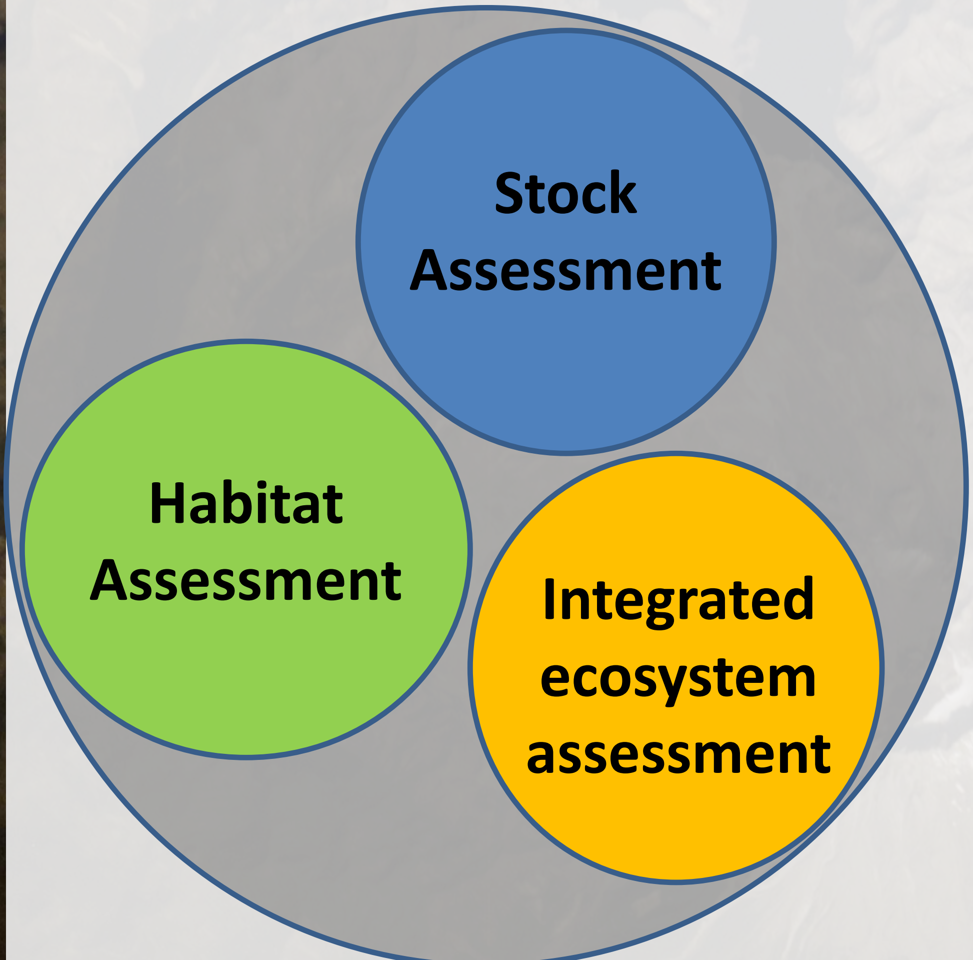
James Thorson
(with guidance from many people in attendance!)

Combining stock assessment, habitat, ecosystem, and climate research using multivariate spatio-temporal models

