Pacific Community
Communauté du Pacifique

Accounting for the effects of oceanography on catchability and recruitment in basin-wide standardized indices of abundance for Pacific yellowfin and bigeye tuna

Laura Tremblay-Boyer, Samuel McKechnie, Graham Pilling and John Hampton

CAPAM spatio-temporal modelling workshop
February 26th to March 2nd
La Jolla, United States

This talk:

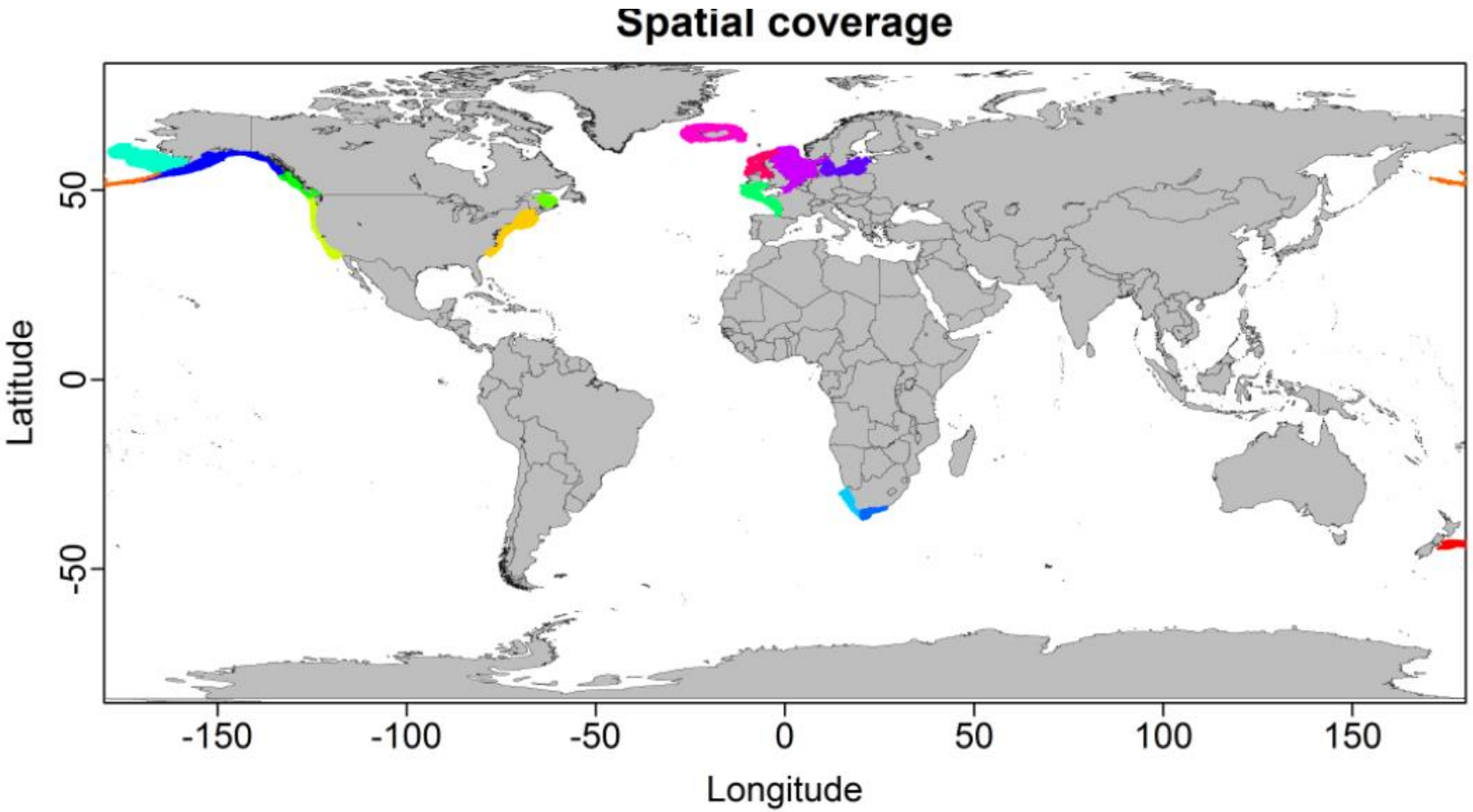
- CPUE indices in WCPO assessments:
  What we are hoping to do with spatio-temporal approaches

