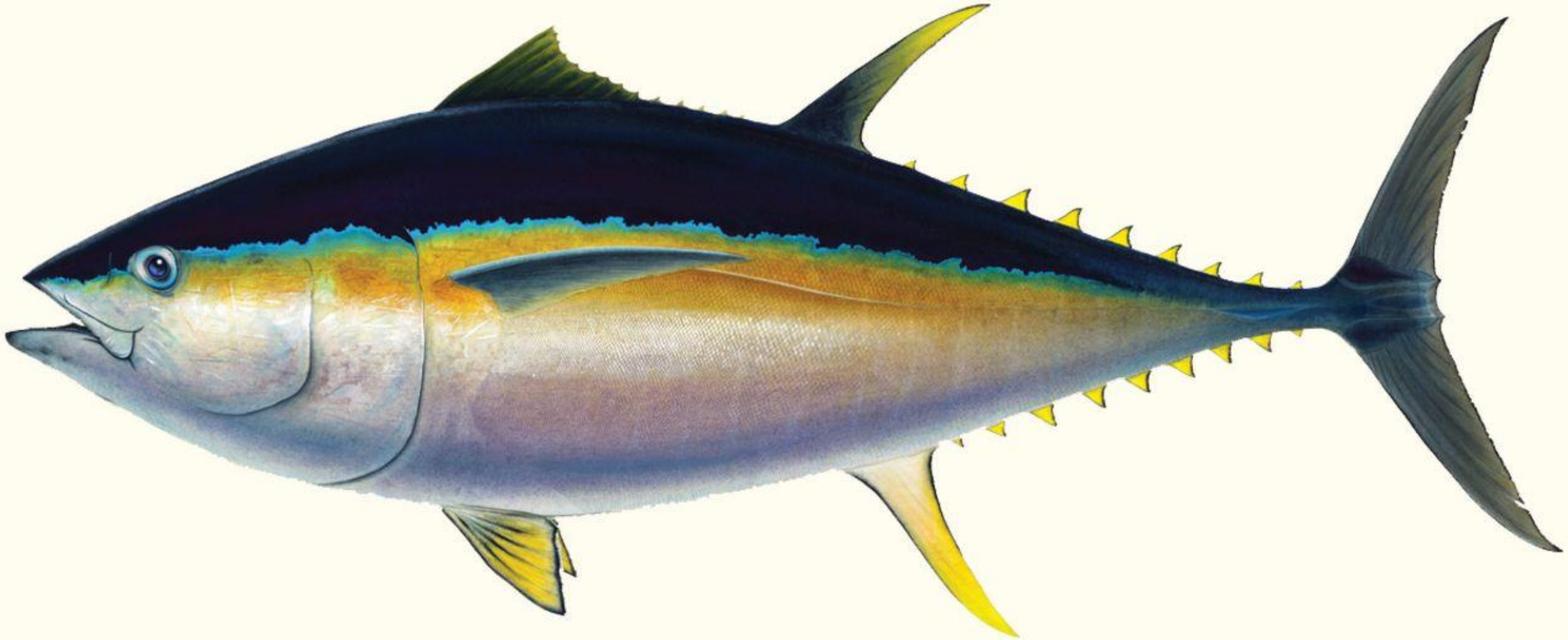
Increase the number of length bins from 9 to 20 leads to improved model fit

Discussion:

- Bad fit or bad diagnose?
- How many number of bins?
- How many number of factors?

