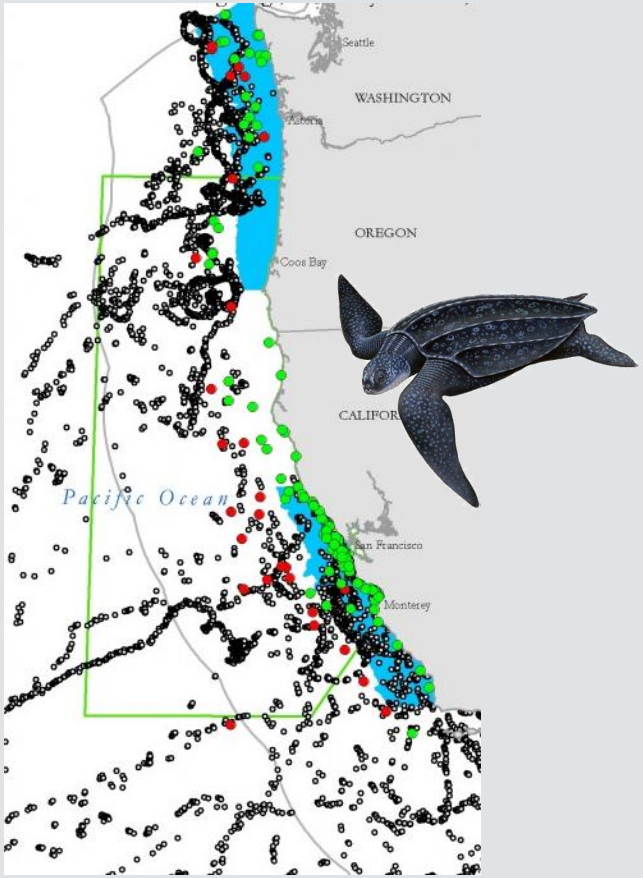
Predicting fisheries bycatch risk for dynamic spatial management

BRIAN STOCK, TOMO EGUCHI, ERIC WARD, JASON JANNOT, ERIC FORNEY, & BRICE SEMMENS

Static vs. dynamic management

Dynamic

Static



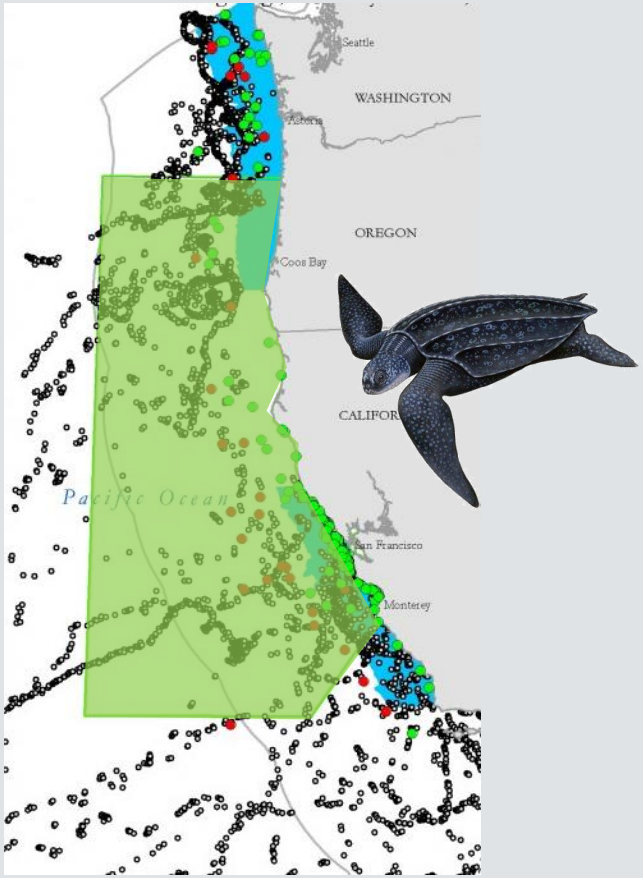
Oceana (2012)

Static vs. dynamic management

Dynamic

1. Effectively protected?
2. Huge loss of fishing area

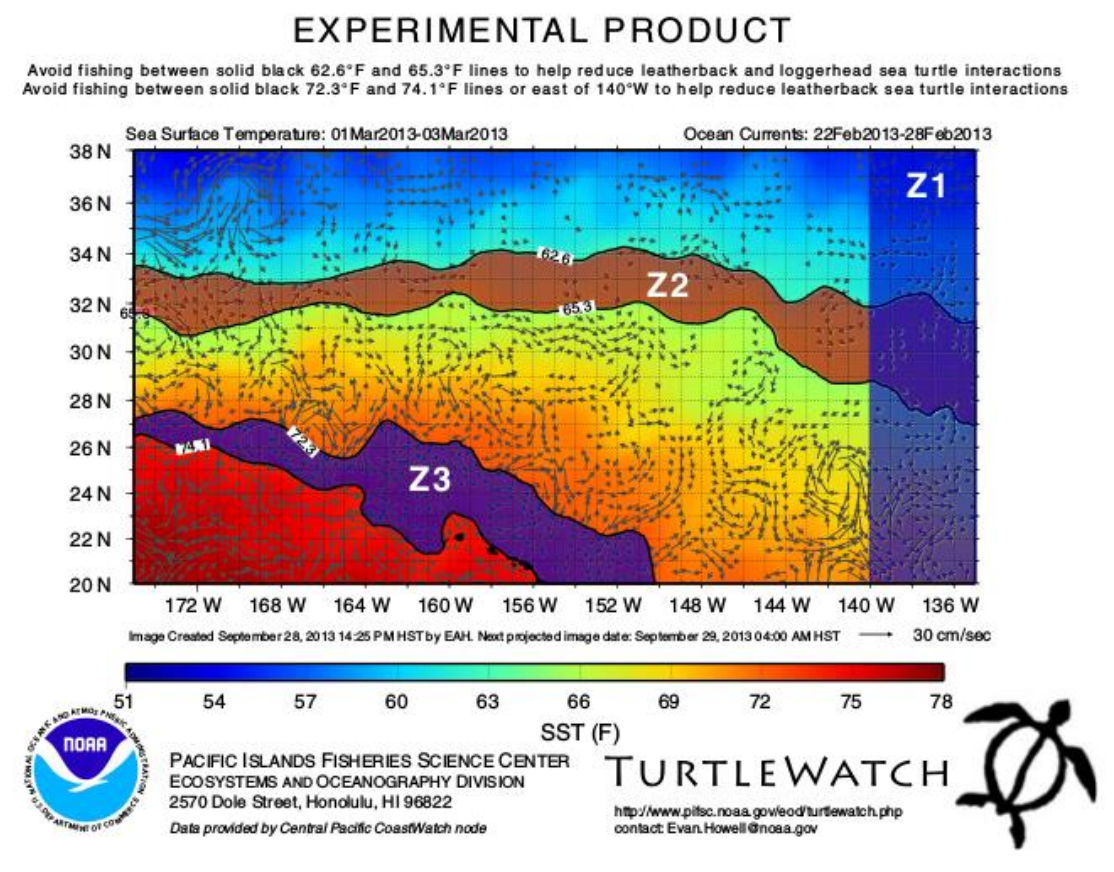
Static



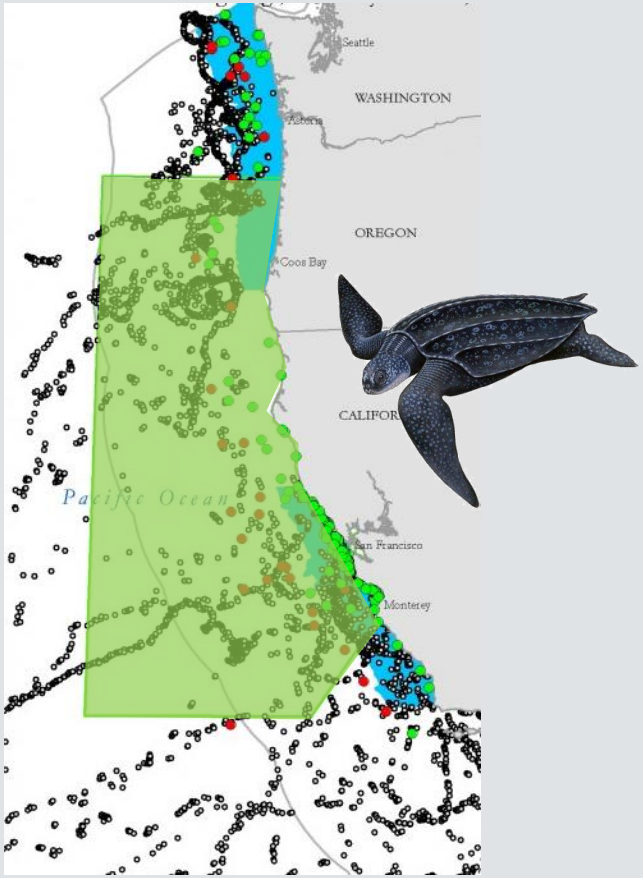
Oceana (2012)