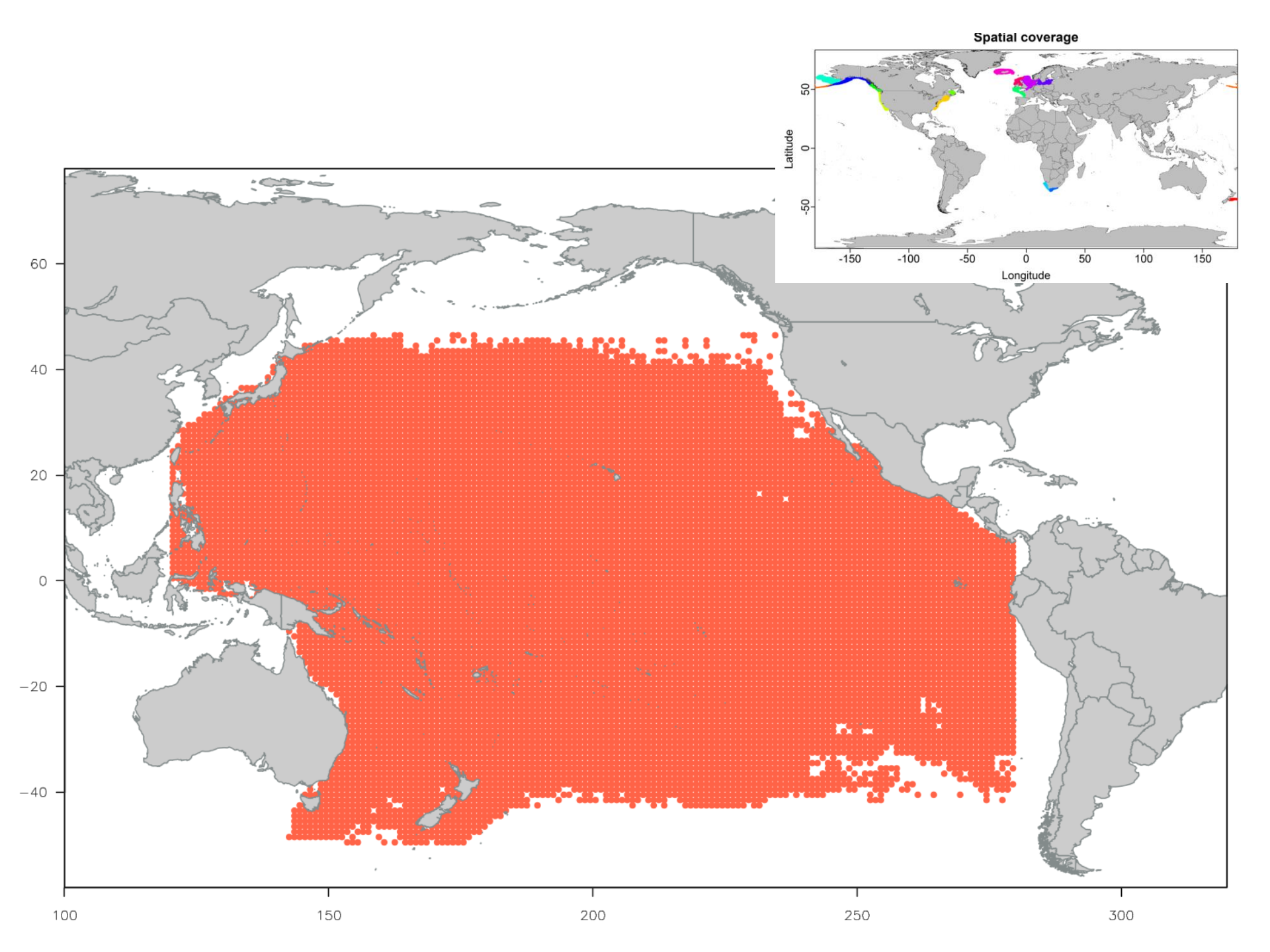
- Lessons learnt: applying VAST to a large, imbalanced dataset

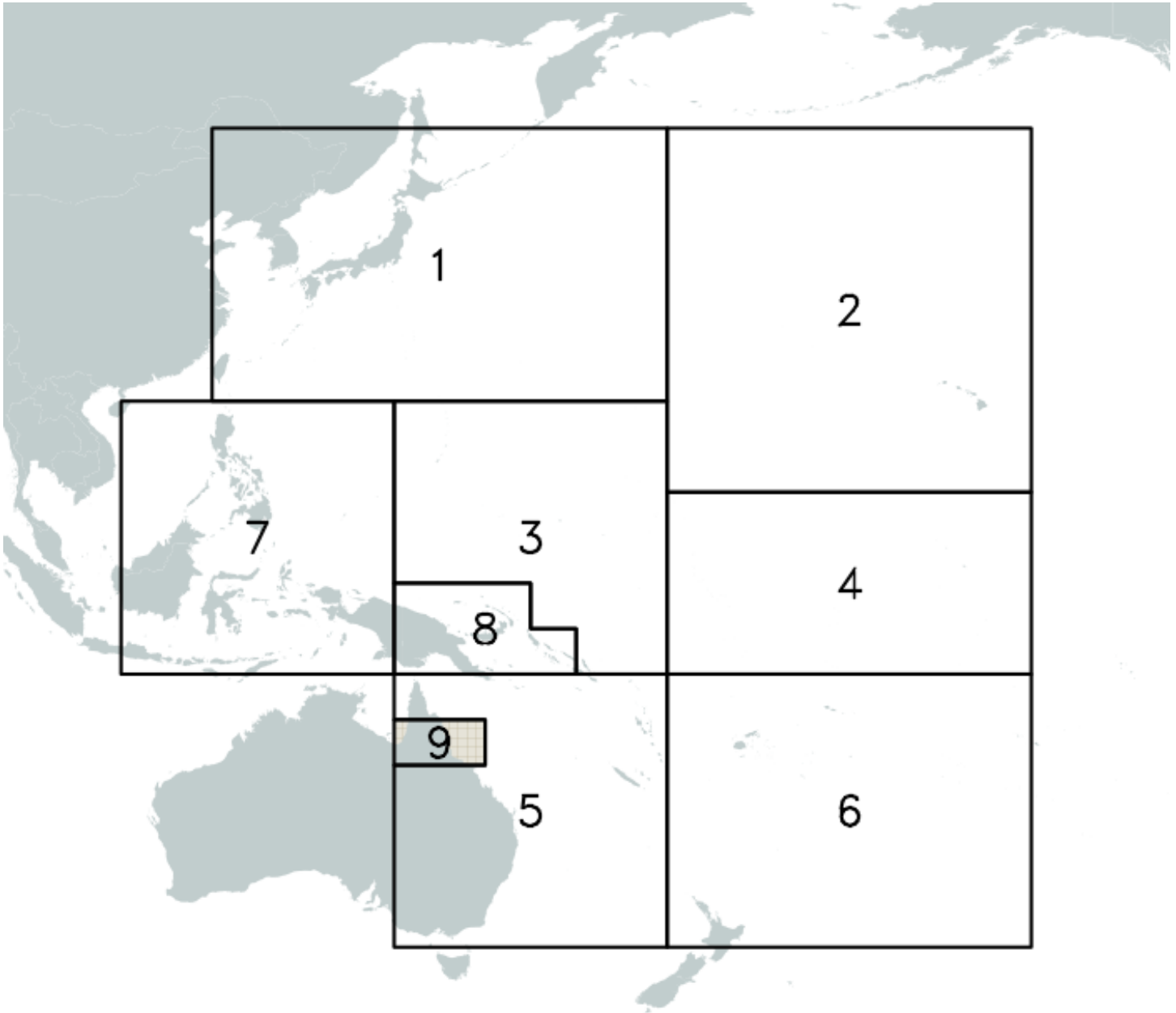
In practice:

- adding a non-linear covariate to a VAST model
- diagnostics

- Abundance vs. catchability covariates: early results


Spatial coverage
Latitude
50
0
-50
-150
-100
-50
0
50
100
150
Longitude

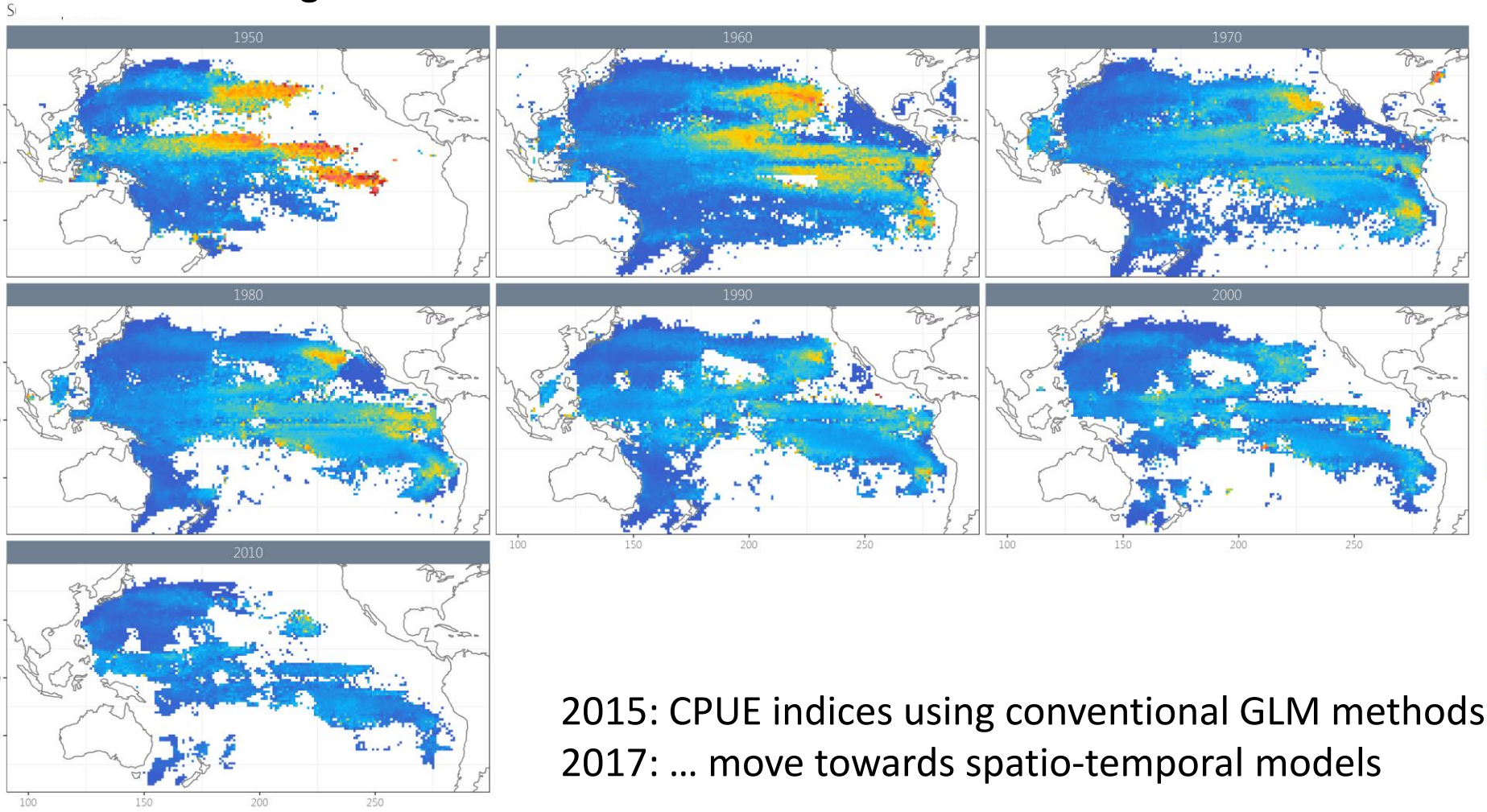

Spatial coverage
Latitude
Longitude
50
0
-50
-150
-100
-50
0
50
100
150
60
40
20
0
−20
−40
100
150
200
250
300


1
2
7
3
4
8
9
5
6

The dataset:

Logsheet longline data for domestic and distant-water fleets in the Pacific 1952-2015