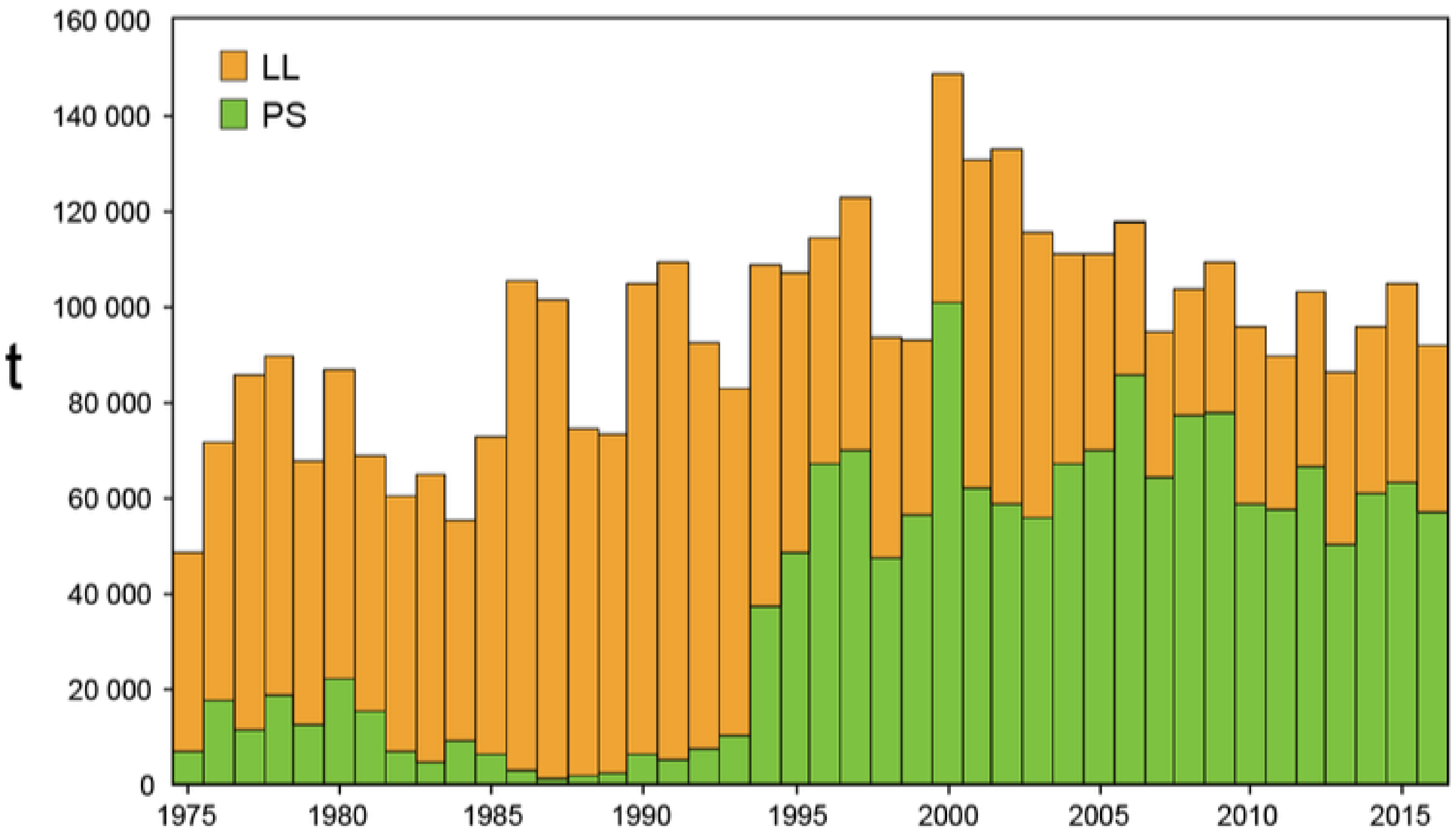
Bigeye tuna (*Thunnus obesus)*

- Assessed as one stock



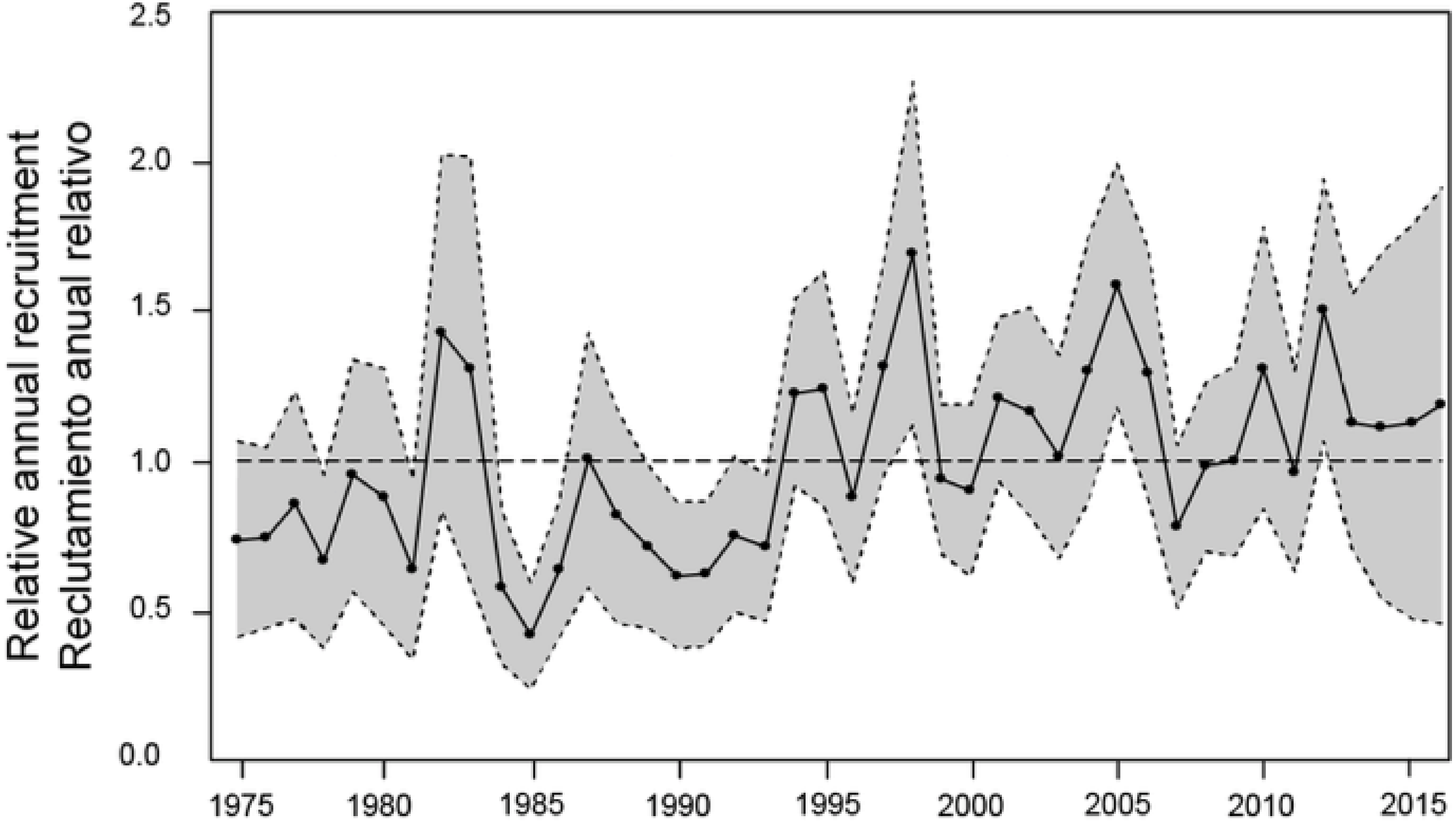
- Decline affected by La Niña events
- Strong recruitment in 2012

Catches



LL
PS
t
160 000
140 000
120 000
100 000
80 000
60 000
40 000
20 000
0
1975
1980
1985
1990
1995
2000
2005