Static vs. dynamic management

Dynamic



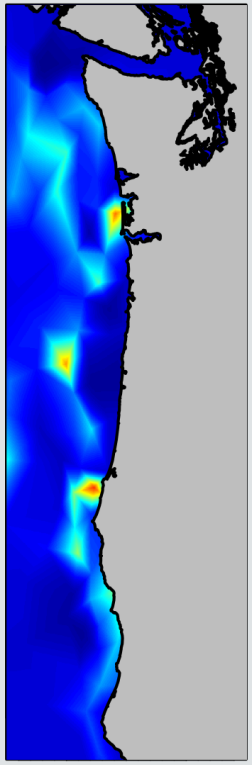
Static



Howell et al. (2008, 2015)

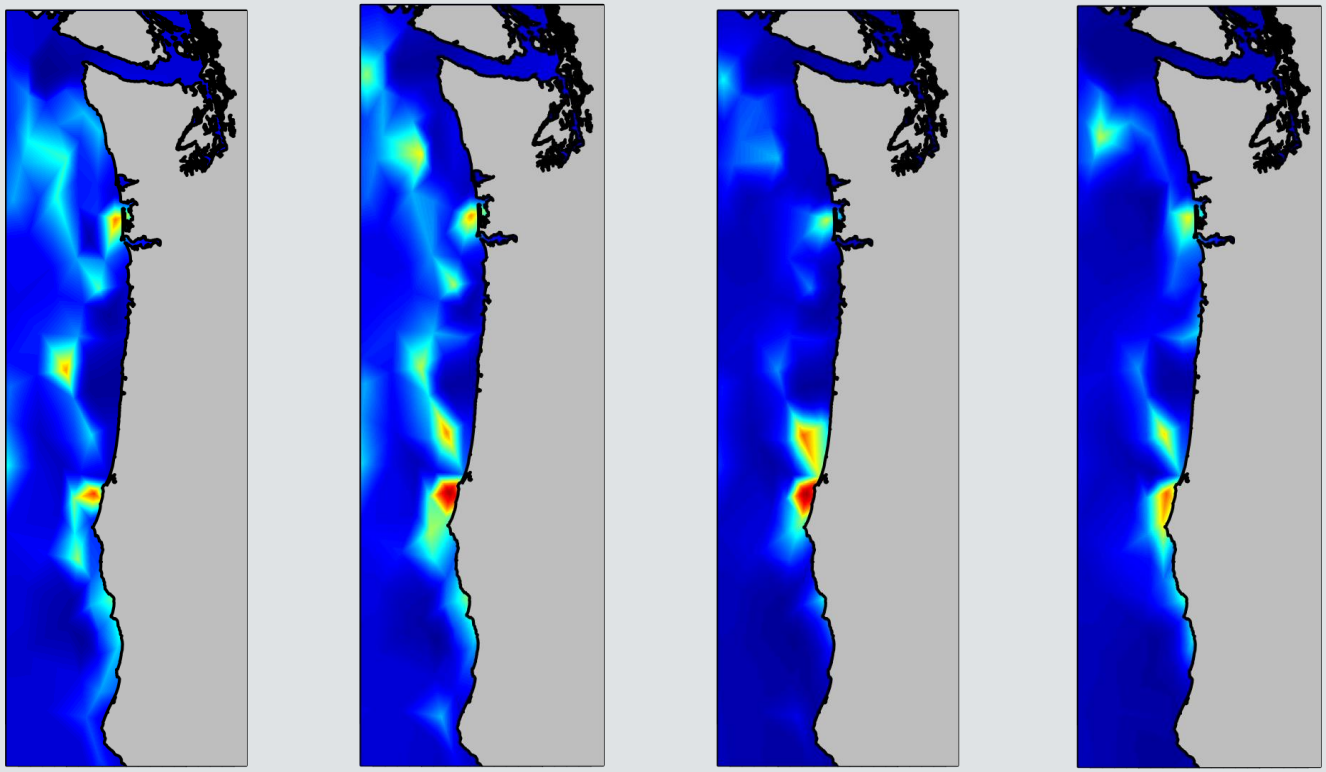
Tools for dynamic management

Need map of bycatch risk



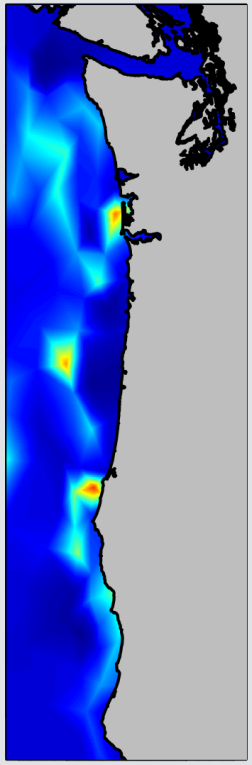
Tools for dynamic management

Need map of bycatch risk



Tools for dynamic management

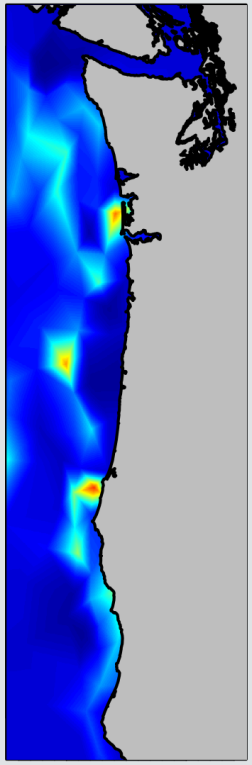
Need map of bycatch risk



- temperature
- depth
- substrate
- spatial field

Tools for dynamic management

Need map of bycatch risk



- temperature
- depth
- substrate
- spatial field

GLM

GAM

INLA

RF

Which spatial model is best?

GLM

GAM

INLA

RF

~ environmental predictors (temp, depth, substrate, …)

Problem:

spatial correlation in residuals (Prediction – Observed)

Research question

Dormann (2007)

Which spatial model is best?

GLM

GAM

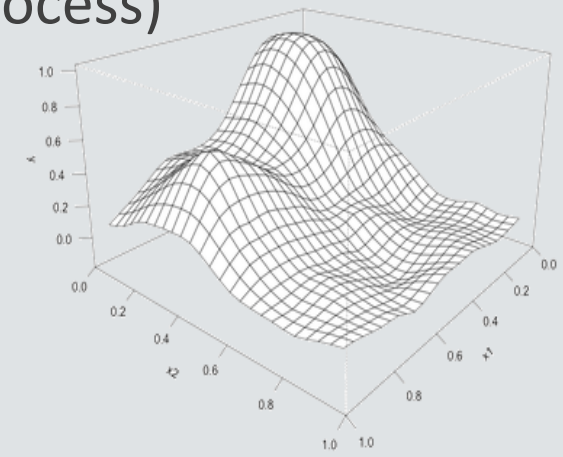
INLA

RF

~ environmental predictors

+ s(lat,lon)

- Common, simple approach
- Helps spatial correlation problem
- Purely statistical (no spatial process)



Research question

Wood (2006)

Which spatial model is best?