~ 11 million longline sets



2015: CPUE indices using conventional GLM methods

2017: … move towards spatio-temporal models

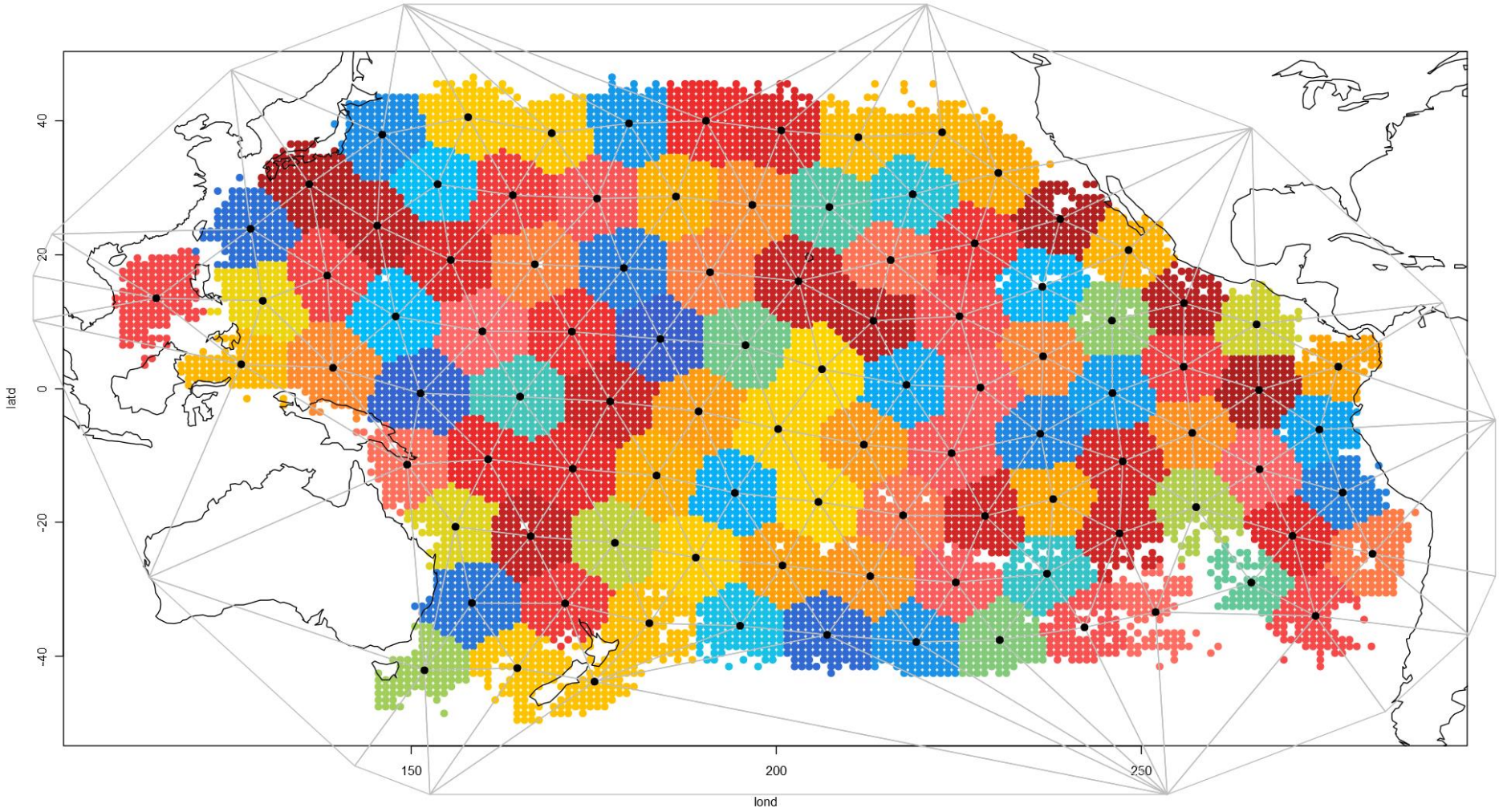
from resp ~ YrQtr + … + cell (*time)

to

resp ~ YrQtr + … + f(knot_i) + f(knot_i,YrQtr)

[…] can be any other covariate

Error distribution: delta-log-normal

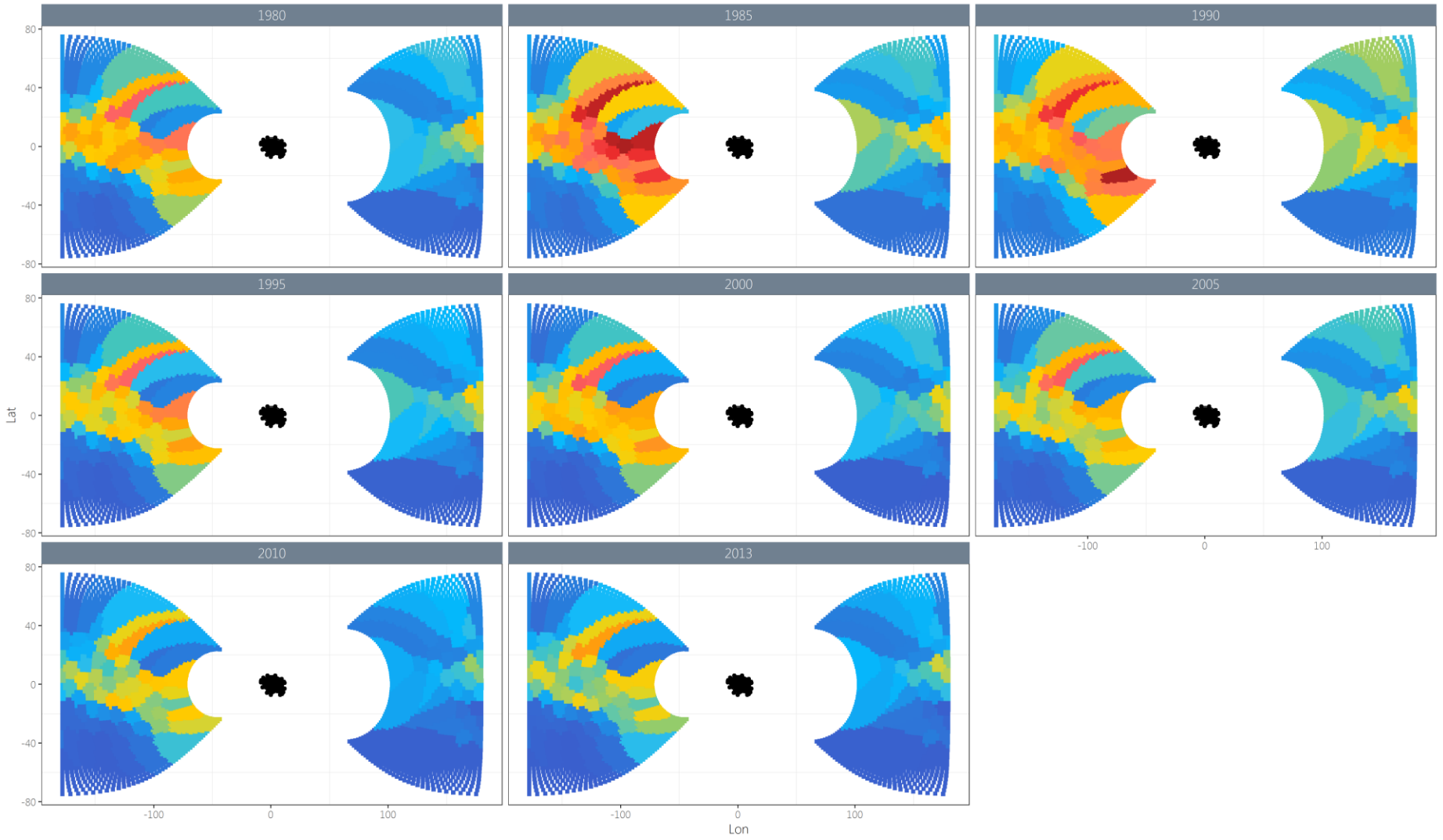


Current status:

From TMB/INLA approach developed by Jim Thorson (SpatialDeltaGLMM/VAST)

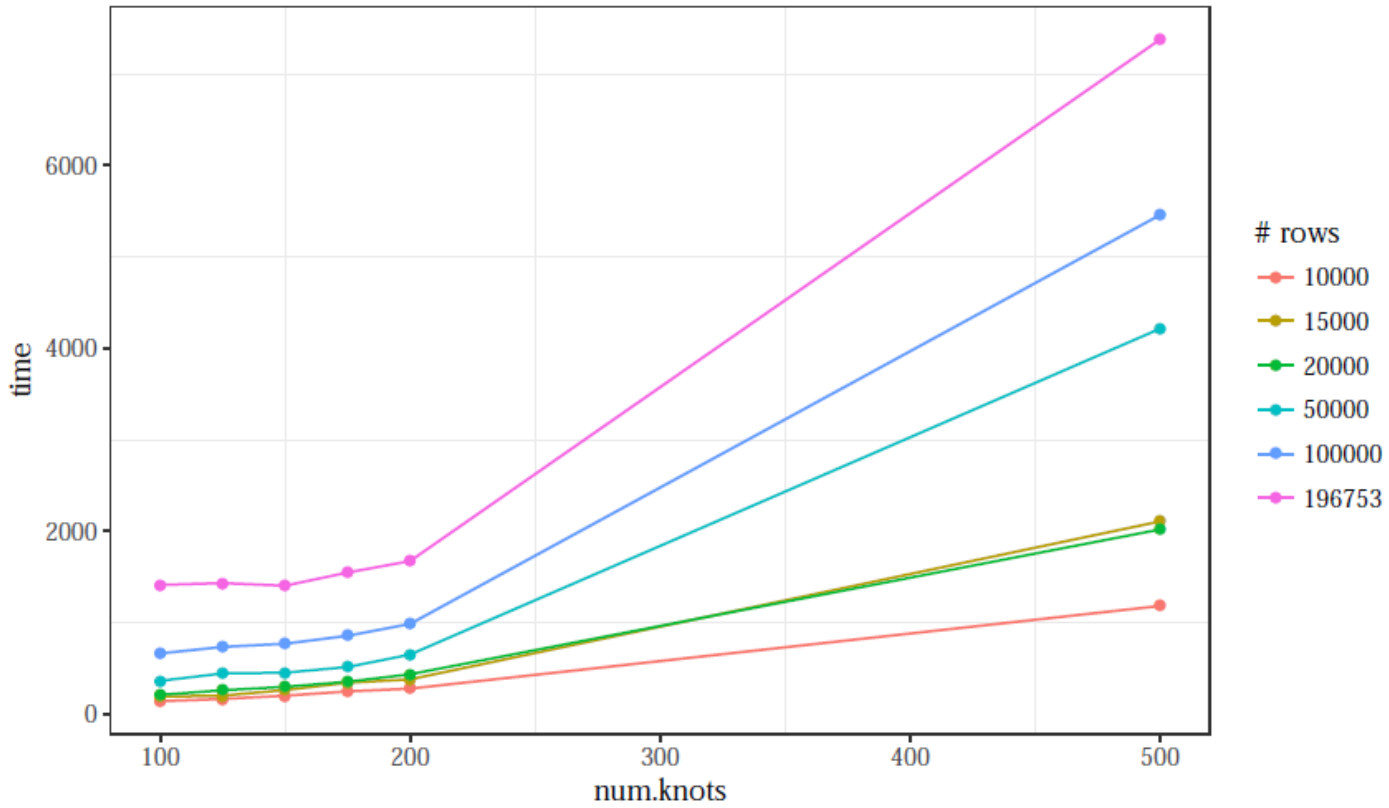
- Area much larger than previous implementations of VAST, so re-projection of coordinates using Two point equidistant projection, but individual cell area calculated by conversion to eastings/northings

- Subsampling of data by knot or by flag
- Anisotropic mesh
- Model structure:
  YrQtr + geostatistical surface with time interaction + s(oceanography covariate) (abundance or catchability)