Spatio-temporal model



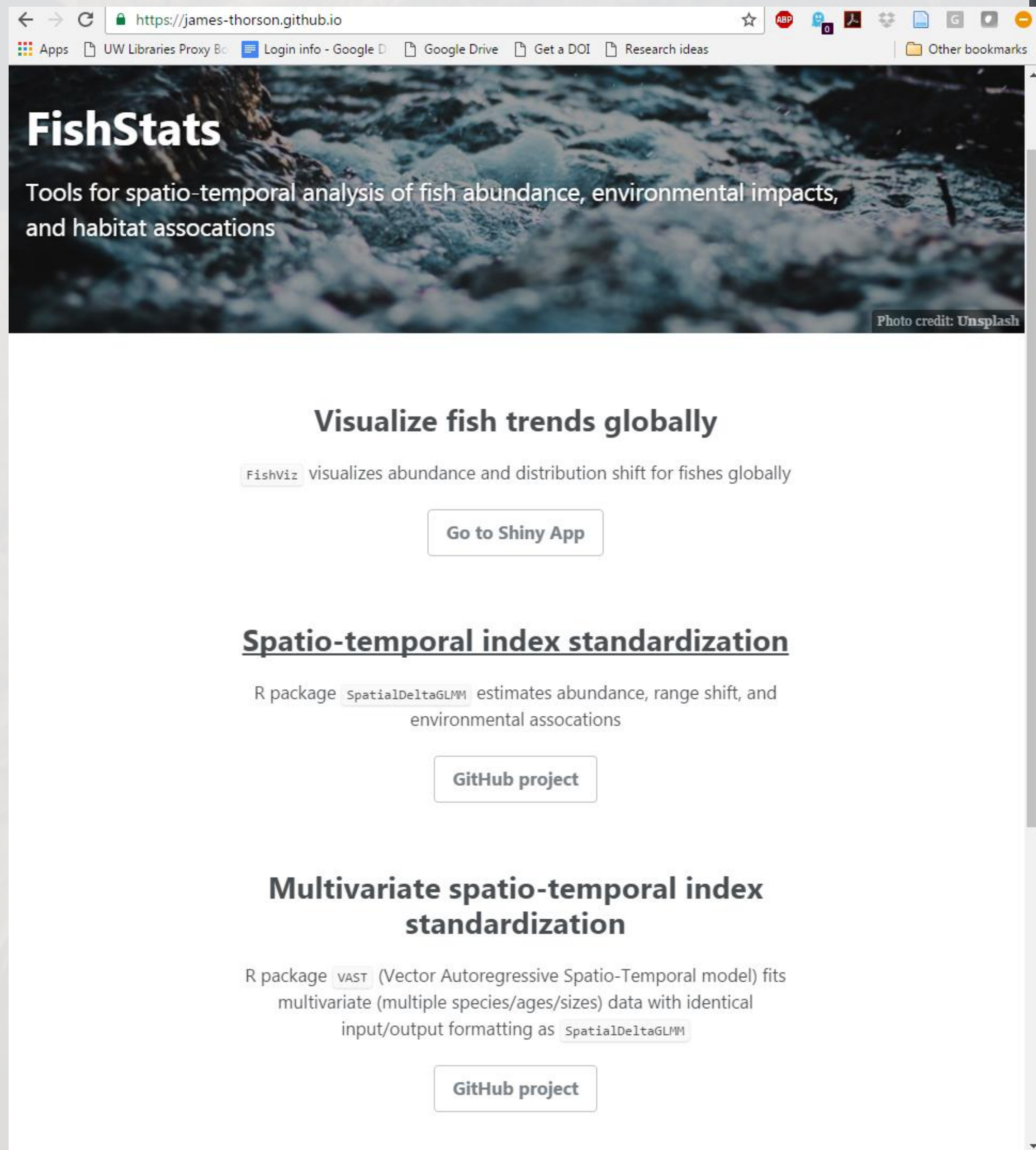
Benefits of single approach

1. Include biological mechanism
2. Improved communication
3. Similar review standards and “burden of proof”

Spatio-temporal fisheries toolbox

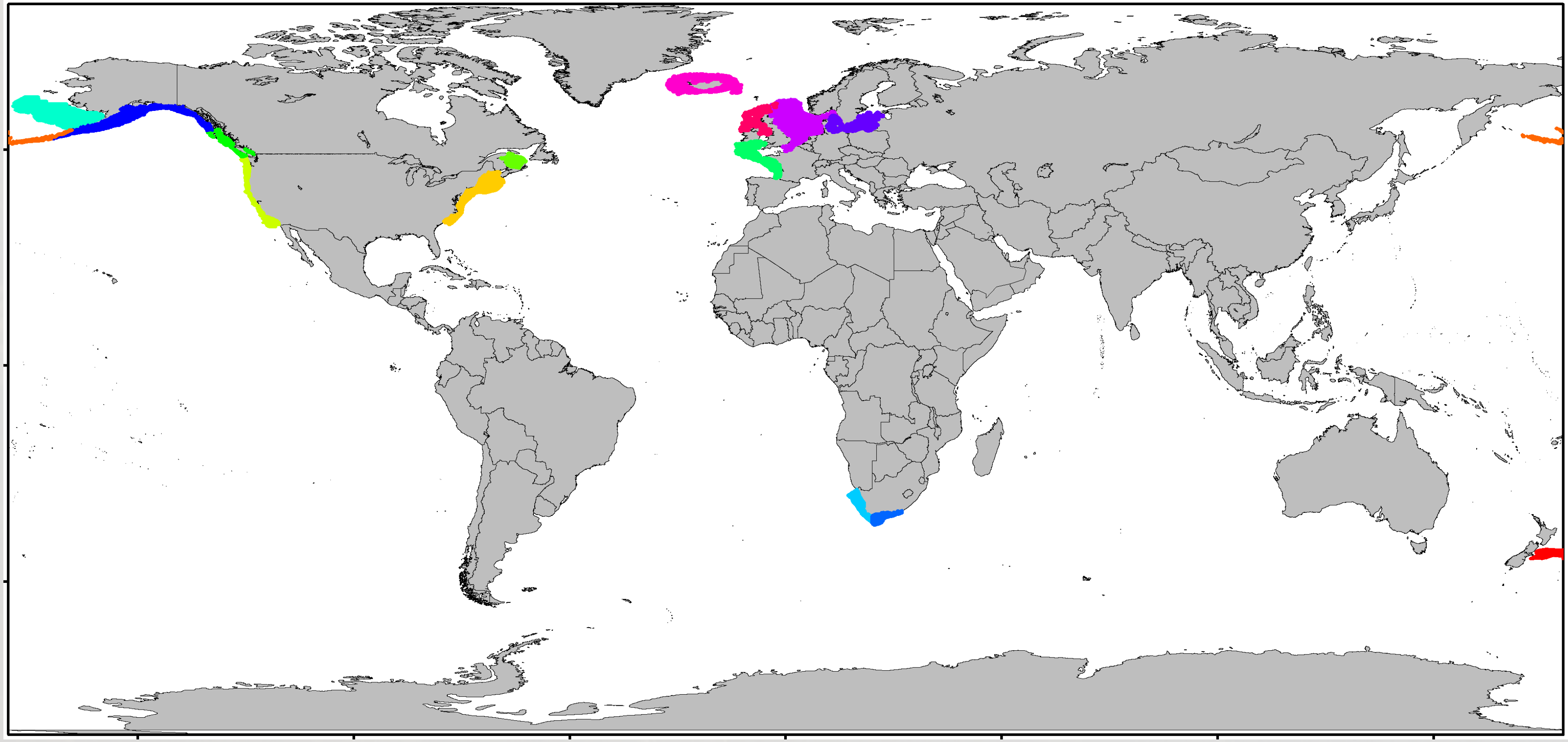
www.FishStats.org

1. FishViz
   – Visualizes results worldwide
2. VAST
   – Multi-species index model
3. MIST
   – Estimate multispecies interactions
4. FishData
   – Scrape data worldwide
5. FishStats-listserv
   – Community updates by email



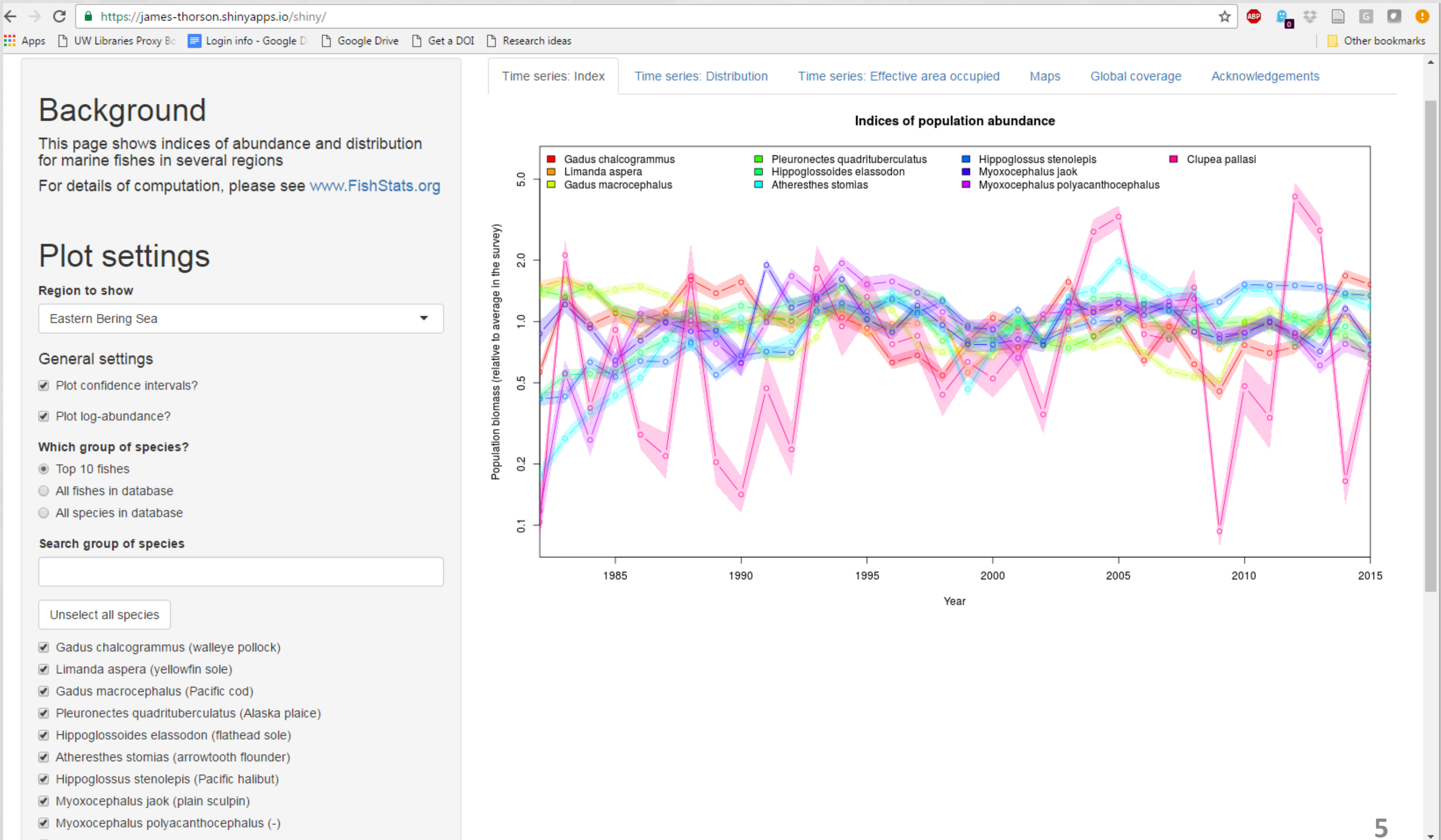
Has been applied to >15 regions worldwide

```
> devtools::install_github("james-thorson/FishData")
Downloading GitHub repo james-thorson/FishData@master
from URL https://api.github.com/repos/james-thorson/FishData/zipball/
Installing FishData
```



Currently showing results for >500 stocks

@ www.FishViz.org



Delta-generalized linear mixed model (Delta-GLMM)

- Delta-model for observations

$$\Pr(B=b)=\begin{cases}1-\gamma(s,t) & \text{if } B=0\\ \gamma(s,t)\times g(B;\lambda(s,t)) & \text{if } B>0\end{cases}$$

  - Where $\gamma(s,t)$ is the probability of encountering the species
  - $g(B;\lambda(s,t))$ is a distribution for positive catches

- Spatio-temporal variation in encounter probability

$$\mathrm{logit}(\gamma(s,t))=\alpha_\gamma(t)+\omega_\gamma(s)+\varepsilon_\gamma(s,t)$$

  - $\alpha_\gamma(t)$ is the intercept for each year
  - Where $\boldsymbol{\omega}_\gamma$ and $\boldsymbol{\varepsilon}_\gamma(t)$ follow a spatial distribution