GLM

GAM

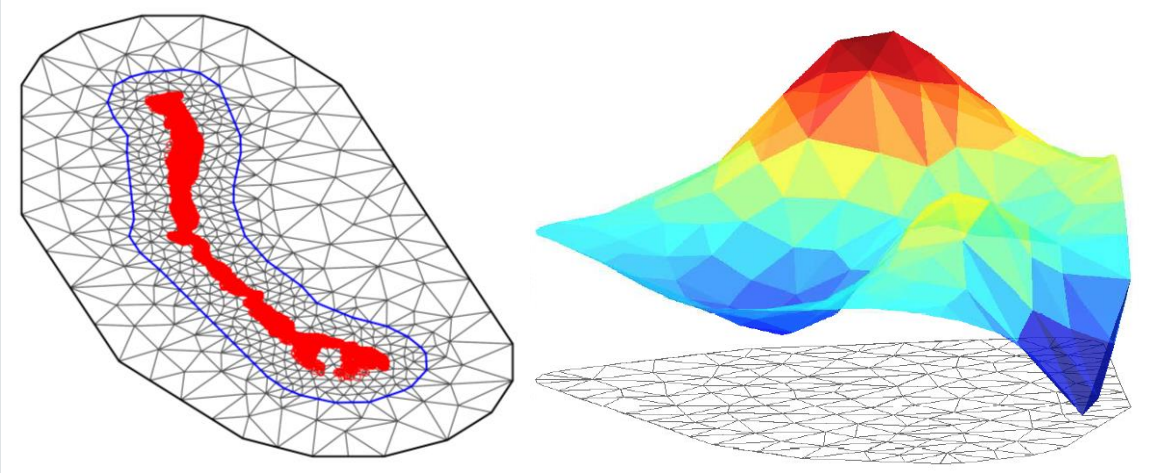
INLA

RF

~ environmental predictors

$+ MVN(0, \boldsymbol{\Sigma})$

- Models *covariance* as function of spatial locations



Research question

Lindgren, Rue, & Lindstrom (2011)

Which spatial model is best?

GLM

GAM

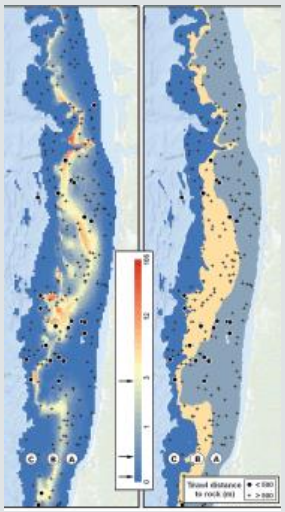
INLA

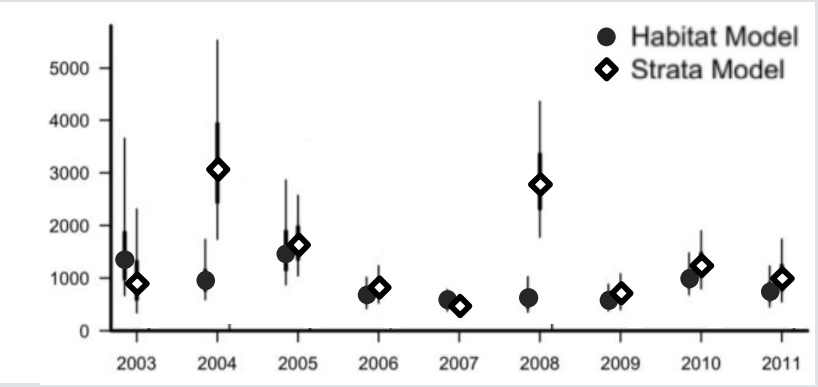
RF

~ environmental predictors

$+ MVN(0, \boldsymbol{\Sigma})$

- Models *covariance* as function of spatial locations
- Increasing adoption in fisheries





Research question

Shelton et al. (2014)

Which spatial model is best?

GLM

GAM

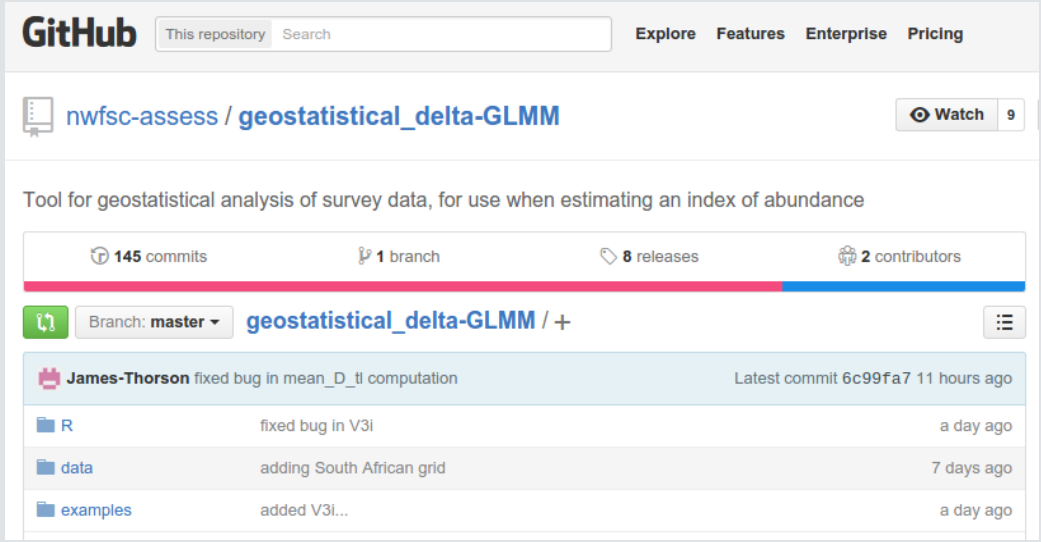
INLA

RF

~ environmental predictors

$+ MVN(0, \boldsymbol{\Sigma})$

- Models *covariance* as function of spatial locations
- Increasing adoption in fisheries



Research question

Thorson et al. (2015)

Which spatial model is best?

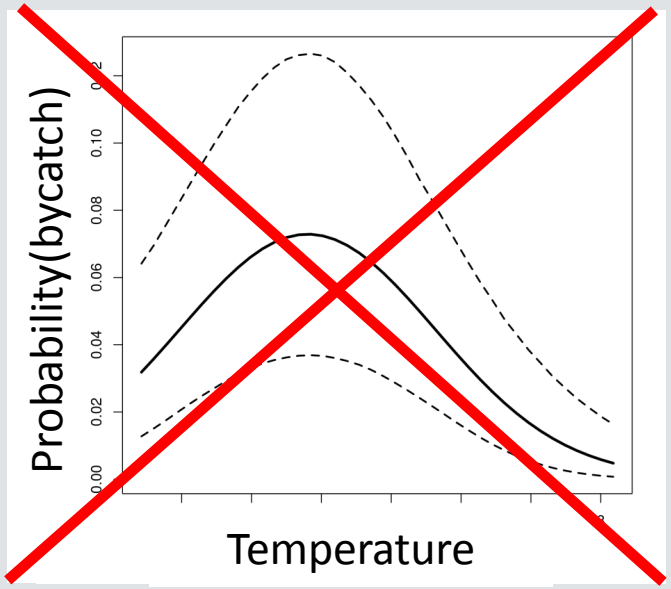
GLM

GAM

INLA

RF

- Machine learning, designed for prediction
- Black box
  - Predictor-bycatch relationships not modeled
  - No spatial field (add LAT, LON)



Research question

Breiman (2001)

Fisheries observer data

West Coast Groundfish Trawl

- 2002-2013
- 55,835 tows

Hawaii Longline

- 1994-2014
- 16,714 sets (swordfish only)

Research question

Does the answer depend on species?

Habitat:	Benthic	Benthic	Benthic
Movement:	Med	Low	Low
Bycatch Rate:	29%	18%	0.3%

Habitat:	Pelagic	Pelagic	Pelagic
Movement:	High	High	High
Bycatch Rate:	96%	1.4%	0.7%

Research question

Delta model

Binomial — Pr(some bycatch)

Positive — E(bycatch | some bycatch)

Binomial x Positive — E(bycatch)

How to gauge model performance?