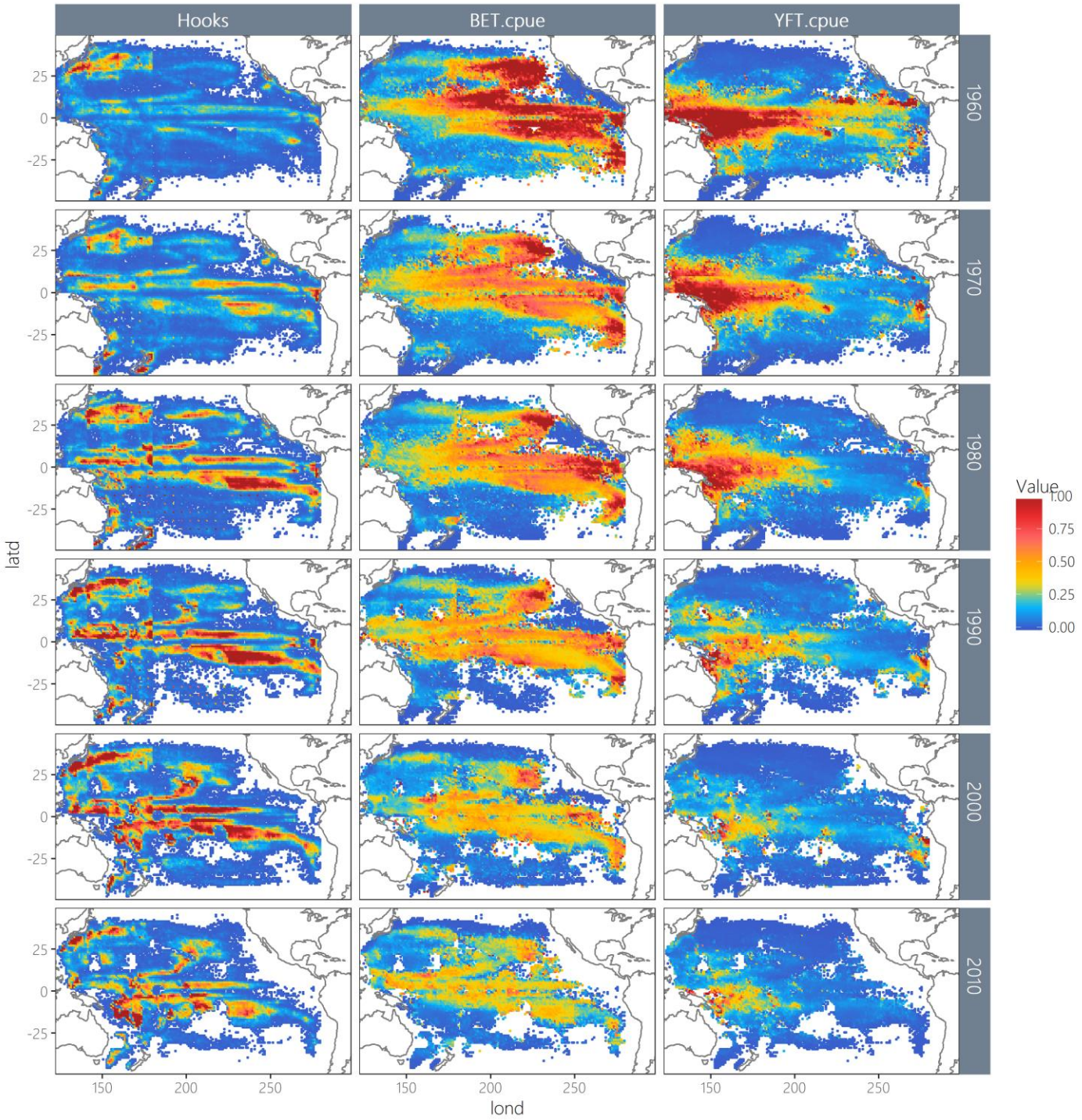

1980
1985
1990
1995
2000
2005
2010
2013
Lat
Lon
Pred
1.5
1.0
0.5
80
40
0
-40
-80
-100
0
100

Challenges for application over the WCPO:

- Computationally intensive



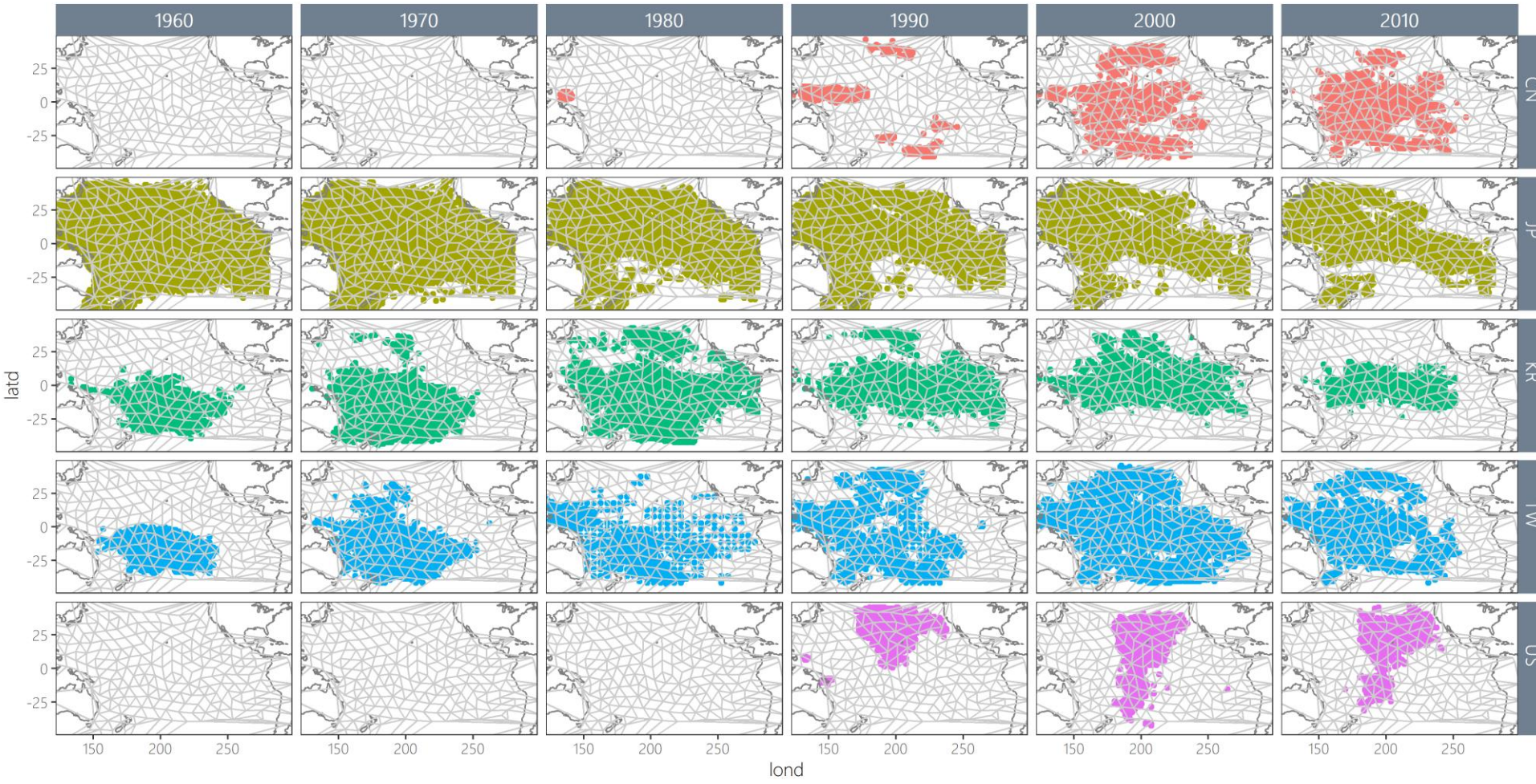
- Subsampling scheme? Flag vs. knot


Hooks
BET.cpue
YFT.cpue
1960
1970
1980
1990
2000
2010
latd
lond
25
0
-25
150
200
250
Value
1.00
0.75
0.50
0.25
0.00