2010
2015

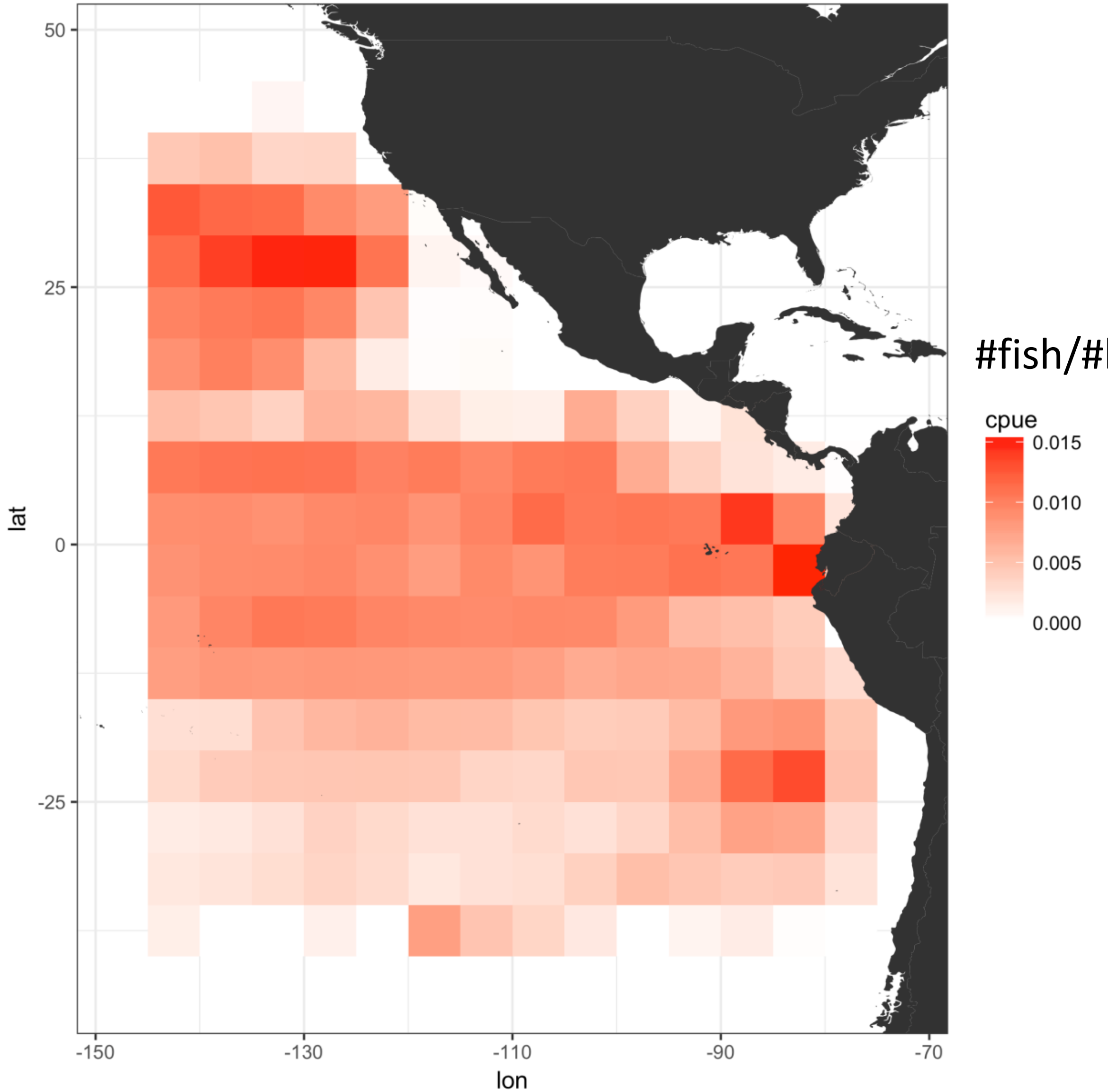
Recruitment



Data (fishery-dependent)

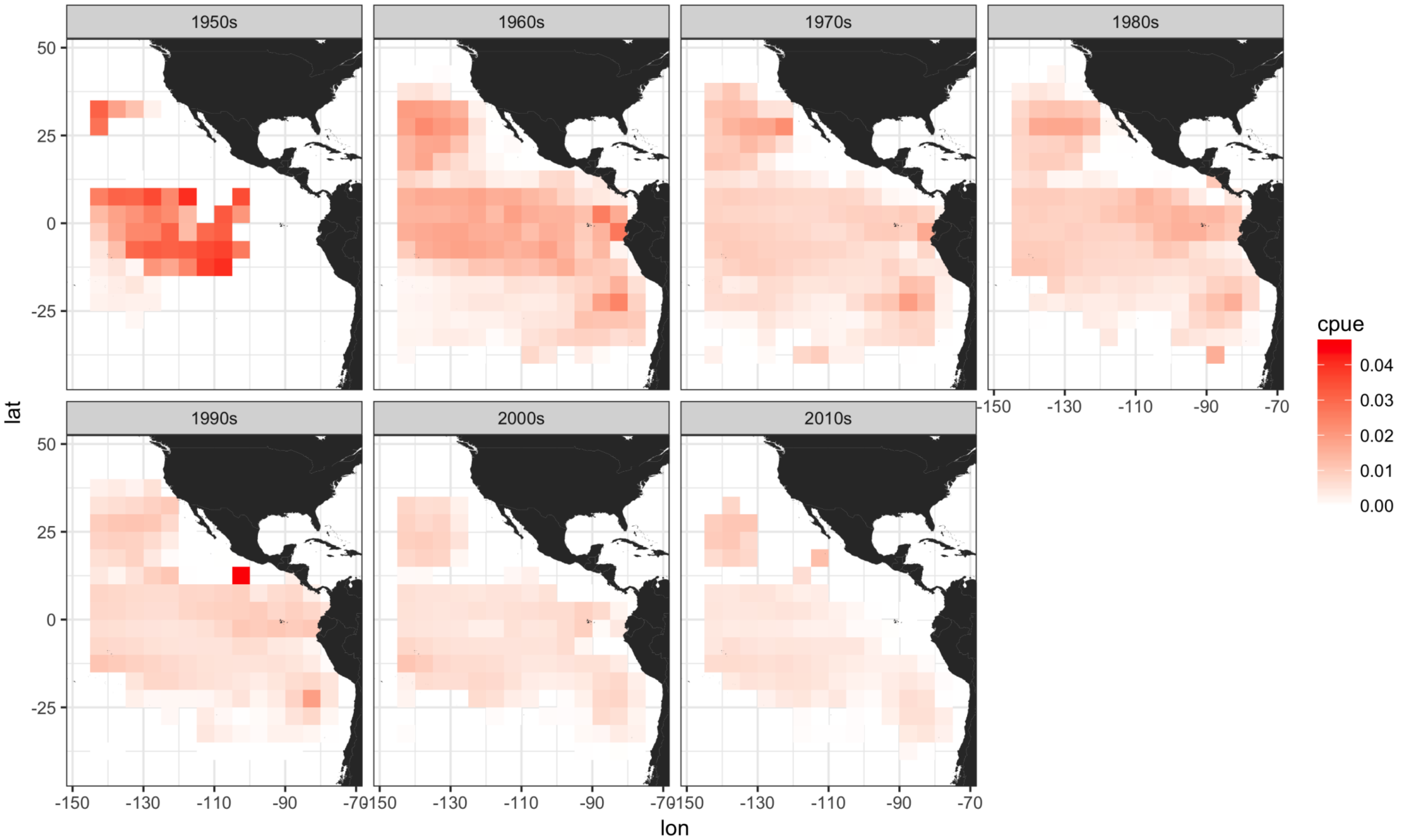
- CPUE
  - #fish/#hooks
  - 5° latitude × 5° longitude
- Length compositions
  - Counts (100% encounters)
  - 5° latitude × 10° longitude
- Japanese catches only
  - Longline
  - Annual