- Spatio-temporal variation in density

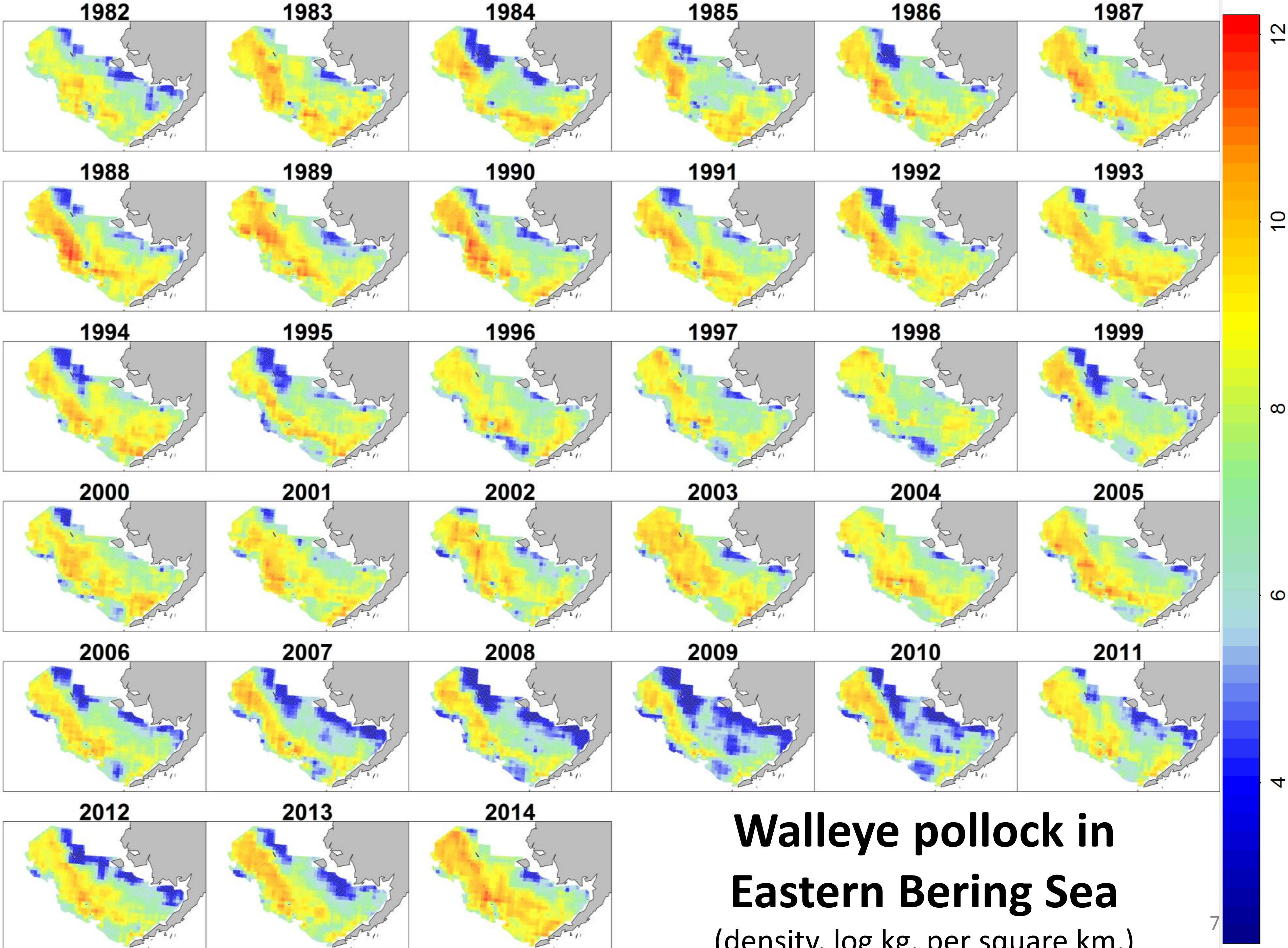
$$\log(\lambda(s,t))=\alpha_\lambda(t)+\omega_\lambda(s)+\varepsilon_\lambda(s,t)$$

  - Where parameters are defined similarly to $\gamma(s,t)$

- Used to predict local density

$$\hat{d}(s,t)=\hat{\gamma}(s,t)\times\hat{\lambda}(s,t)$$

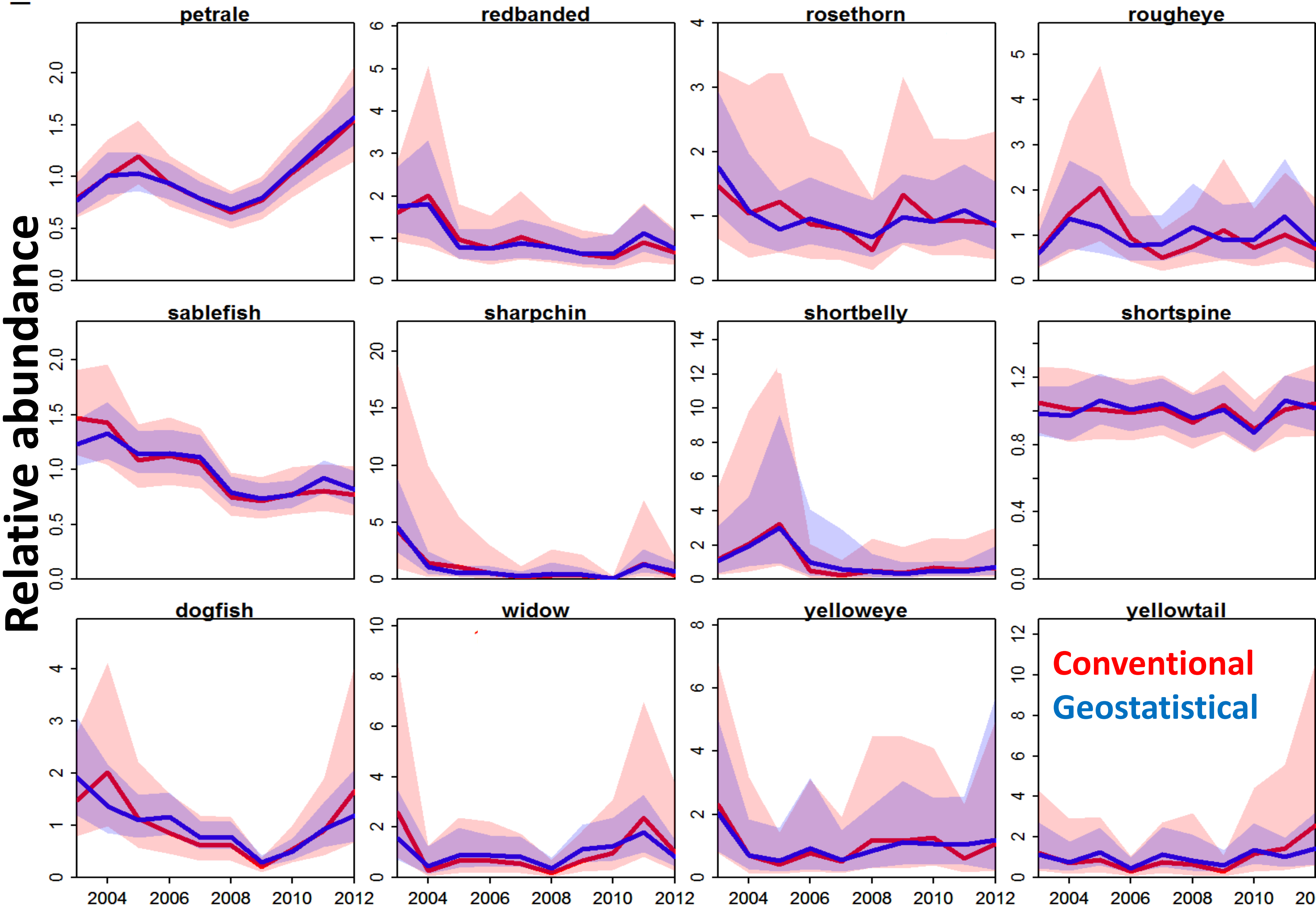
  - Where $\hat{\gamma}(s,t)$ and $\hat{\lambda}(s,t)$ are predictions conditioned on data



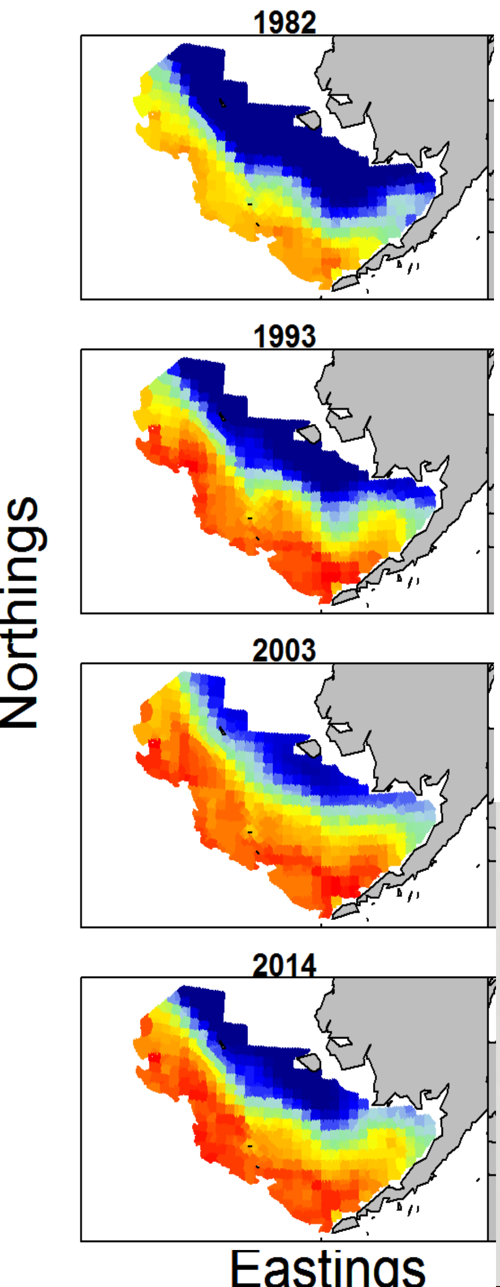
Walleye pollock in Eastern Bering Sea

(density, log kg. per square km.)