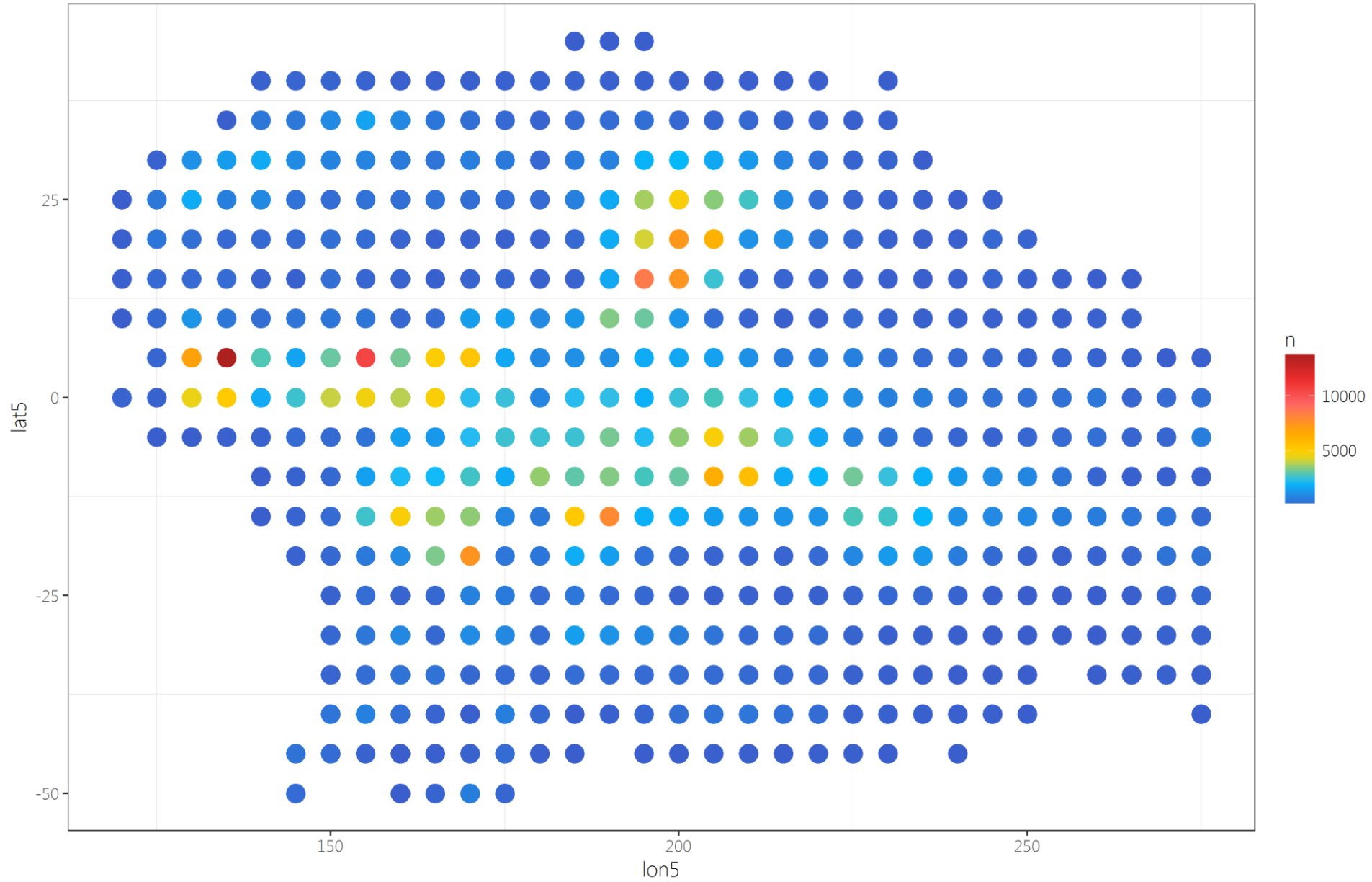
Knots sampled unevenly over time

Knots sampled unevenly by fleets

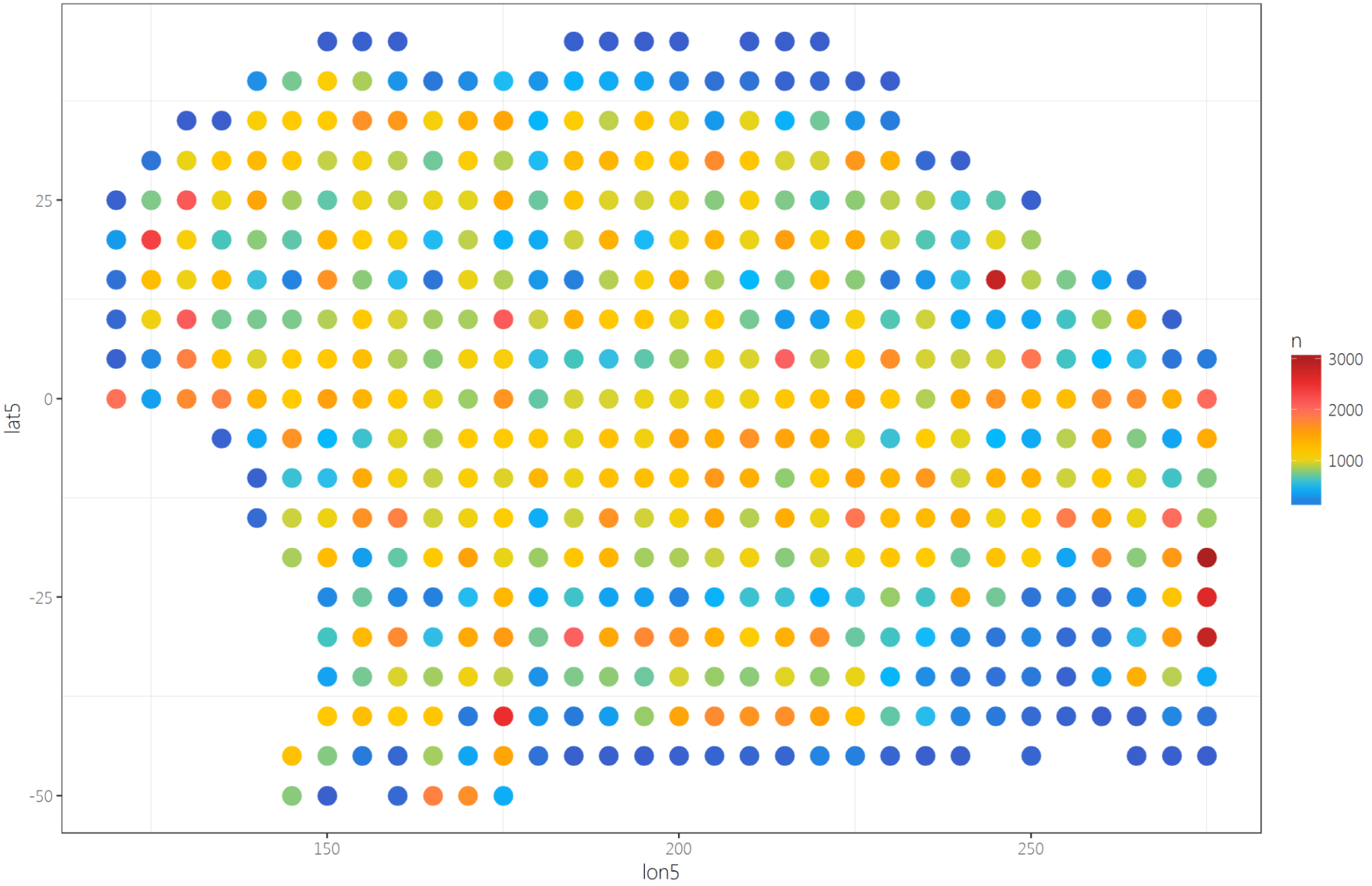