CPUE

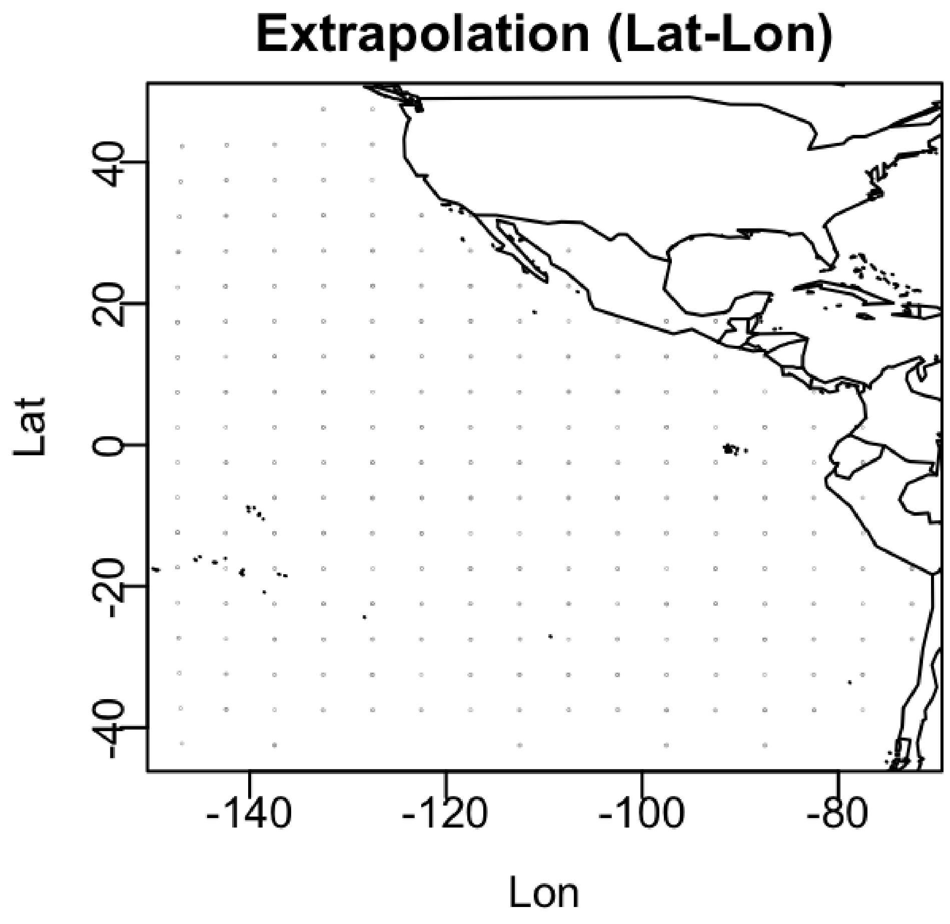


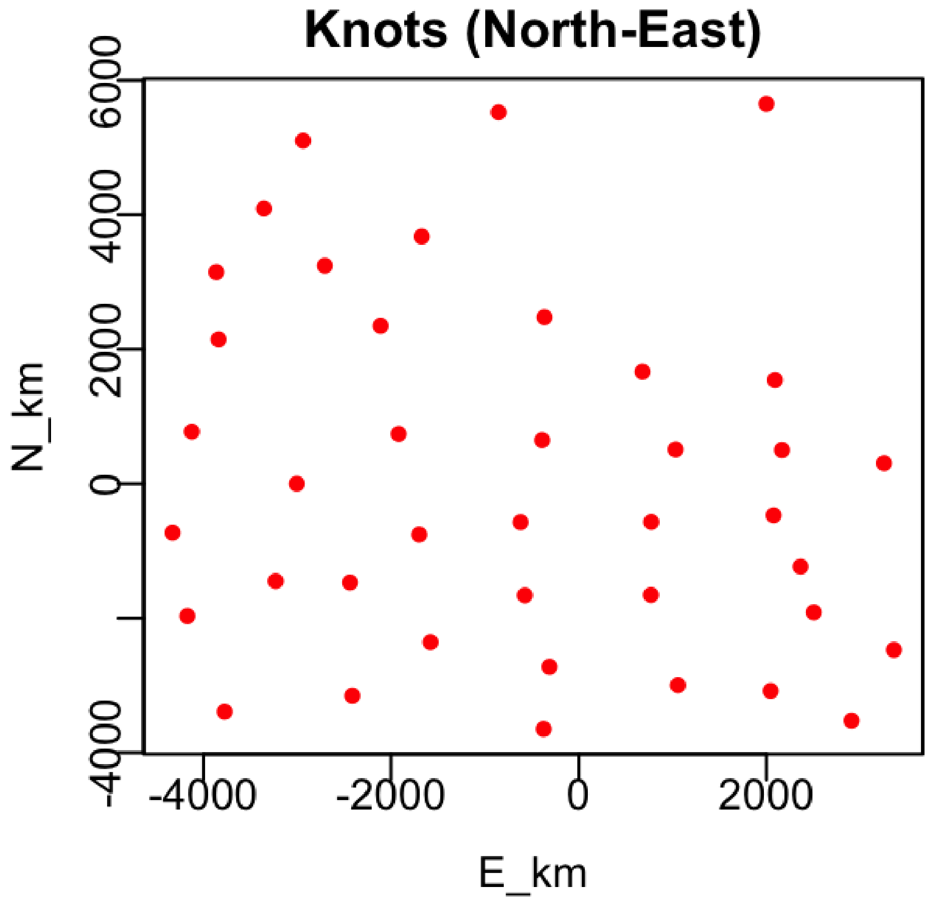
#fish/#hooks

CPUE by decade

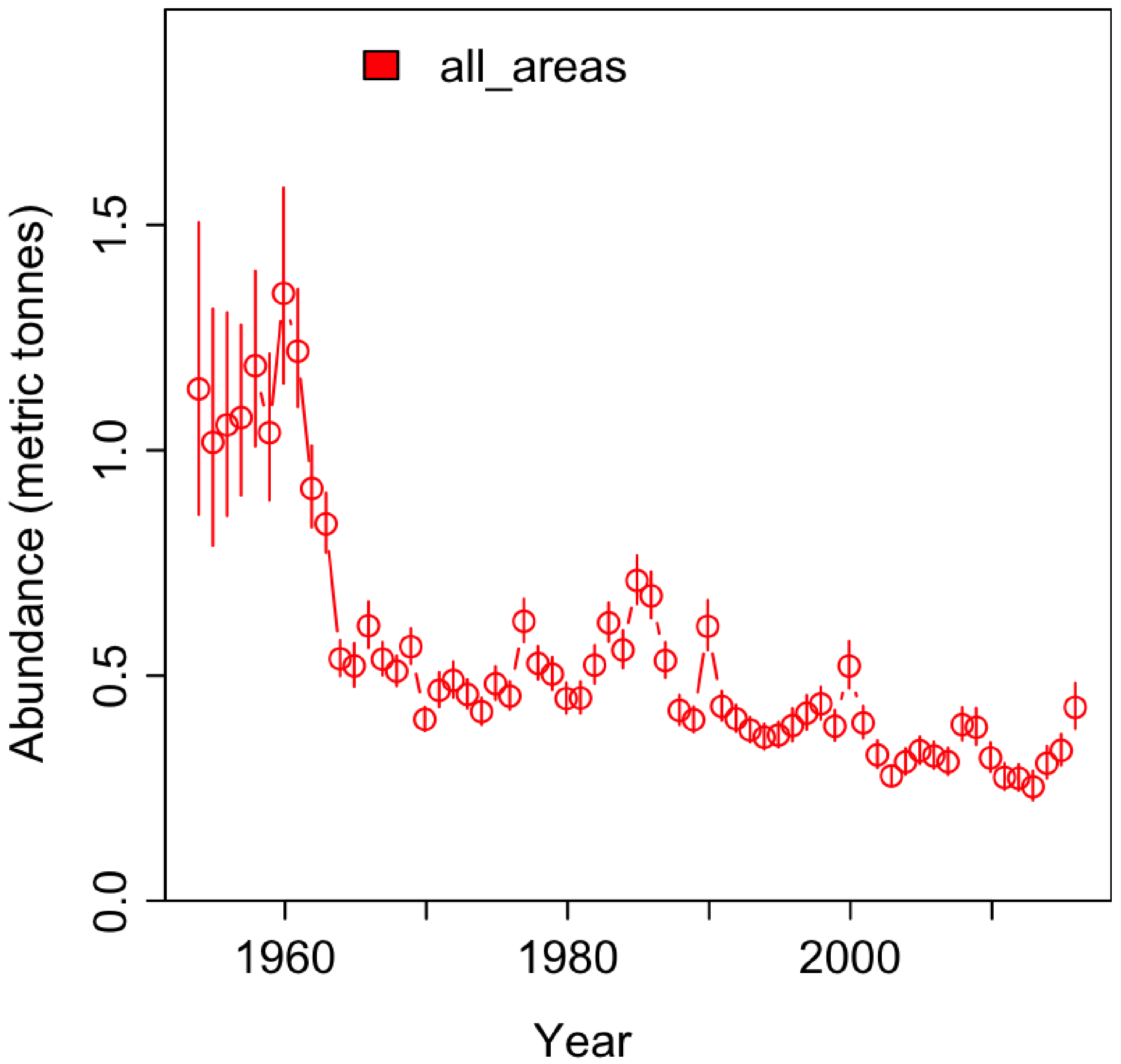


VAST model - CPUE