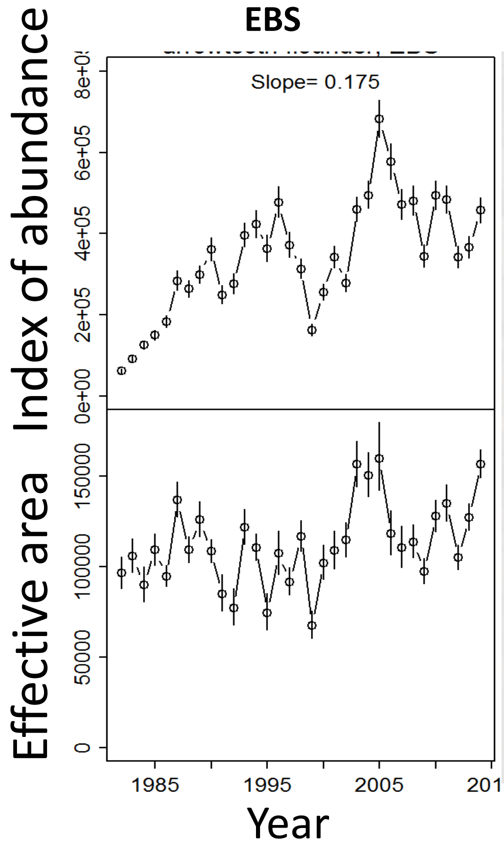
Abundance indices



Arrowtooth flounder
Eastern Bering Sea



Arrowtooth flounder
EBS



Thorson, Rindorf, Gao, Hanselman, and Winker. 2016. Density-dependent changes in effective area occupied for sea-bottom-associated marine fishes. Proc R Soc B **283**(1840).

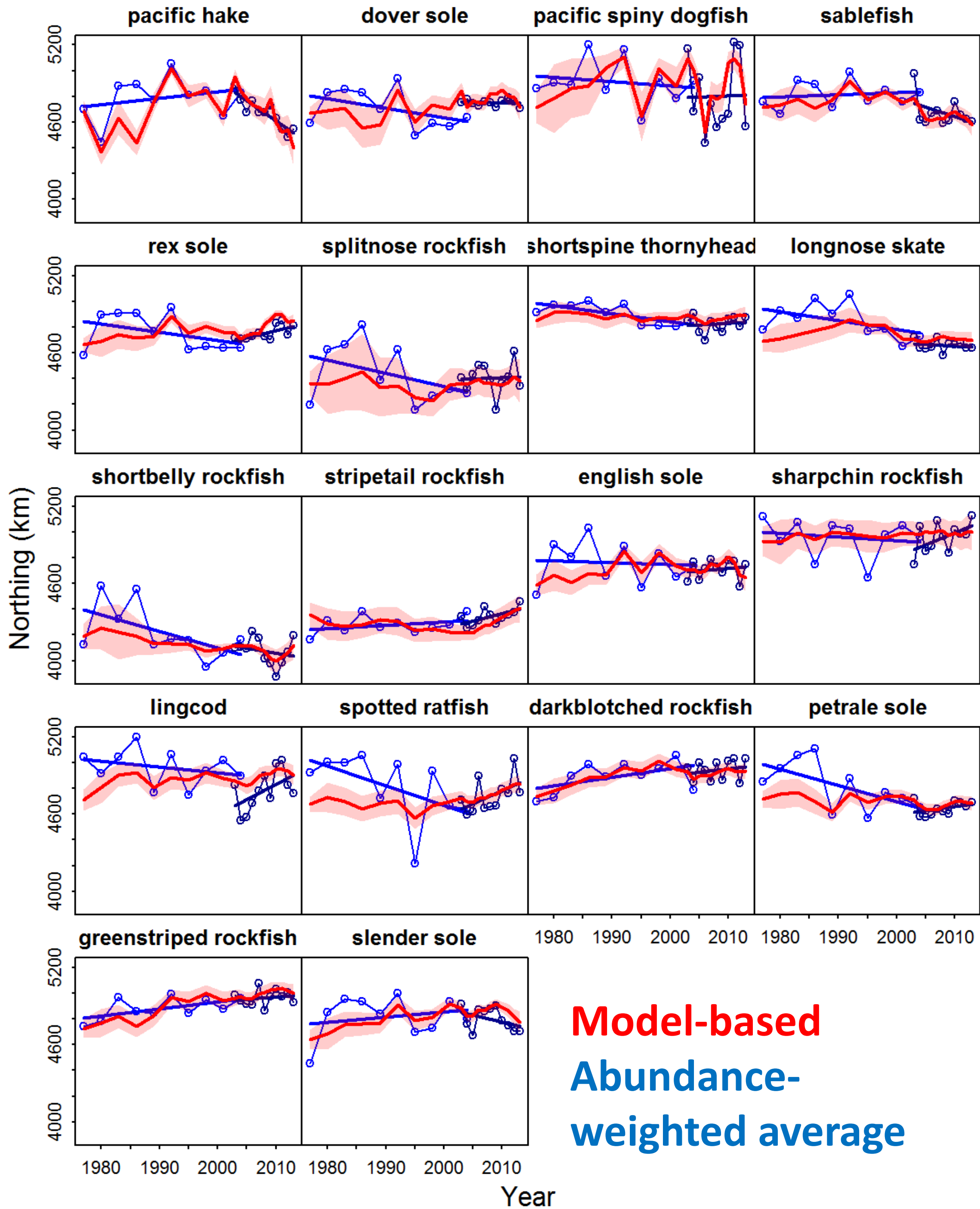
Density-dependent habitat selection

- Do populations shrink their range when abundance is low?
- Average
  - Small contraction in range
  - Greatest in Eastern Bering Sea

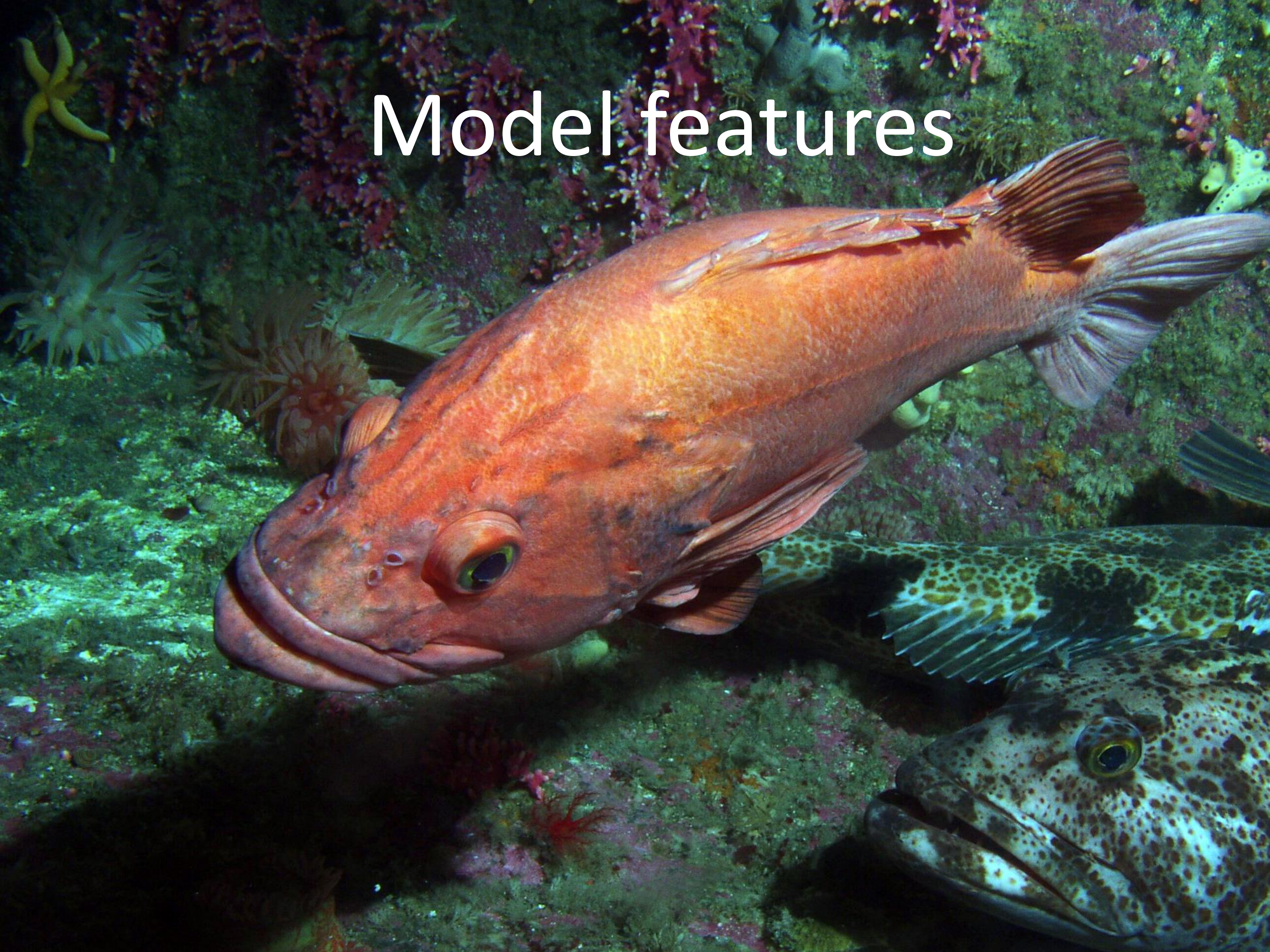
Distribution shifts

- Highly variable distribution for semi-pelagic species
  - Dogfish
  - Sablefish
  - Hake
- Few clear trends
  - Depends on time-scale

Thorson, Pinsky, and Ward. 2016. Model-based inference for estimating shifts in species distribution, area occupied and centre of gravity. Methods Ecol. Evol. **7**(8): 990–1002.