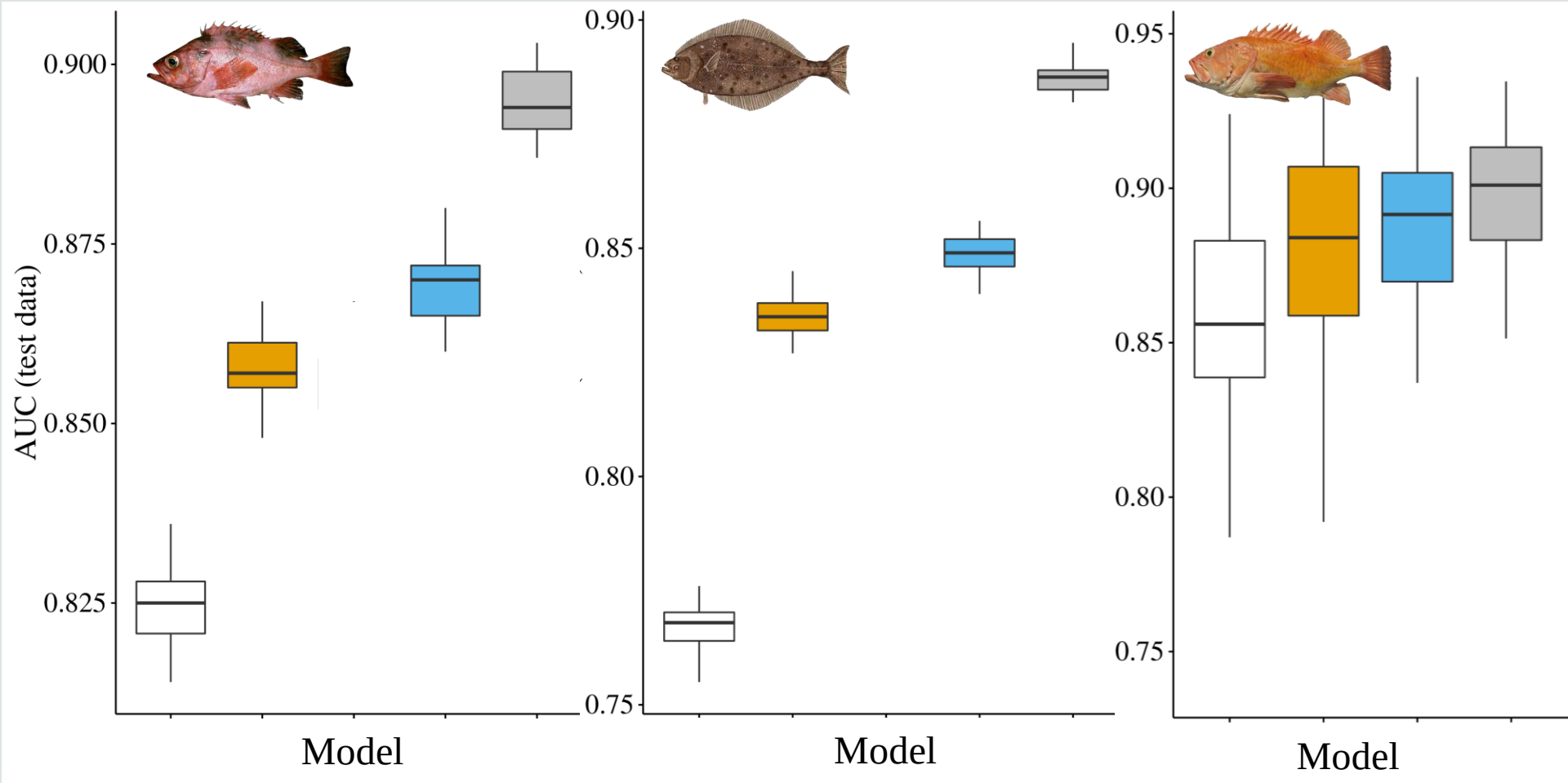
Goal: prediction

5-fold cross validation repeated 10x

Binomial — ROC curve (AUC)