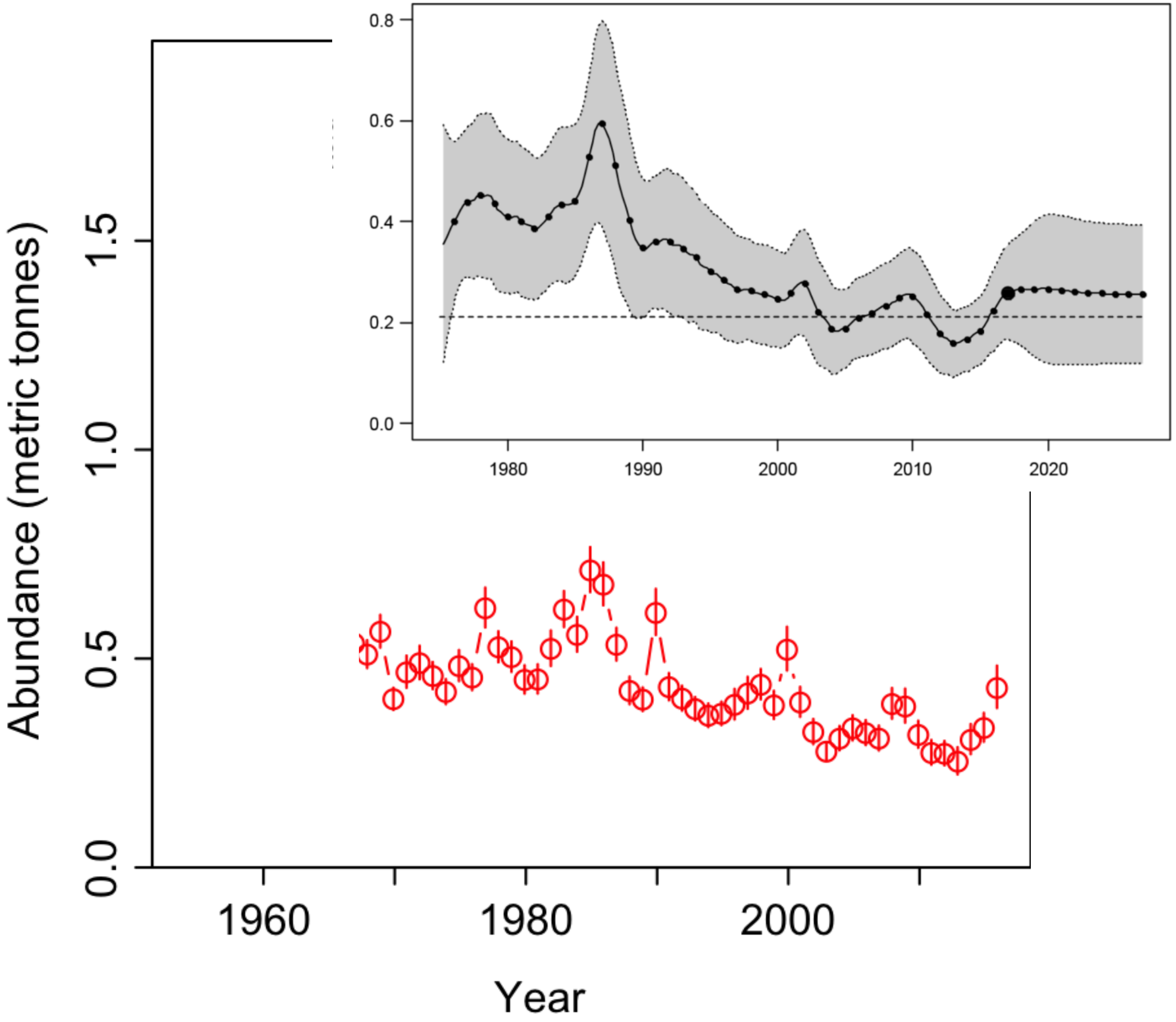




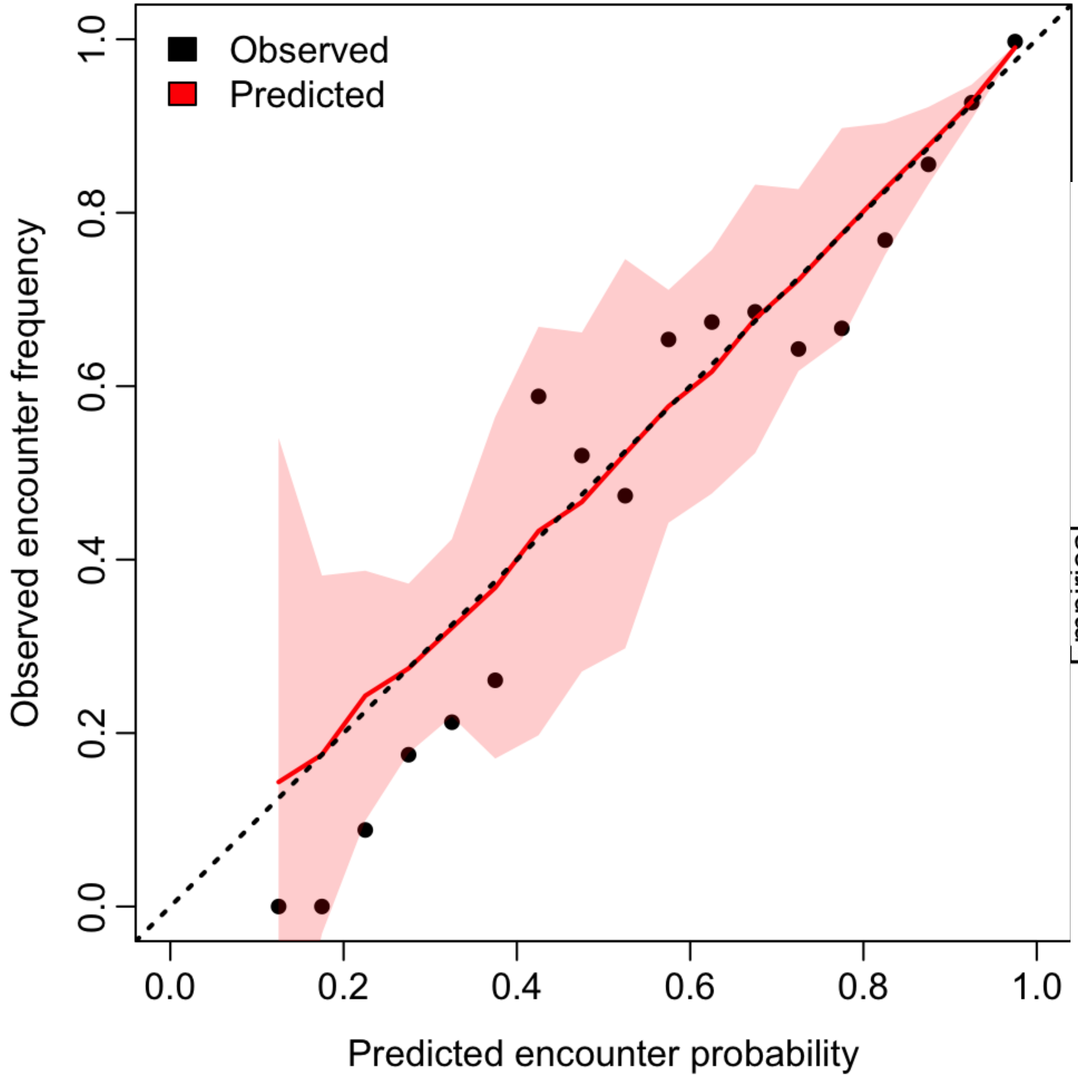
Index of abundance

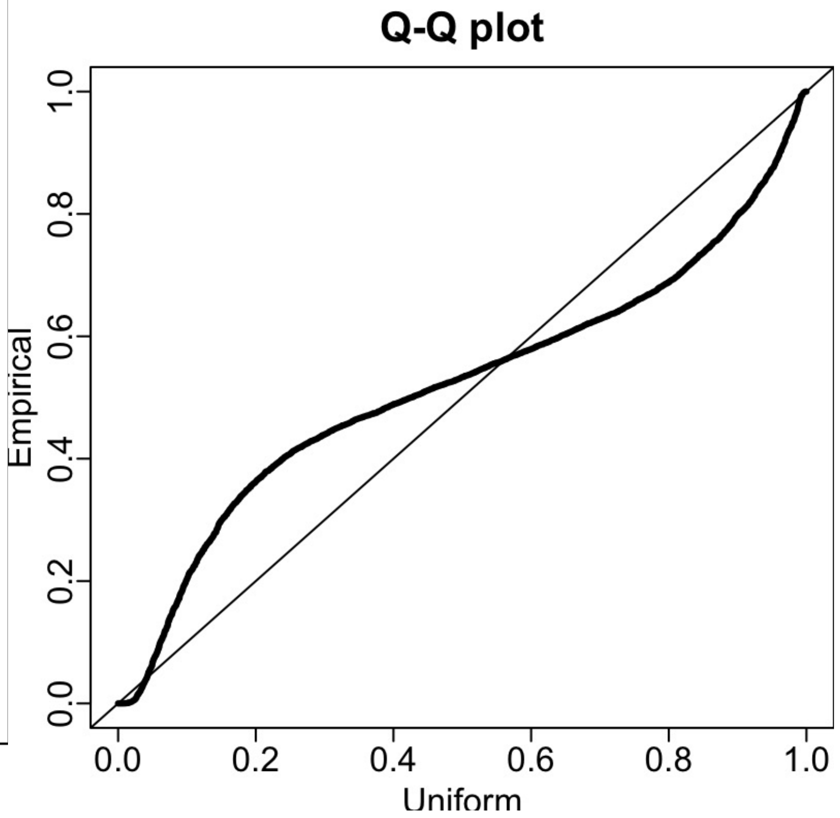


Compared to spawning biomass ratio

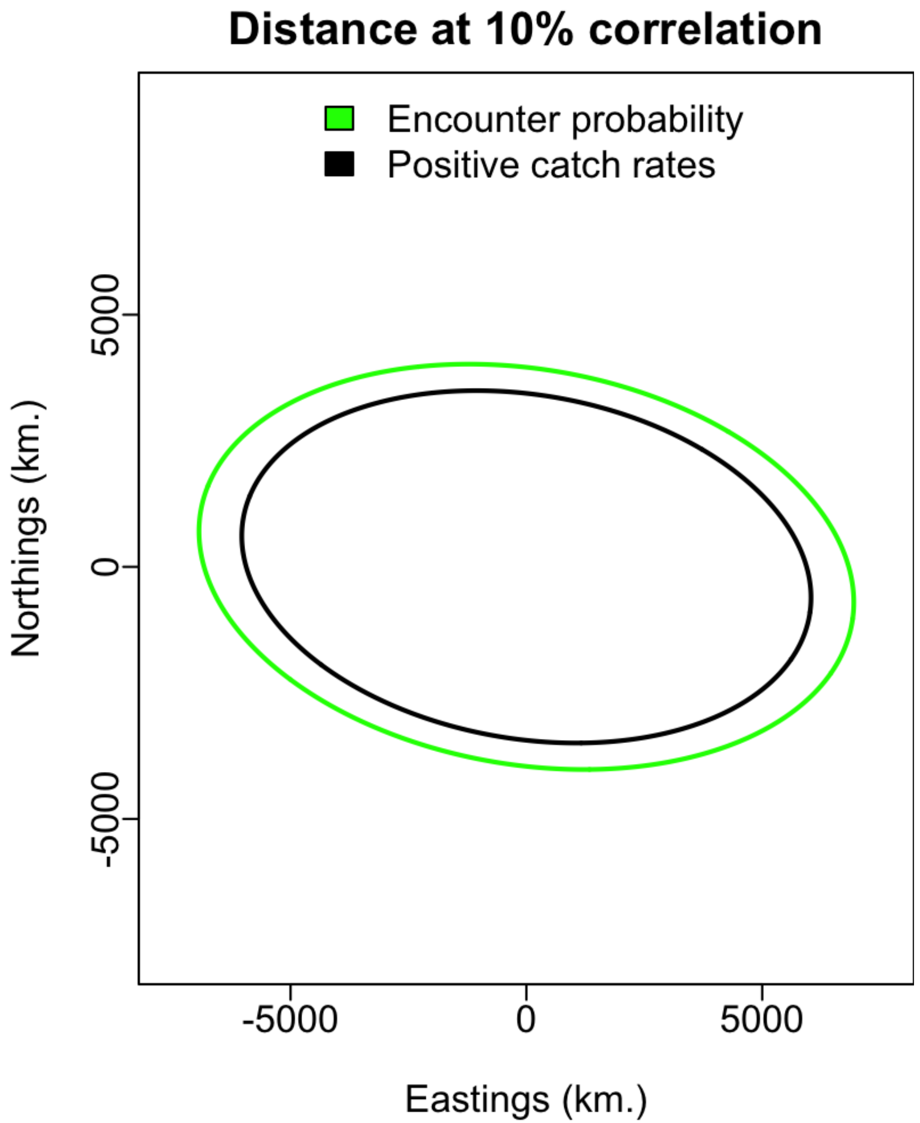