Model features



use ObsModel[2]=3

Delta-generalized linear mixed model

Basic features

- Specifying distributions for data
- Specifying link functions for predicting data given linear predictors
- Dynamic habitat covariates
- Catchability covariates

use ObsModel[2]=3

Delta-generalized linear mixed model

Spatio-temporal features

- Define “extrapolation grid”
- Specifying a spatial smoother
- Specifying number of “knots”
- Include/exclude spatial variation for each of two linear predictors
- Include/exclude spatio-temporal variation for each of two linear predictors

use ObsModel[2]=3

Delta-generalized linear mixed model

Derived quantities

- Specifying strata for derived quantities;
- Select "derived quantities" to calculate from:
  - range shift
  - effective area occupied
  - abundance indices
  - covariance among categories within a multivariate model
  - synchrony among categories.

use ObsModel[2]=3

Delta-generalized linear mixed model

Temporal structure

- Annual intercepts being estimated as
  - fixed effects in every year
  - fixed at the same value for all years
  - random effect by year
  - first-order autoregressive structure
  - random-walk structure.
- Spatio-temporal variation being estimated as
  - independent deviations in each year
  - first-order autoregressive structure over time
  - random-walk structure over time.

use ObsModel[2]=3

Delta-generalized linear mixed model

Multivariate analysis

- Include a “multivariate” structure with multiple responses
  - Rank of covariance chosen by user covary due to a specified number of “factors” for spatial and spatio-temporal terms;
- Rotate results prior to visualization
  - Principle component rotation
  - Varimax rotation

use ObsModel[2]=3

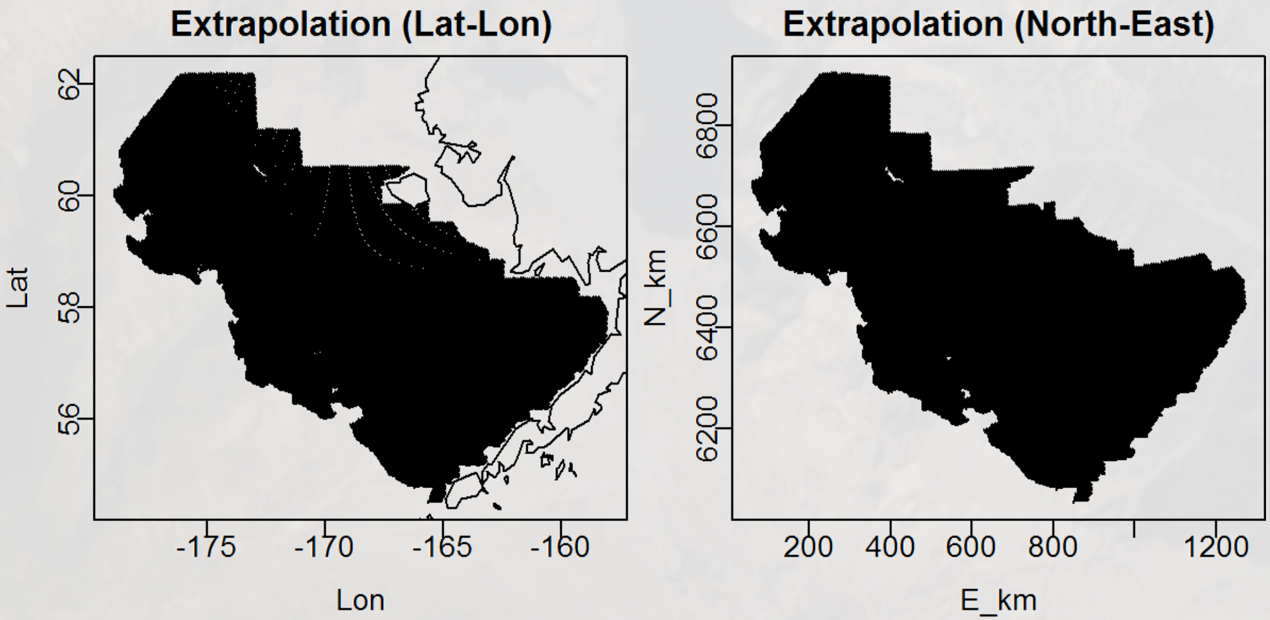
Delta-generalized linear mixed model

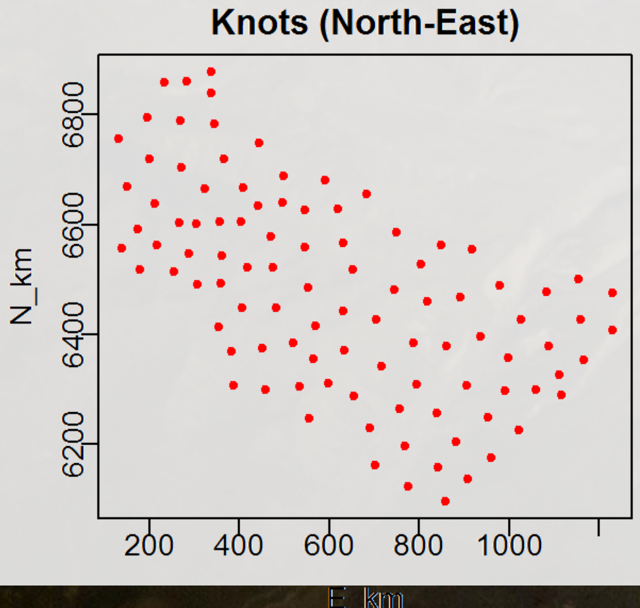
Unusual circumstances and spatial cases

- Specifying multiple distributions for response
- Specifying that some data are predicted based on summing linear predictors across multiple variables
- Specifying multiple “seasons”