Positive — R^2 (pred-obs), RMSE

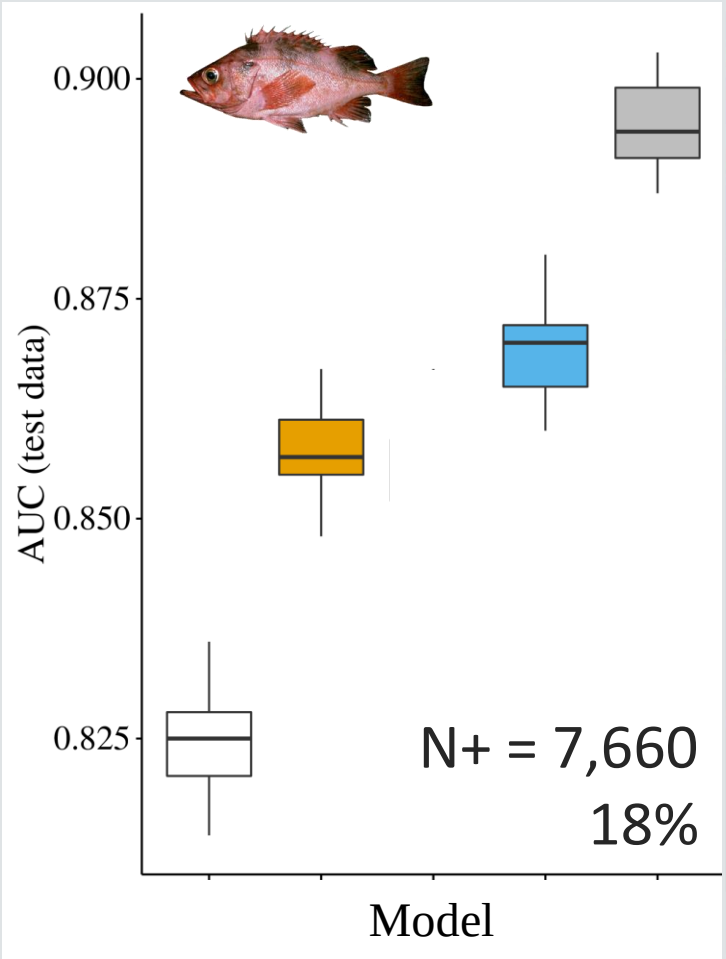
Generally: GLM < GAM < INLA < RF

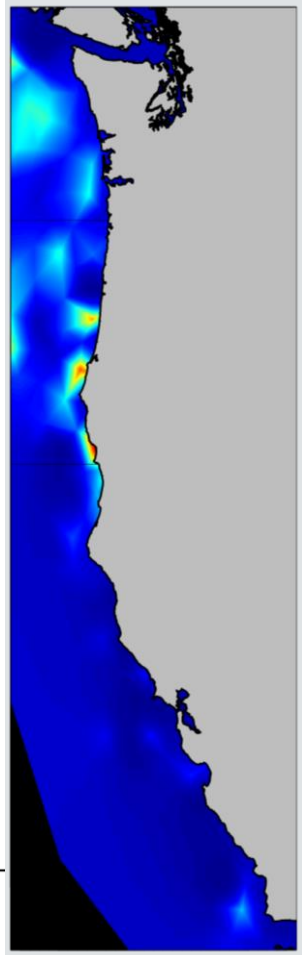


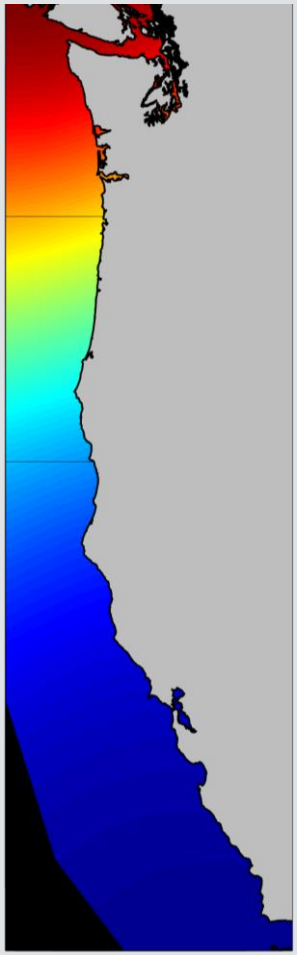
Results | Binomial

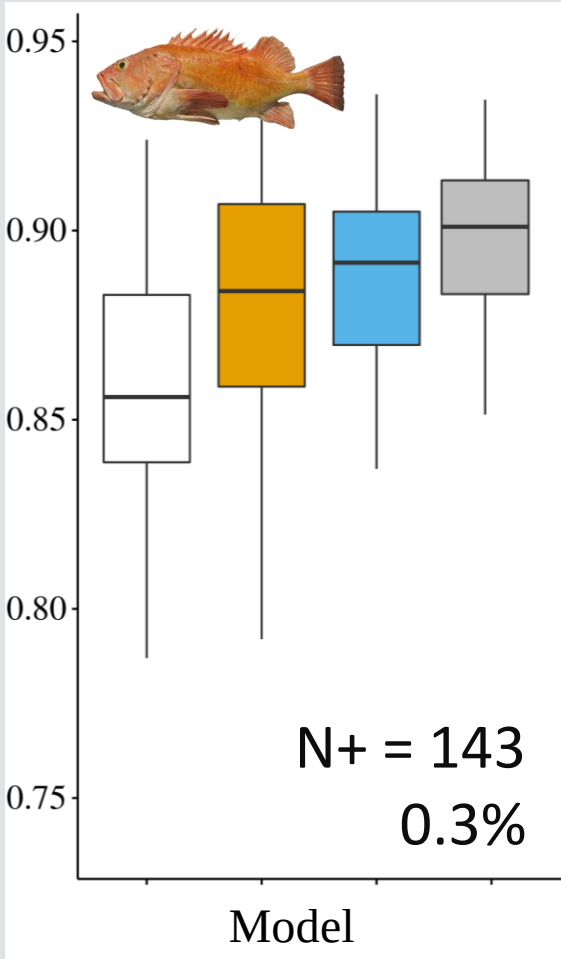
Generally: GLM < GAM < INLA < RF

Less clear for rarer species



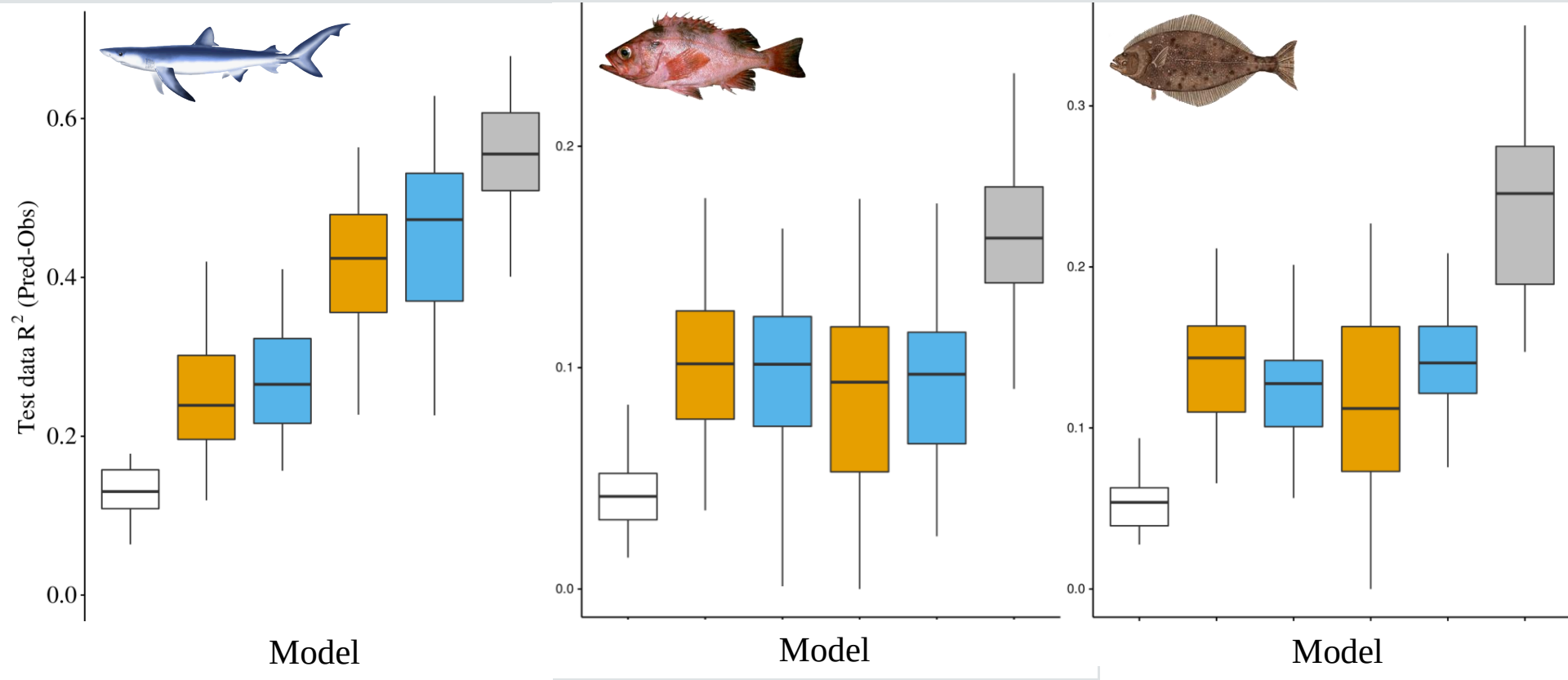






Results Binomial

Generally: GLM < GAM < INLA < RF



Results Positive

Conclusion

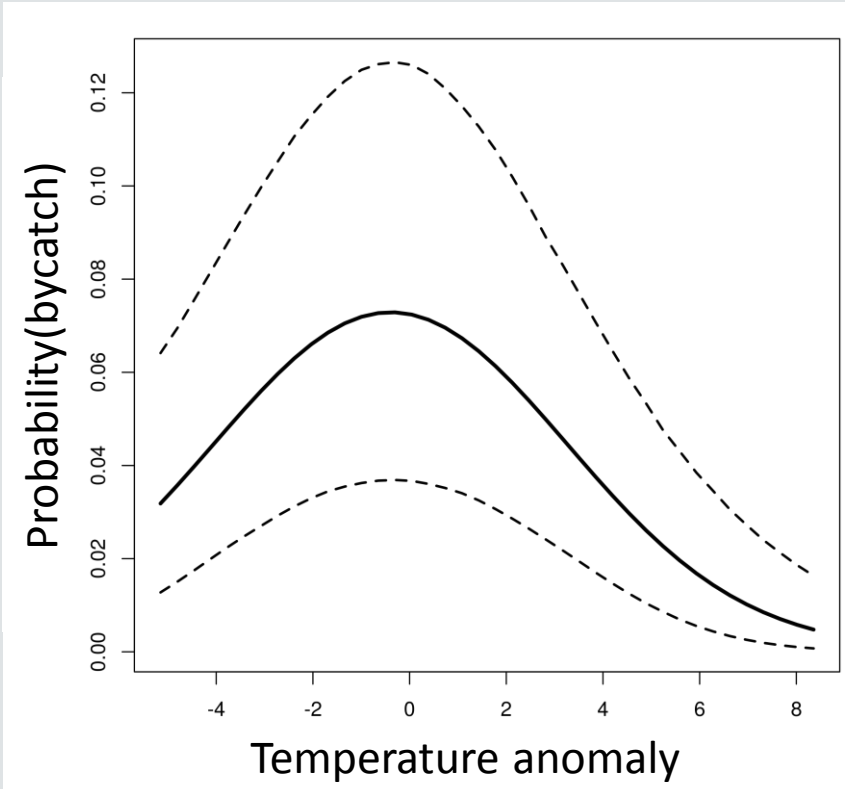
If the goal is purely *prediction*:

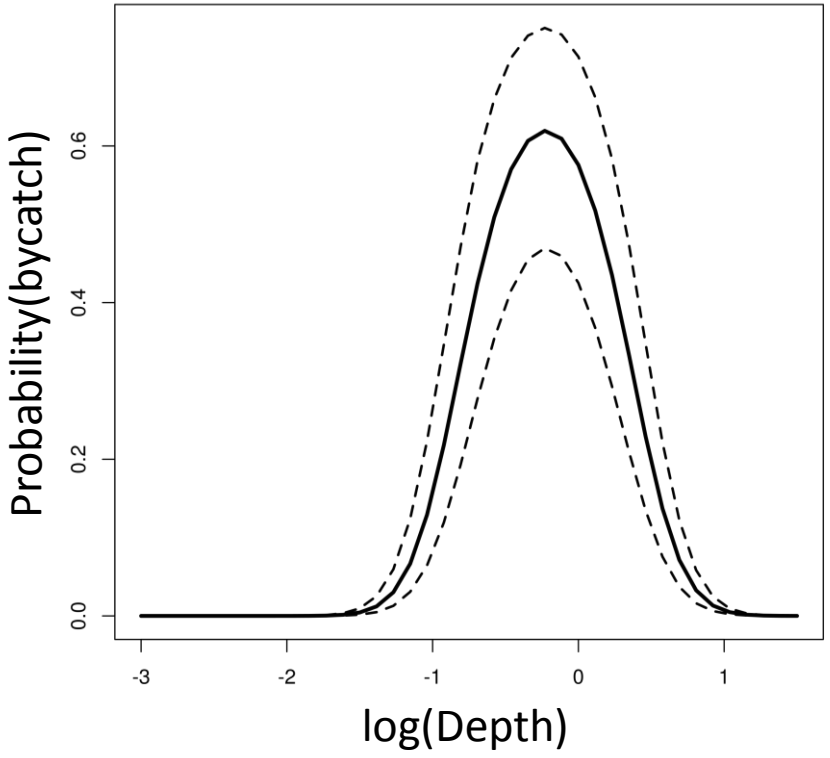
GLM < GAM < INLA < RF

…but if we care about *inference on processes* affecting bycatch:

Marginal posteriors

INLA

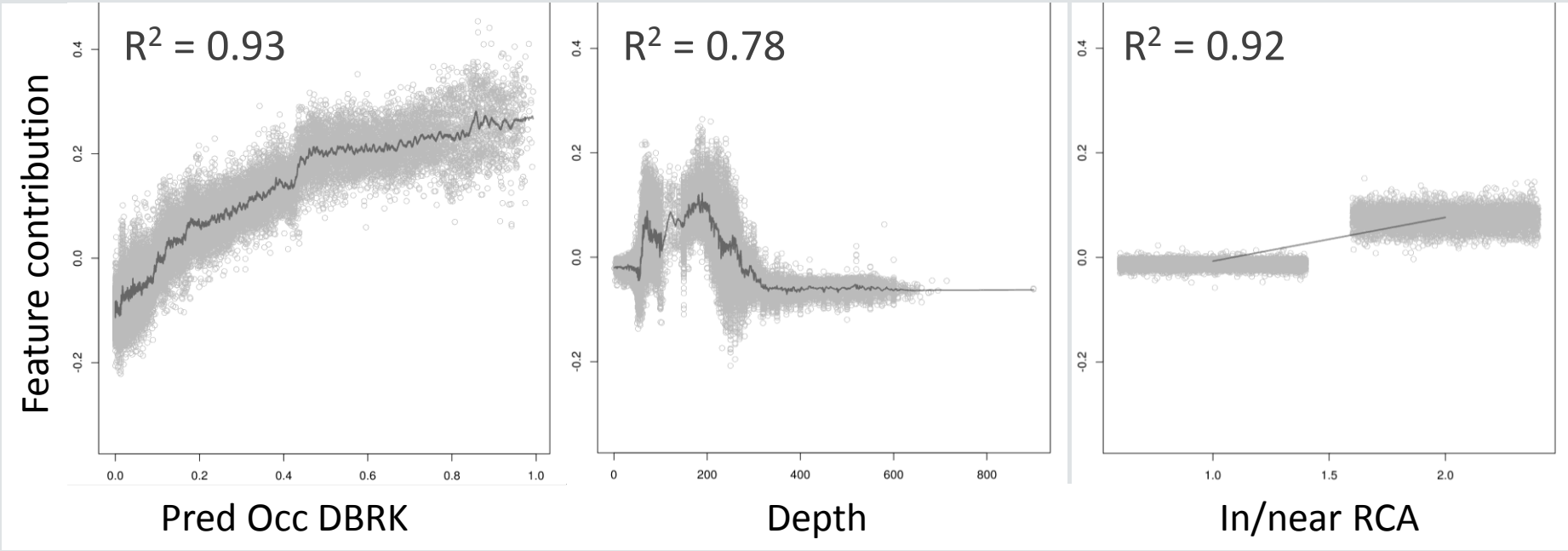




Can random forests do the same?

RF

Are random forests really “black boxes”?



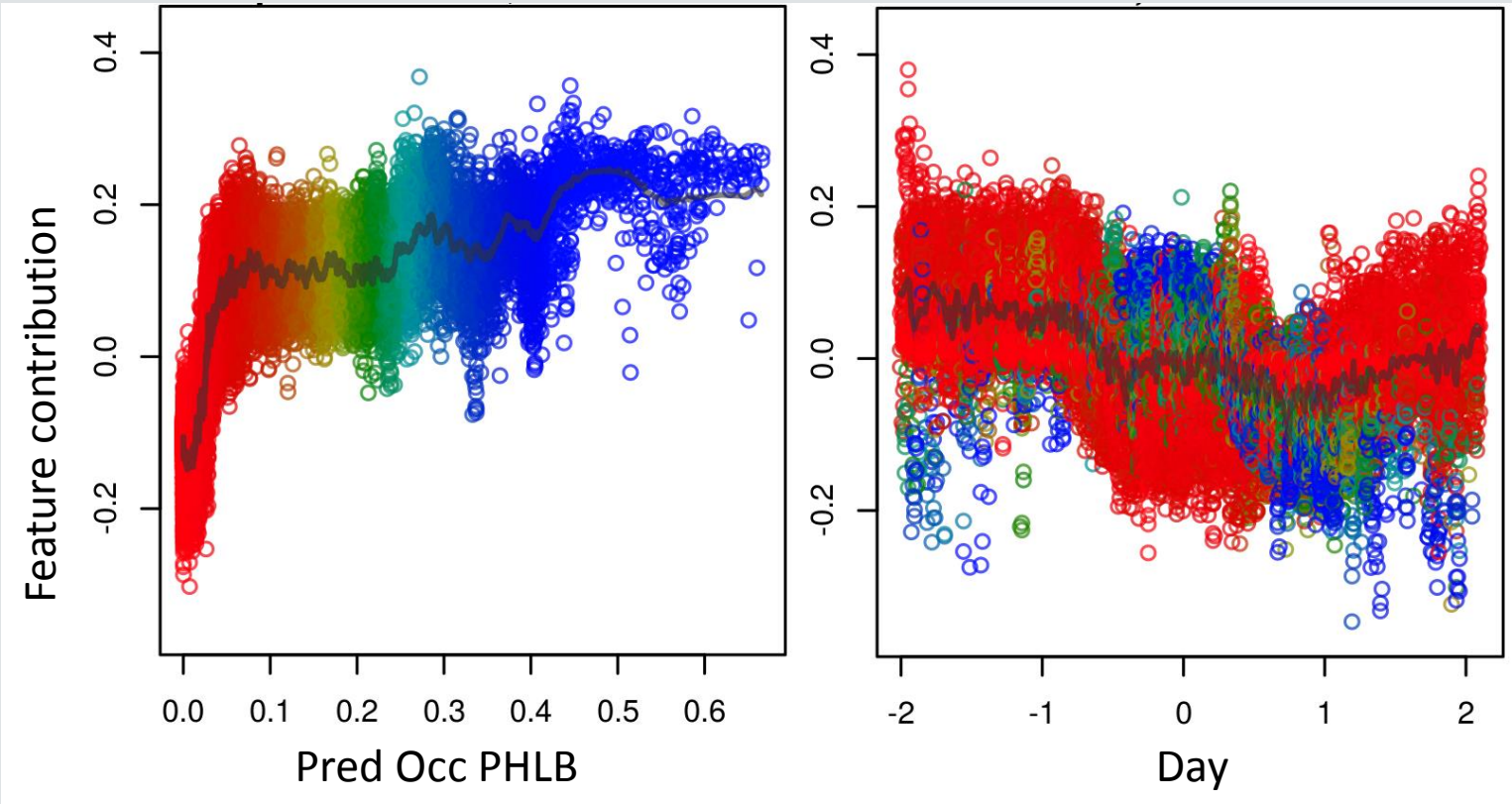
Results

Palczewksa (2013), Welling (2016)

Can random forests do *better*?