Diagnostics





Anisotropy

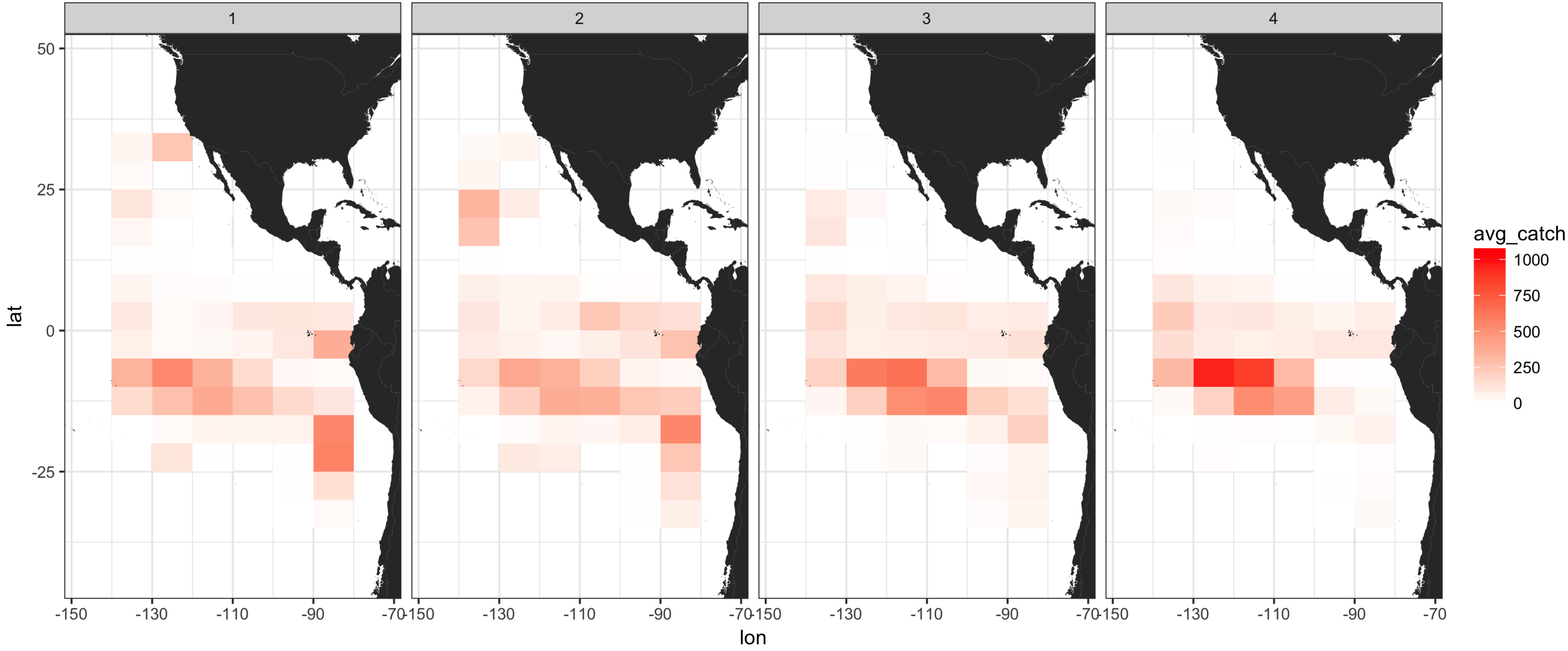


Distance at 10% correlation
Encounter probability
Positive catch rates
Northings (km.)
5000
0
-5000
-5000
0
5000
Eastings (km.)