Model diagnostics

Diagnostics

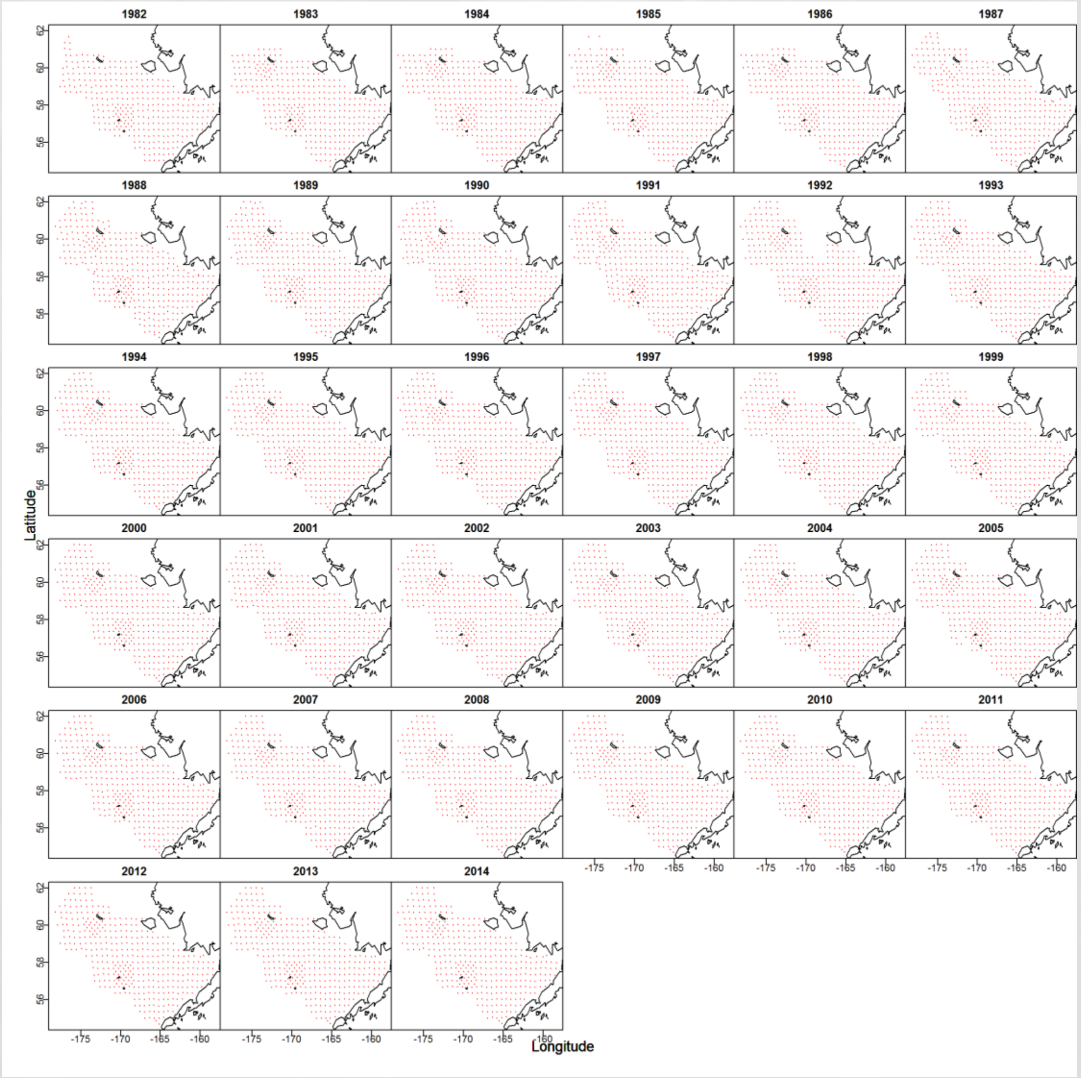




Advice: Inspect extrapolation footprint and knots

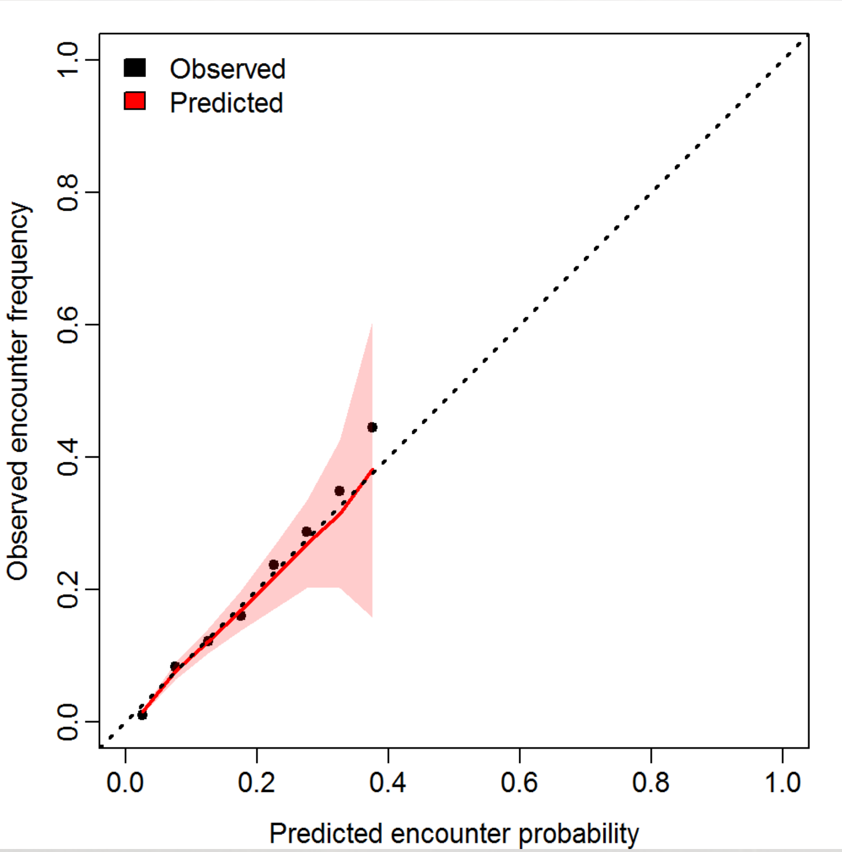
Diagnostics

Advice: Inspect spatial distribution of data

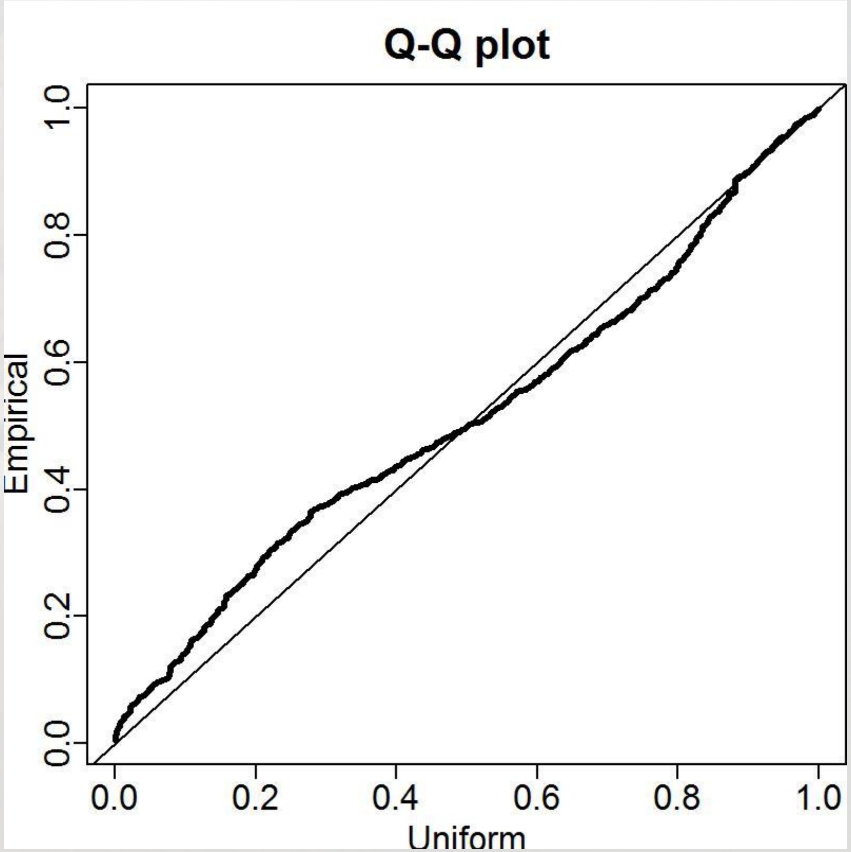


Diagnostics

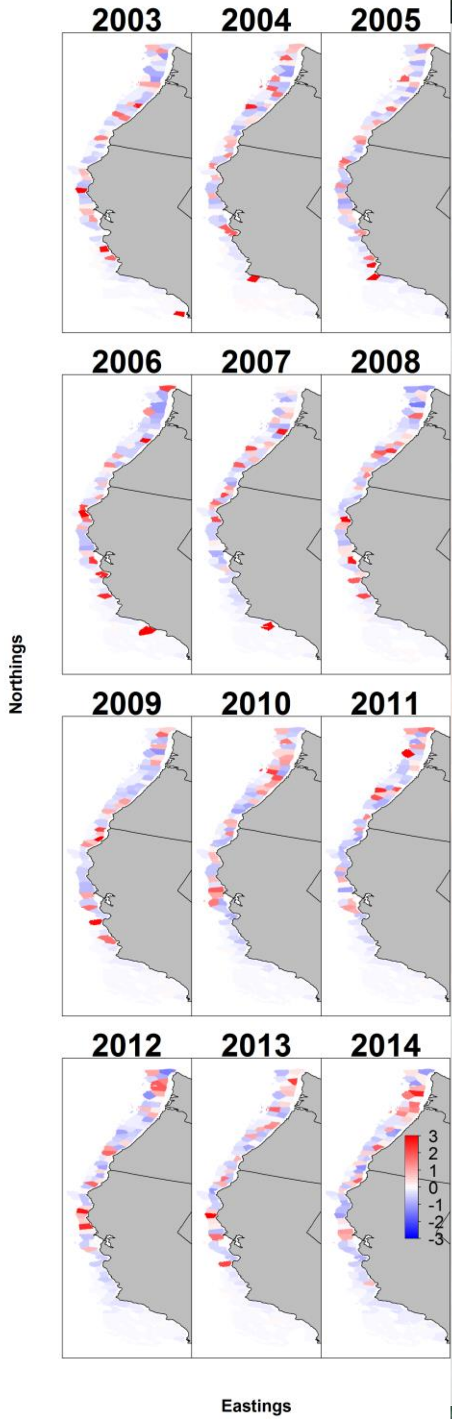
Encounter probability vs. frequency



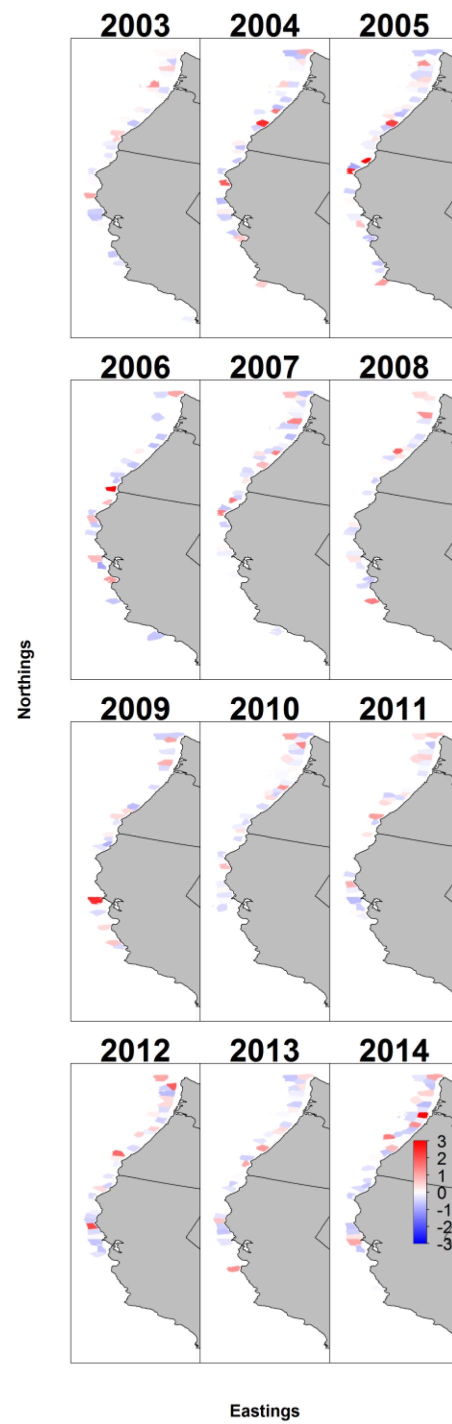
Quantile-quantile plot for positive catch rates



Pearson residuals for encounter



Pearson residuals for positive catch rates

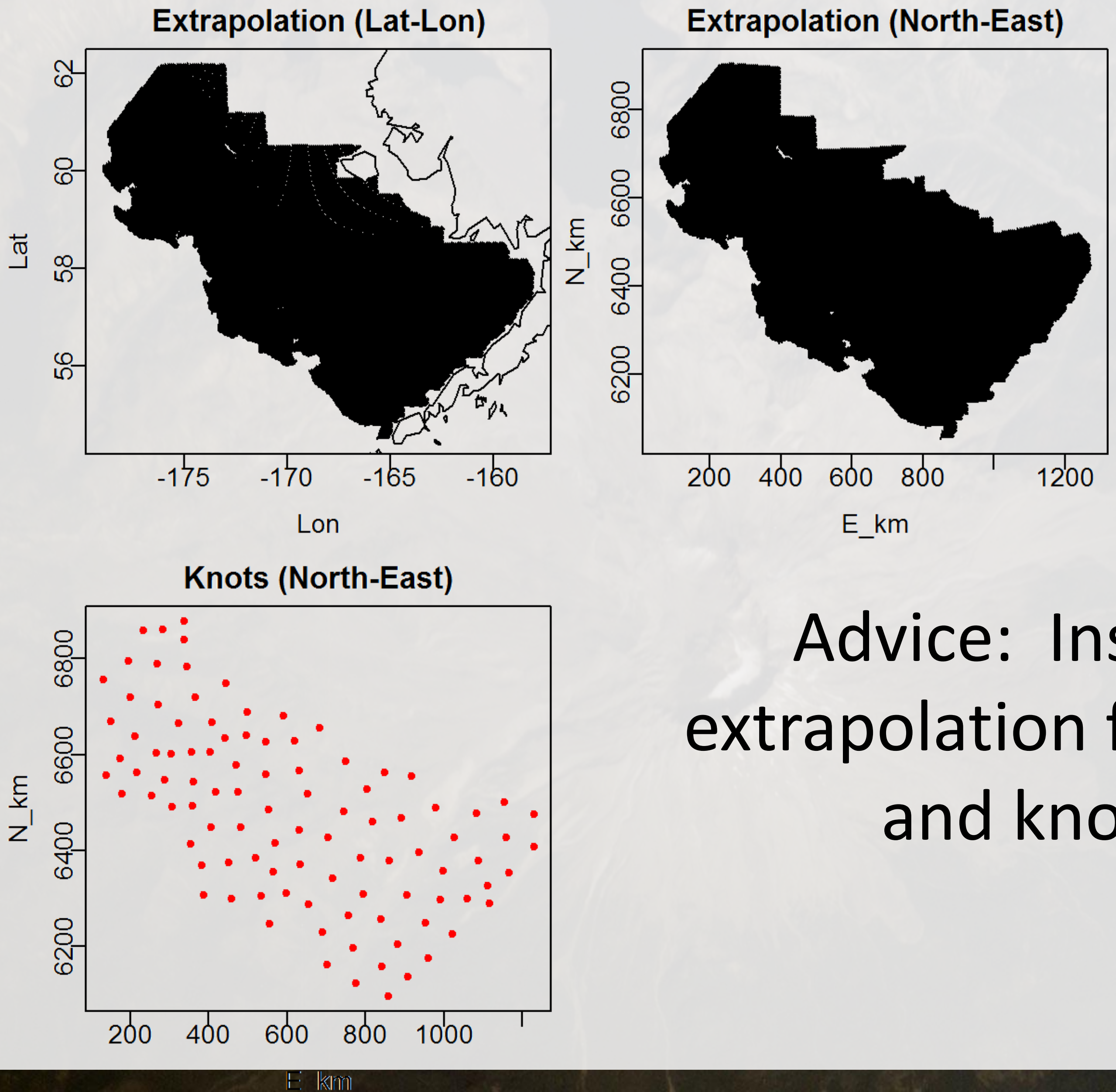


Advice – Look at bounds and gradients

	Param	starting_value	Lower	MLE	Upper	final_gradient
1	ln_H_input	0	-50	0.231528	50	-6.19E-08
2	ln_H_input	0	-50	-0.96568	50	-8.88E-08
3	beta1_ct	-4.64096	-50	4.120475	50	2.44E-09
4	beta1_ct	-4.64096	-50	4.228782	50	1.90E-09
5	beta1_ct	-4.64096	-50	4.322799	50	-5.30E-10
6	beta1_ct	-4.64096	-50	5.093036	50	4.60E-09
7	beta1_ct	-4.64096	-50	5.428053	50	-5.17E-08
8	beta1_ct	-4.64096	-50	4.105238	50	-2.95E-09
9	beta1_ct	-4.64096	-50	5.056347	50	4.29E-09
10	beta1_ct	-4.64096	-50	4.168261	50	-5.53E-09
11	beta1_ct	-4.64096	-50	4.333523	50	7.79E-09
12	beta1_ct	-4.64096	-50	5.989274	50	-1.02E-08