RF

Identifying covariate interactions

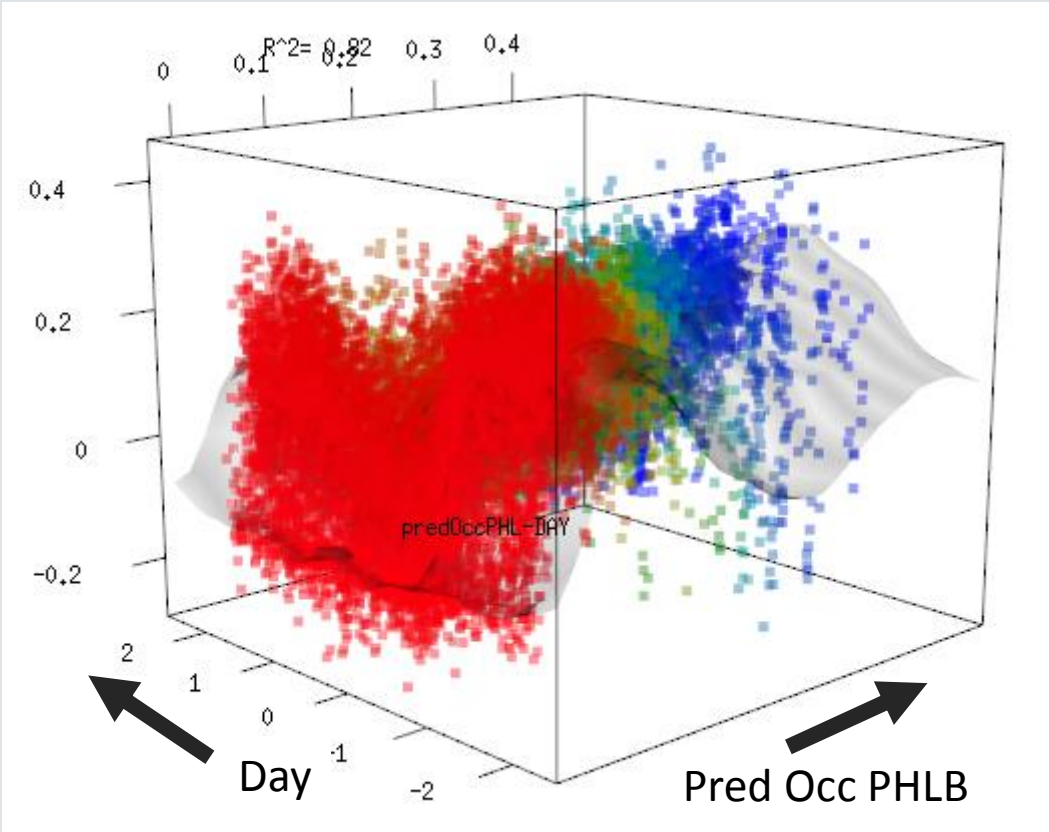


Results

Can random forests do *better*?

RF

Identifying covariate interactions



Results

Discussion

RF

+ Better at prediction

+ More complex covariate relationships (incl. interactions)

+ Much quicker to set up and run (~2 min vs. 5-15 hours)

+ Not just a “black box”?

INLA

+ Statistical inference, marginal posteriors for covariate effects

+ Ability to include observation error

Thank you!

XSEDE
Extreme Science and Engineering Discovery Environment

SIO
- Brice Semmens

SWFSC
- Tomo Eguchi

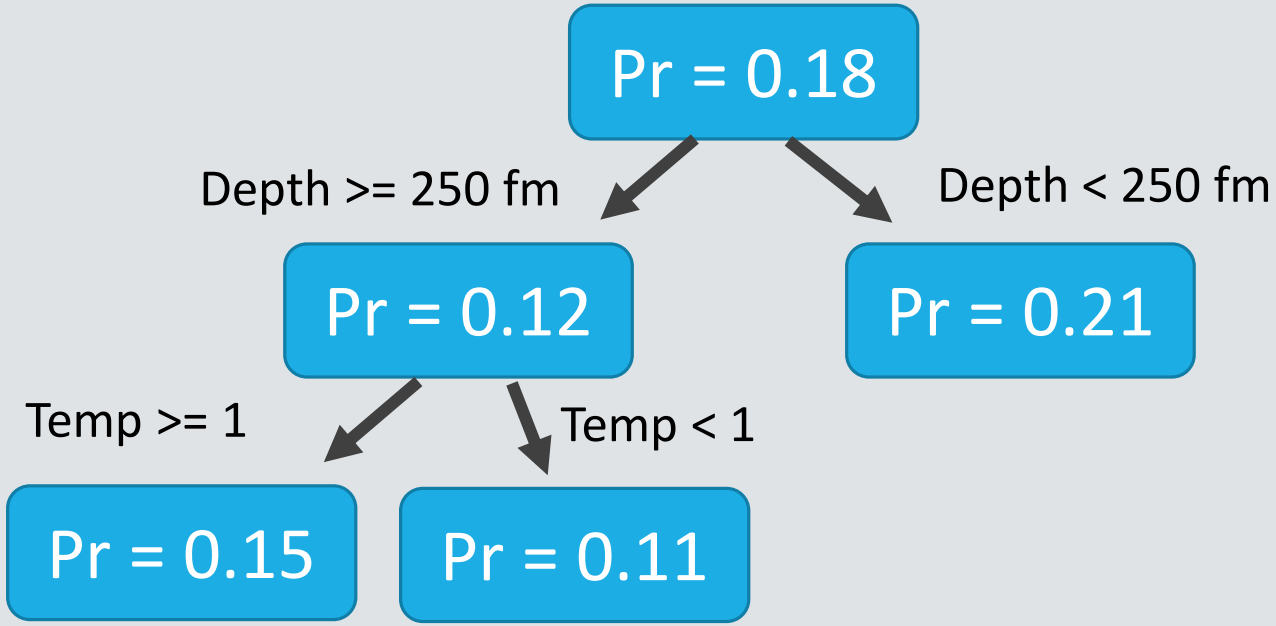
NWFSC
- Eric Ward
- Essential Fish Habitat (Blake Feist)
- West Coast Groundfish Observer Program (Jason Jannot)

PIFSC
- Hawaii Longline Observer Program (Eric Forney)

Can random forests do the same?

RF

What is a “feature contribution”??

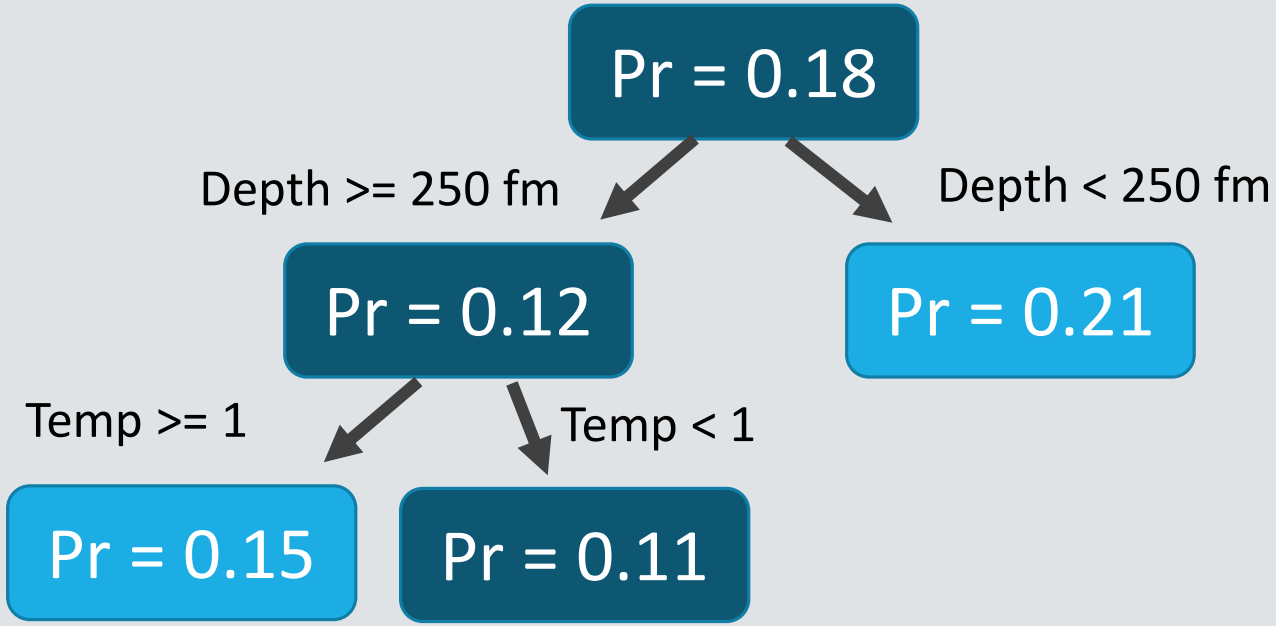


Palczewksa (2013), Welling (2016)

Can random forests do the same?