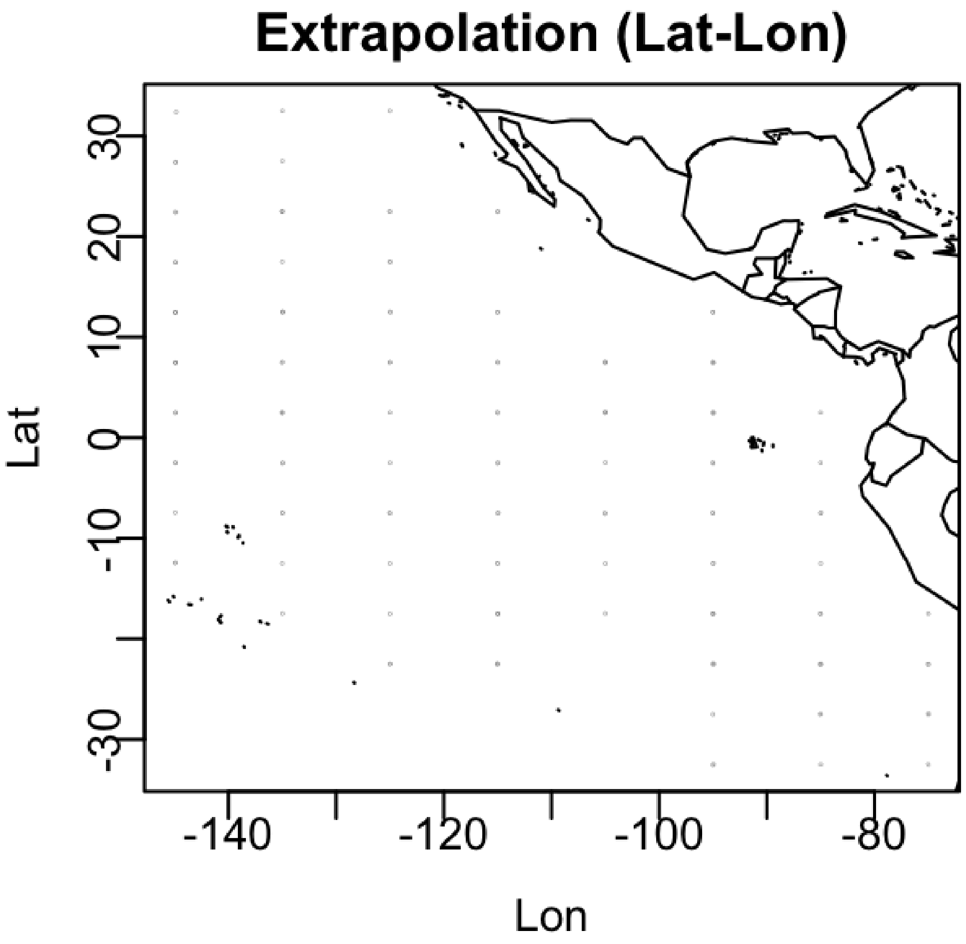
VAST model – length comps

- Counts (100% encounters)
- 5° latitude × 10° longitude
- Length quartiles

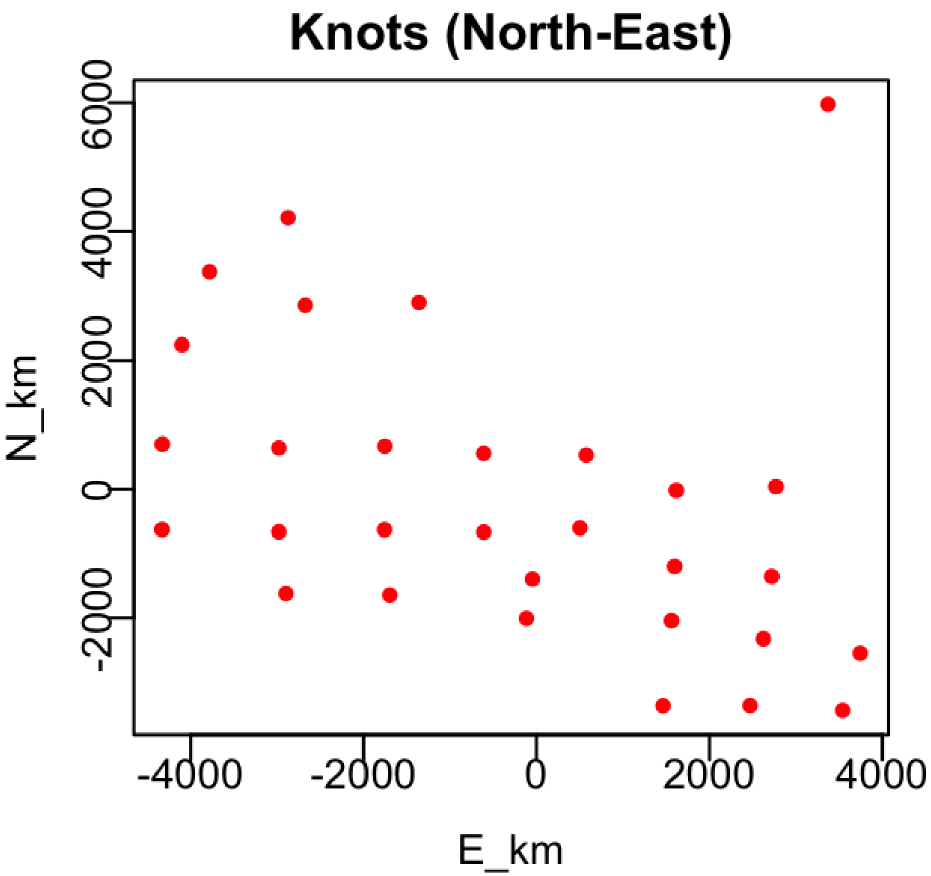
Data-length compositions



30 knots

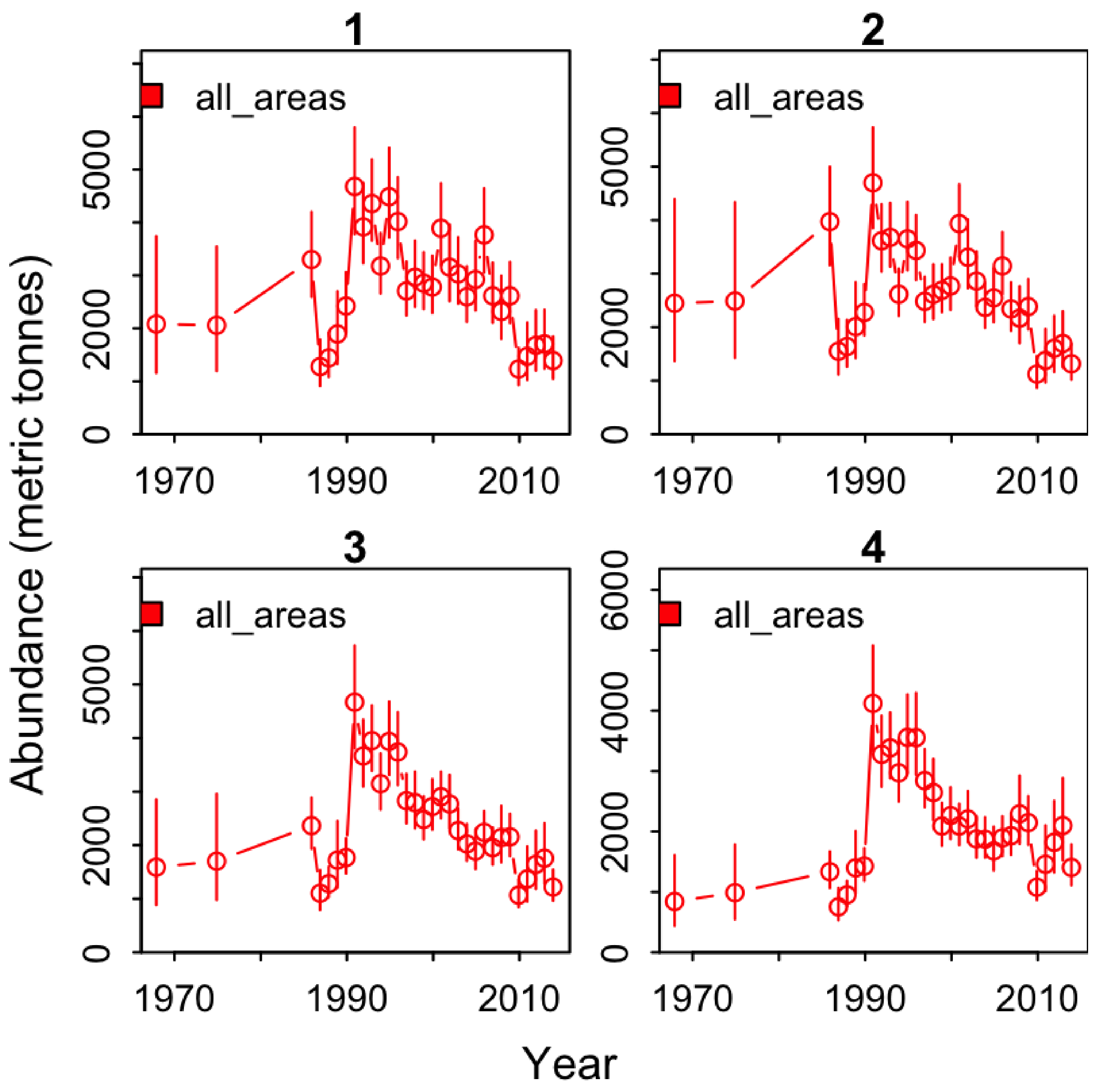


Extrapolation (Lat-Lon)
Lat
30
20
10
0
-10
-30
-140
-120
-100
-80
Lon

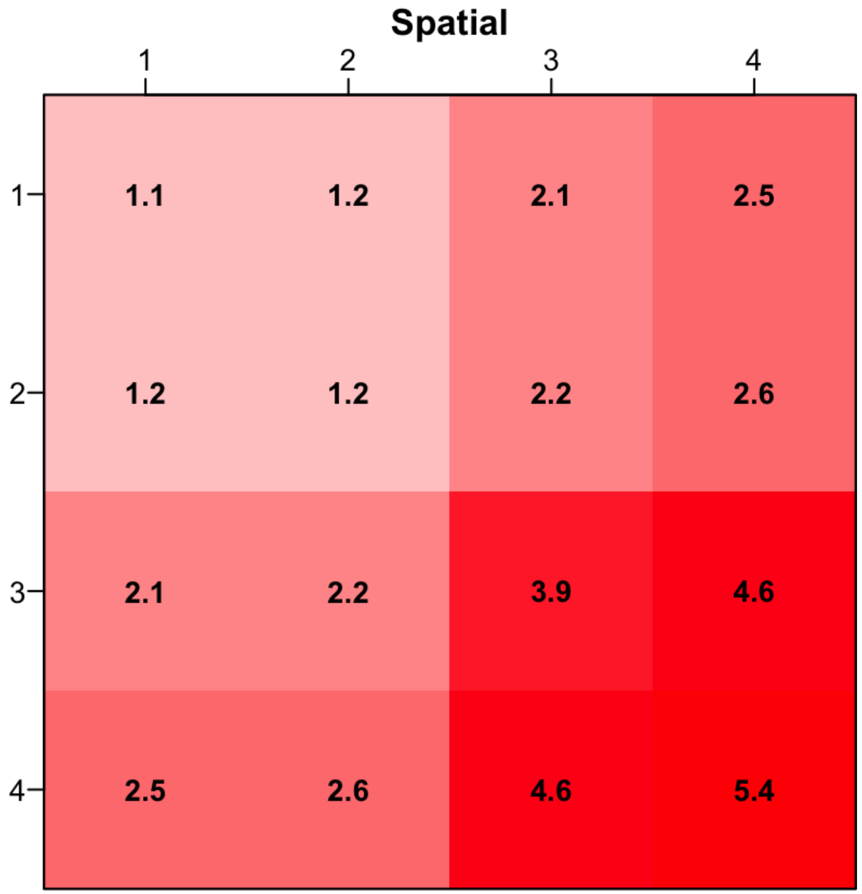


Knots (North-East)
N_km
6000
4000
2000
0
-2000
-4000
-2000
0
2000
4000
E_km

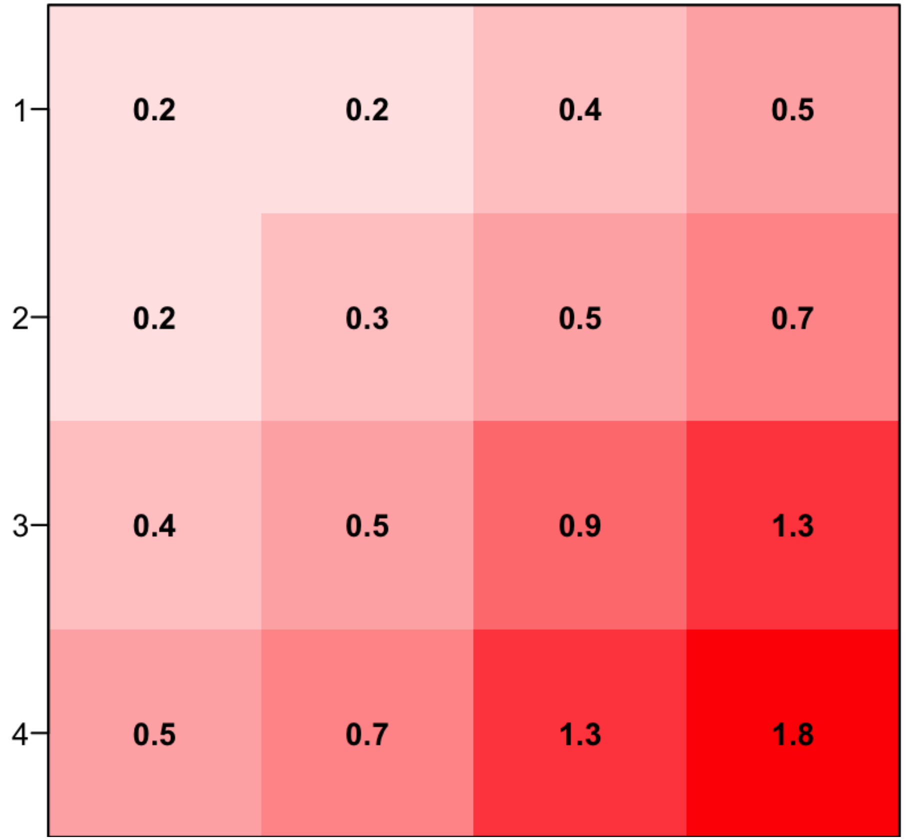
Index


1
2
3
4
all_areas
Abundance (metric tonnes)
Year
0
2000
5000
4000
6000
1970
1990
2010