13	beta1_ct	-4.64096	-50	4.524008	50	3.45E-09
14	beta1_ct	-4.64096	-50	5.265399	50	3.09E-09
15	beta1_ct	-4.64096	-50	5.646847	50	-9.95E-11
16	beta1_ct	-4.64096	-50	4.886118	50	3.52E-09
17	beta1_ct	-4.64096	-50	5.073619	50	4.40E-09
18	beta1_ct	-4.64096	-50	4.753279	50	5.61E-09
19	beta1_ct	-4.64096	-50	4.996536	50	4.81E-09
20	beta1_ct	-4.64096	-50	6.218751	50	1.38E-09
21	beta1_ct	-4.64096	-50	5.124685	50	-3.66E-09
22	beta1_ct	-4.64096	-50	5.706784	50	-8.00E-09
23	beta1_ct	-4.64096	-50	4.80919	50	4.36E-09
24	beta1_ct	-4.64096	-50	4.534566	50	6.15E-09
25	beta1_ct	-4.64096	-50	5.45406	50	-1.36E-09
26	beta1_ct	-4.64096	-50	4.746618	50	4.78E-09
27	beta1_ct	-4.64096	-50	4.572286	50	8.21E-09
28	beta1_ct	-4.64096	-50	4.198098	50	1.19E-08
29	beta1_ct	-4.64096	-50	2.877037	50	1.34E-08
30	beta1_ct	-4.64096	-50	3.426151	50	8.29E-09
31	beta1_ct	-4.64096	-50	2.986486	50	5.37E-09
32	beta1_ct	-4.64096	-50	4.659832	50	4.15E-09
33	beta1_ct	-4.64096	-50	4.656848	50	7.09E-09
34	beta1_ct	-4.64096	-50	5.18952	50	4.60E-09
35	beta1_ct	-4.64096	-50	6.231048	50	1.60E-09
36	L_omega1_z	-0.83795	-50	-1.94641	50	1.61E-07
37	L_epsilon1_z	1.037078	-50	0.975252	50	-3.53E-07
38	logkappa1	-0.10536	-6.01012	-4.12041	-2.57395	4.33E-07