RF

What is a “feature contribution”??

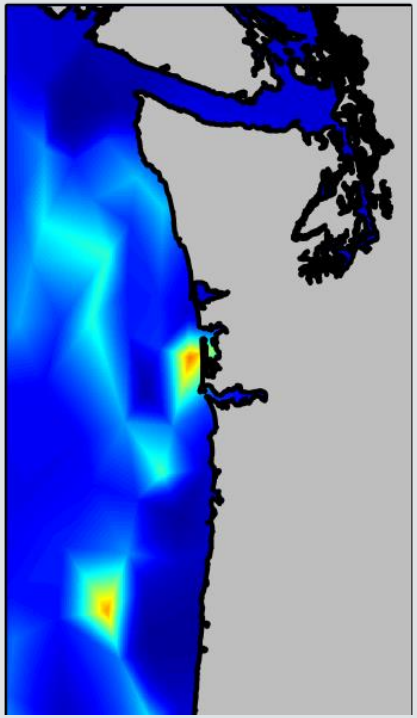


$\text{Prediction}_i = 0.11 = 0.18 - 0.06\ (\text{Depth}) - 0.01\ (\text{Temp})$

Palczewksa (2013), Welling (2016)

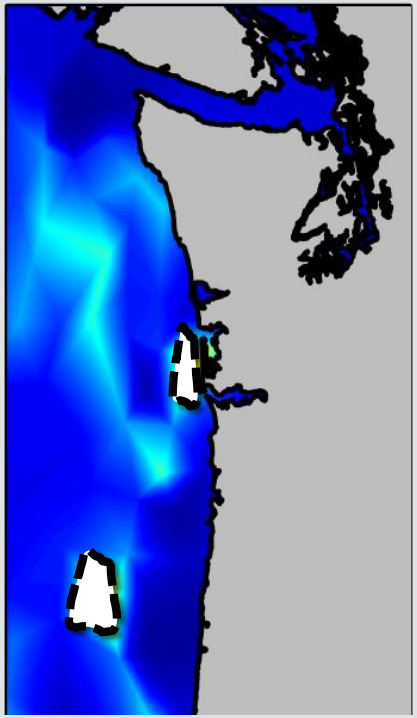
Generally: GLM < GAM < INLA < RF

Remove X% of fishing effort with highest bycatch risk



Generally: GLM < GAM < INLA < RF

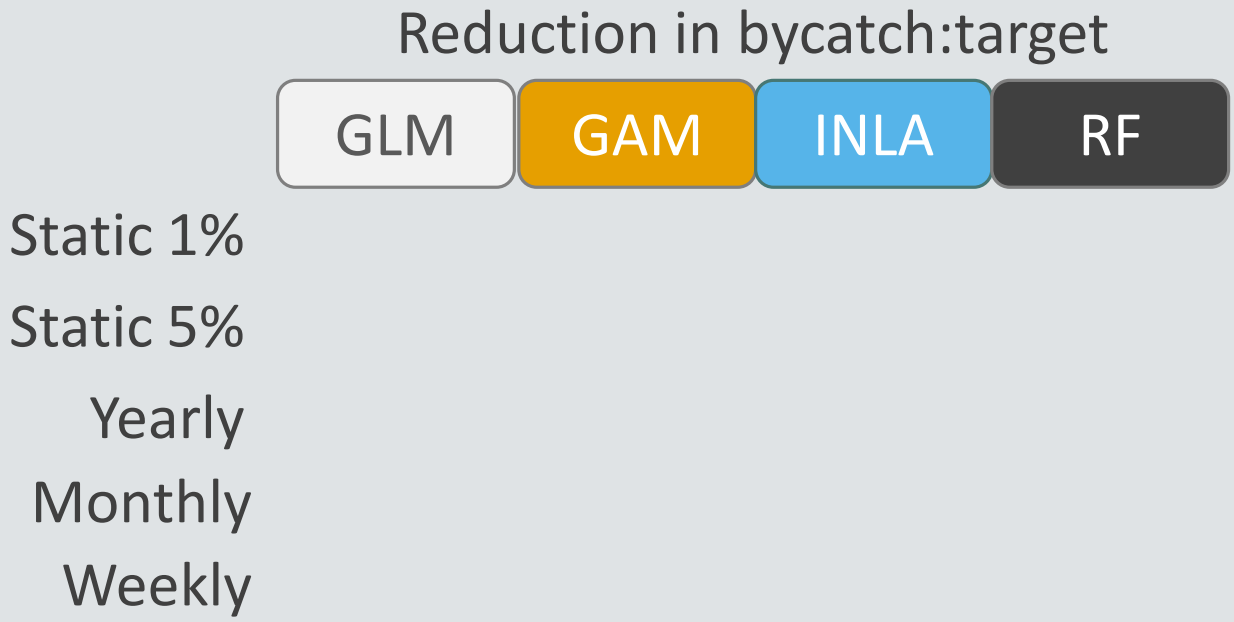
Remove X% of fishing effort with highest bycatch risk



Bycatch removed	Fishing removed 1%	Fishing removed 5%
GAM	6.2%	23.6%
INLA	7.3%	29.4%
RF	5.7%	32.7%

To do: dynamic vs. static management

1. Consider bycatch:target ratio
2. Simulate management strategy performance

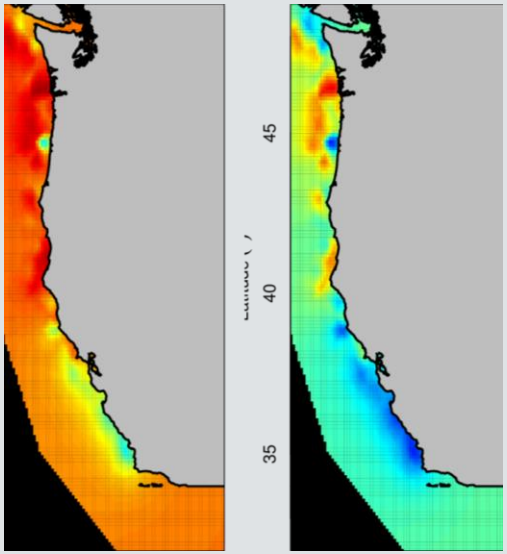


Future work

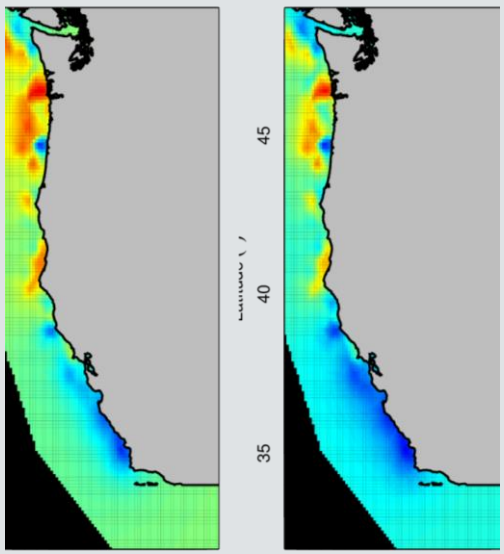
Research questions

2. What *type* of spatial model best predicts bycatch?

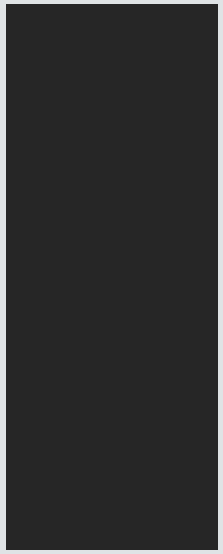
NULL



GAM
INLA-SPDE



GAM
INLA-SPDE



Random Forest

West Coast Groundfish covariates

Binomial

Positive

$\sim$ sst + sst^2 +

depth + $depth^2$ +

distance to rocky substrate +

size of rocky patch +

in Rockfish Conservation Area +

predicted occurrence (survey) +

day of year +

spatial field

Shelton et al. (2014)

Hawaii Longline covariates

Binomial

Positive

$\sim$ sst + sst^2 +

day of year +

spatial field

Shelton et al. (2014)