Covariance

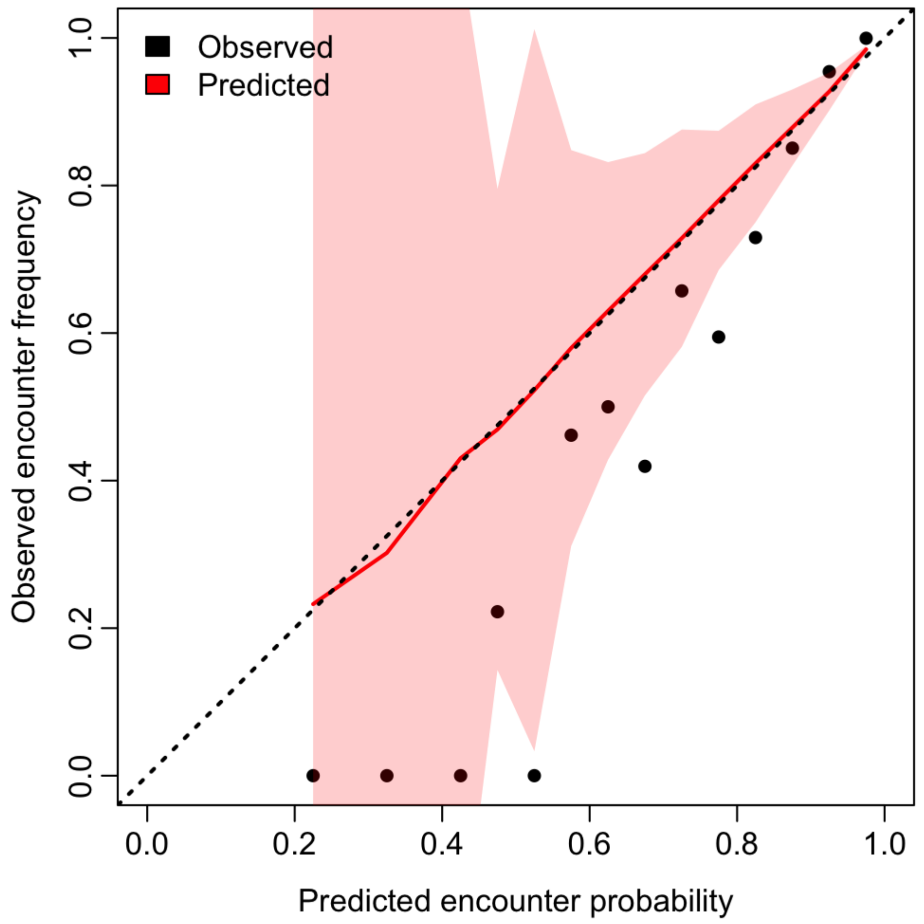


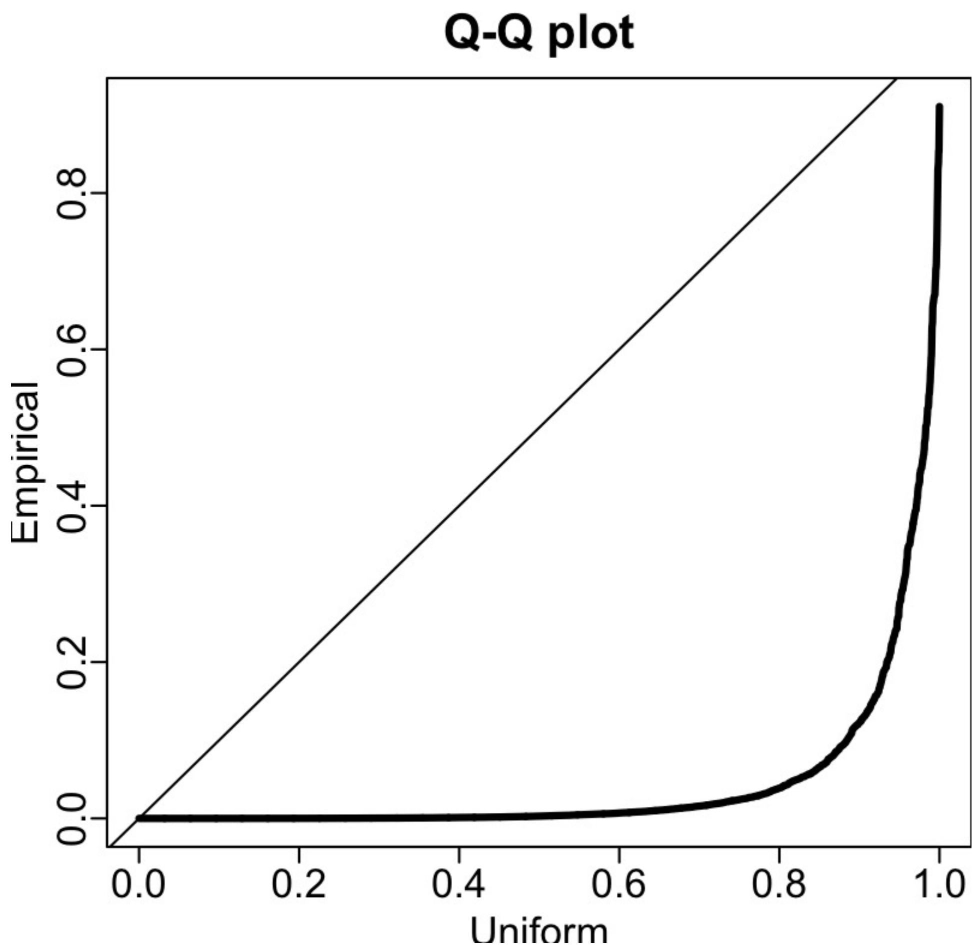
Encounter



Catch rate

Problems


Observed
Predicted
Observed encounter frequency
Predicted encounter probability
0.0
0.2
0.4
0.6
0.8
1.0


Q-Q plot
Empirical
Uniform
0.0
0.2
0.4
0.6
0.8
1.0

Open questions

- Handling 0% and 100% encounter rate
- Spatial modeling over ~500km by ~500km grids?
  - Bigger in some cases
- Discussion:
- Bad fit or bad diagnostics?
- How many number of bins?
- How many number of factors?

- Backup slides

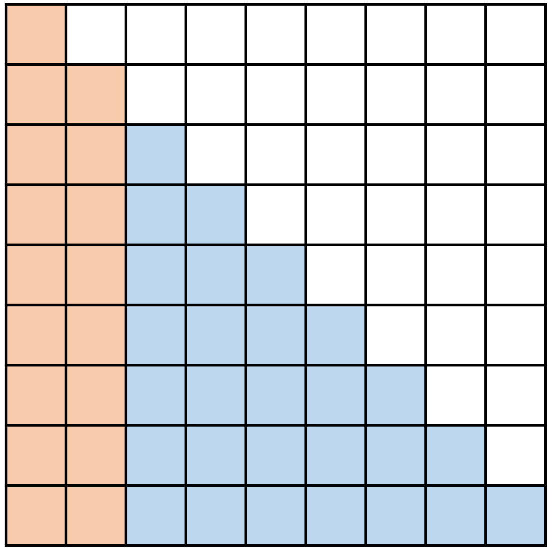
Problems in this length comp modelling

- How to eliminate the possibility that the size segregation is not caused by including only 2 loading factors for 9 length bins?

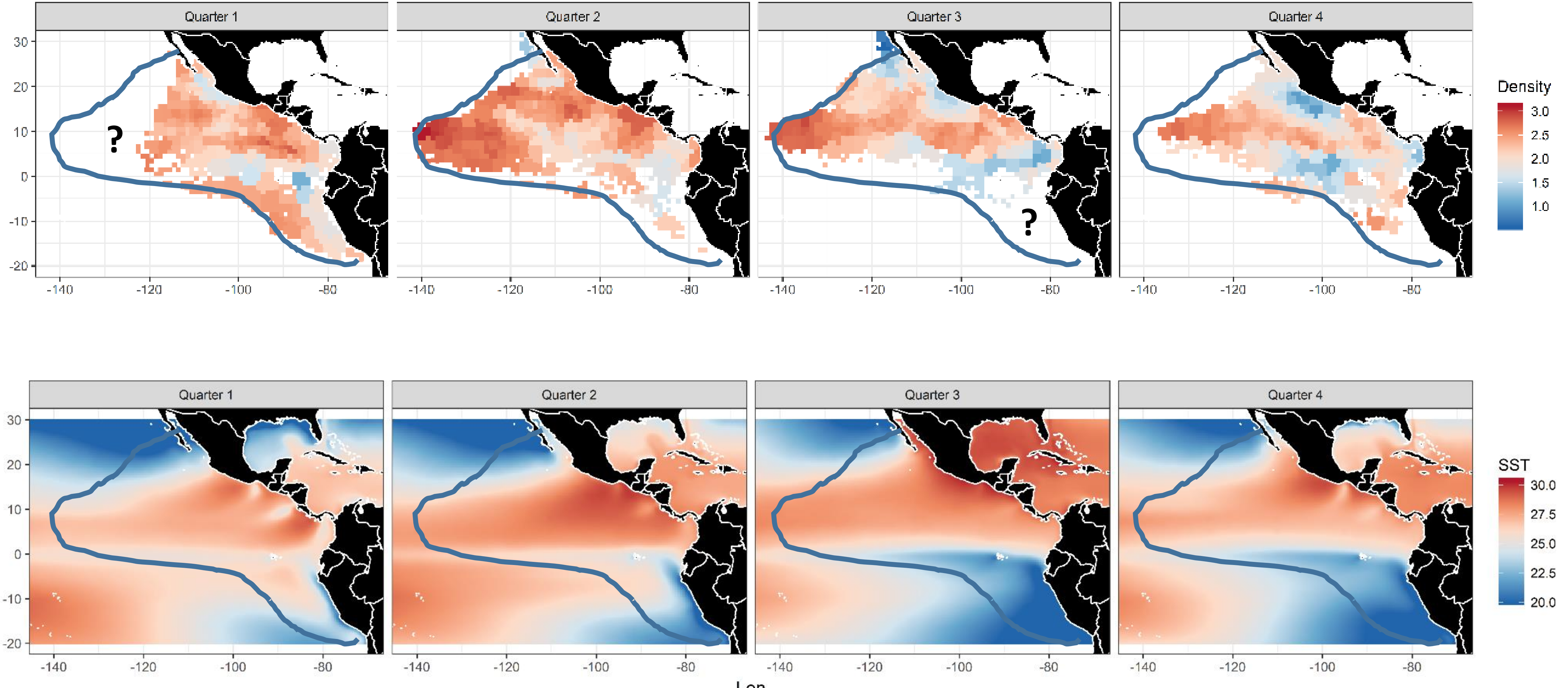
$$p_i = \text{logit}^{-1}\left(\beta_1(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_1}(l_i, f)\omega_1(s_i, f) + \sum_{f=1}^{n_f} L_{\varepsilon_1}(l_i, f)\varepsilon_1(s_i, f, t_i)\right)$$

$$\lambda_i = \exp\left(\beta_2(l_i, t_i) + \sum_{f=1}^{n_f} L_{\omega_2}(l_i, f)\omega_2(s_i, f) + \sum_{f=1}^{n_f} L_{\varepsilon_2}(l_i, f)\varepsilon_2(s_i, f, t_i)\right)$$

L =



The fishery follows the movement of warm waters



Plan for the next step

1. Find the key env driver(s) that causes the size-segregation in spatial and spatiotemporal distribution
2. Calculate the mean predicted catch rate at length as a function of the driver
3. Find the SST range in which

$$\log(d_{large}) - \log(d_{small}) = \log\left(\frac{d_{large}}{d_{small}}\right) \text{ is largest}$$

4. Improve spatial management: predicting preferred fishing grounds using real-time env observations