	Param	starting_value	Lower	MLE	Upper	final_gradient
39	beta2_ct	9.21962	-50	7.516834	50	2.20E-10
40	beta2_ct	9.21962	-50	8.739776	50	-1.49E-10
41	beta2_ct	9.21962	-50	7.843733	50	1.06E-09
42	beta2_ct	9.21962	-50	8.534672	50	1.02E-09
43	beta2_ct	9.21962	-50	8.097048	50	1.03E-09
44	beta2_ct	9.21962	-50	8.458756	50	3.47E-10
45	beta2_ct	9.21962	-50	8.286936	50	-1.06E-10
46	beta2_ct	9.21962	-50	8.242662	50	2.49E-11
47	beta2_ct	9.21962	-50	8.045717	50	9.28E-10
48	beta2_ct	9.21962	-50	8.170187	50	3.41E-11
49	beta2_ct	9.21962	-50	8.06304	50	9.76E-10
50	beta2_ct	9.21962	-50	8.212494	50	-5.63E-10
51	beta2_ct	9.21962	-50	8.008228	50	-1.80E-10
52	beta2_ct	9.21962	-50	7.516414	50	-7.77E-10
53	beta2_ct	9.21962	-50	7.730084	50	-2.10E-10
54	beta2_ct	9.21962	-50	7.886746	50	2.17E-10
55	beta2_ct	9.21962	-50	7.662614	50	-5.12E-10
56	beta2_ct	9.21962	-50	7.40508	50	7.91E-10
57	beta2_ct	9.21962	-50	8.197652	50	-7.12E-10
58	beta2_ct	9.21962	-50	8.165989	50	-2.23E-10
59	beta2_ct	9.21962	-50	7.847344	50	-2.63E-09
60	beta2_ct	9.21962	-50	8.542195	50	1.28E-10
61	beta2_ct	9.21962	-50	7.982901	50	-1.07E-09
62	beta2_ct	9.21962	-50	7.832611	50	-4.09E-10
63	beta2_ct	9.21962	-50	7.129841	50	-1.78E-10
64	beta2_ct	9.21962	-50	6.996498	50	4.58E-10
65	beta2_ct	9.21962	-50	6.544465	50	-1.44E-09
66	beta2_ct	9.21962	-50	6.056259	50	8.04E-10
67	beta2_ct	9.21962	-50	7.290846	50	1.02E-09
68	beta2_ct	9.21962	-50	7.545933	50	5.06E-10
69	beta2_ct	9.21962	-50	7.247531	50	-1.24E-09
70	beta2_ct	9.21962	-50	7.513136	50	1.12E-09
71	beta2_ct	9.21962	-50	8.565377	50	-9.49E-10
72	L_omega2_z	-0.58178	-50	-1.10638	50	-9.35E-09
73	L_epsilon2_z	-0.45241	-50	-1.12348	50	1.43E-07
74	logkappa2	-0.10536	-6.01012	-4.53498	-2.57395	3.22E-08
75	logSigmaM	1.609438	-50	0.168158	10	-2.42E-07

Diagnostics

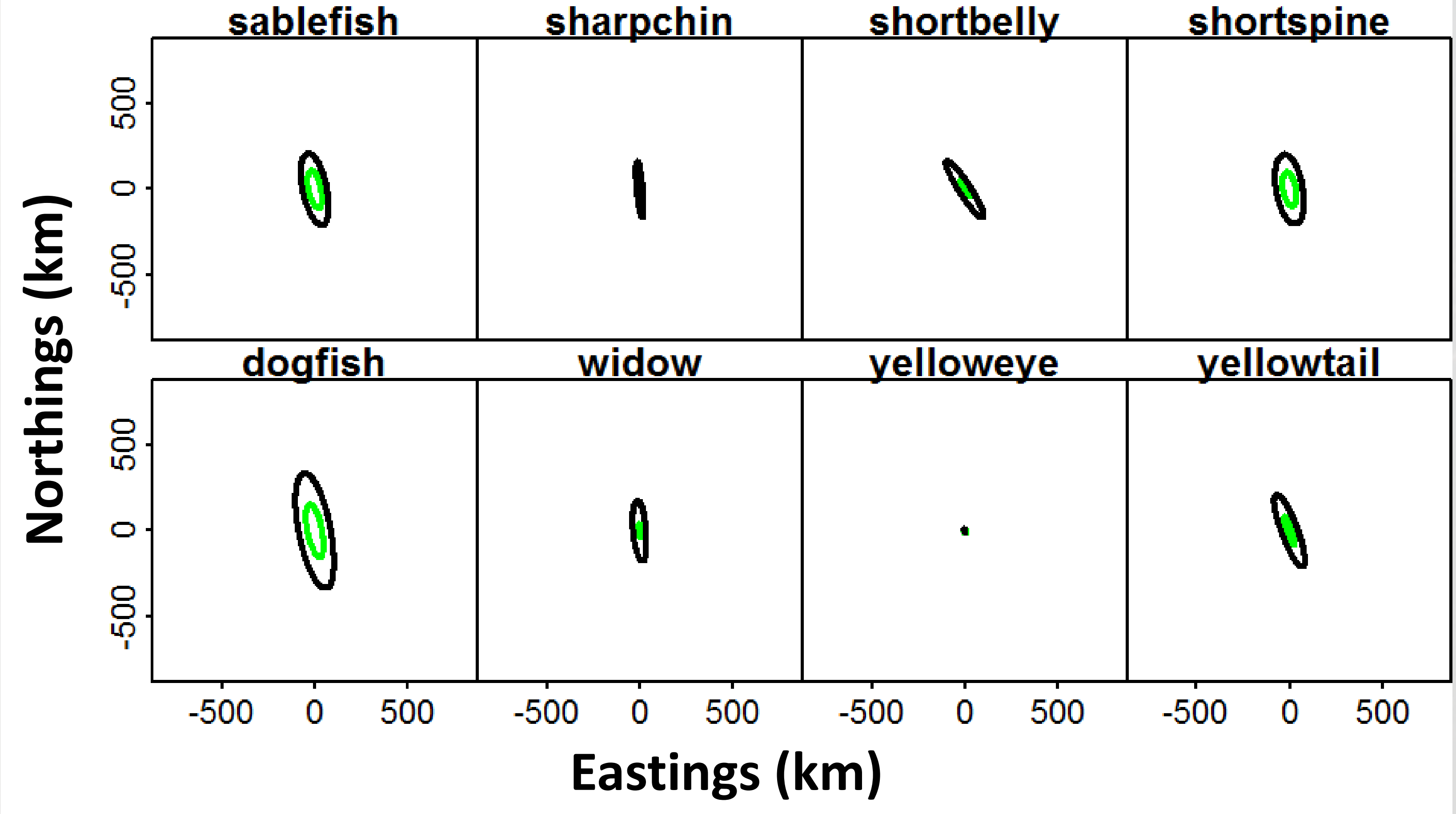


Advice: Inspect extrapolation footprint and knots

Geometric anisotropy

Decorrelation distance varies by direction

- **Probability of encounter**; **Positive catch rates**



Potential problems

Delta-generalized linear mixed model

Main difficulty

- Defining covariates X_xtp for every knot and year
  - Hurdle for analysts
  - Might be missing for some knots and/or years -> Imputation
  - Might lose variance when aggregating to know

Solutions

- Include mean and variance of samples for each knots/year
- Fill in missing values in consistent manner and check sensitivity
- Treat covariate as observation in multivariate model

Delta-generalized linear mixed model

Potential problems

1. Some combination of species and year has 0% or 100% encounter rate
   - If 100% encounter rate in year t, then $\beta_p(t) \rightarrow \infty$ and/or $\varepsilon_p(s,t) \rightarrow \infty$ for that year
   - If 0% encounter rate in year t, then $\beta_p(t) \rightarrow -\infty$ and/or $\varepsilon_p(s,t) \rightarrow -\infty$ and there's no information to estimate β_r or $\varepsilon_r(s,t)$ for that year t

use ObsModel[2]=3

Delta-generalized linear mixed model

Potential problems

- Some combination of species and year has 0% or 100% encounter rate

Solutions

1. If a few years with 100% encounter rate, try `ObsModel[2]=3`
   – indicates that VAST should check for species-years combinations with 100% encounter rates and fix those intercept for encounter probability to an extremely high value
2. If a few years with either 100% of 0% encounter rate, add temporal structure
   – Shrinks intercept and slope towards adjacent years
3. Four other special cases available

use ObsModel[2]=3

Delta-generalized linear mixed model

Potential problems

- Parameters hit bounds

Solutions

1. Simplify model
   - Check parameter estimates when turning off standard errors and newtonstep
     - Identify parameter at bounds
   - E.g., ` ln_H_input` sometimes gets to strange values
     - Turn off using `Data_Fn(…, Aniso=FALSE)