Relationship between effort and CPUE by decade



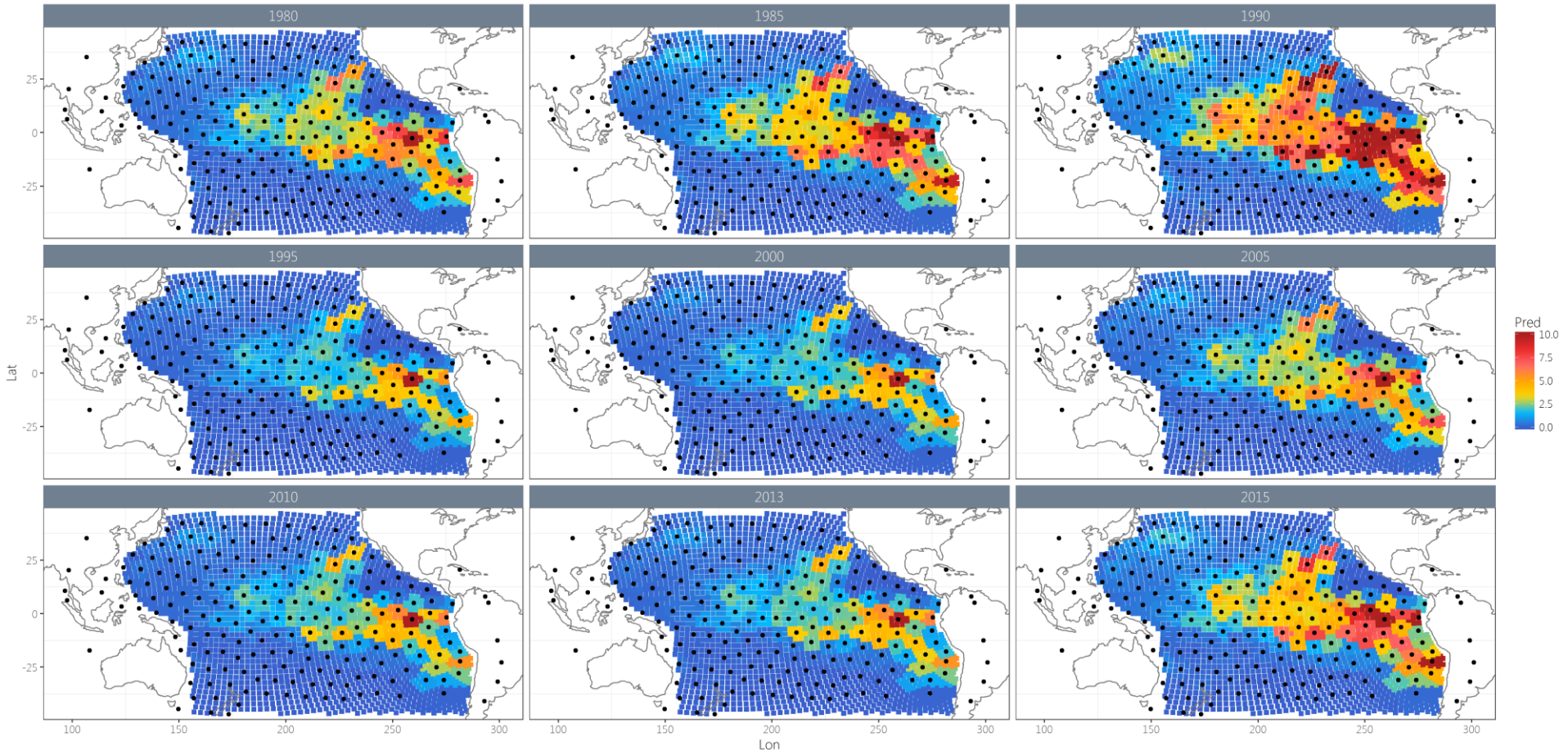
Subsampling by flag



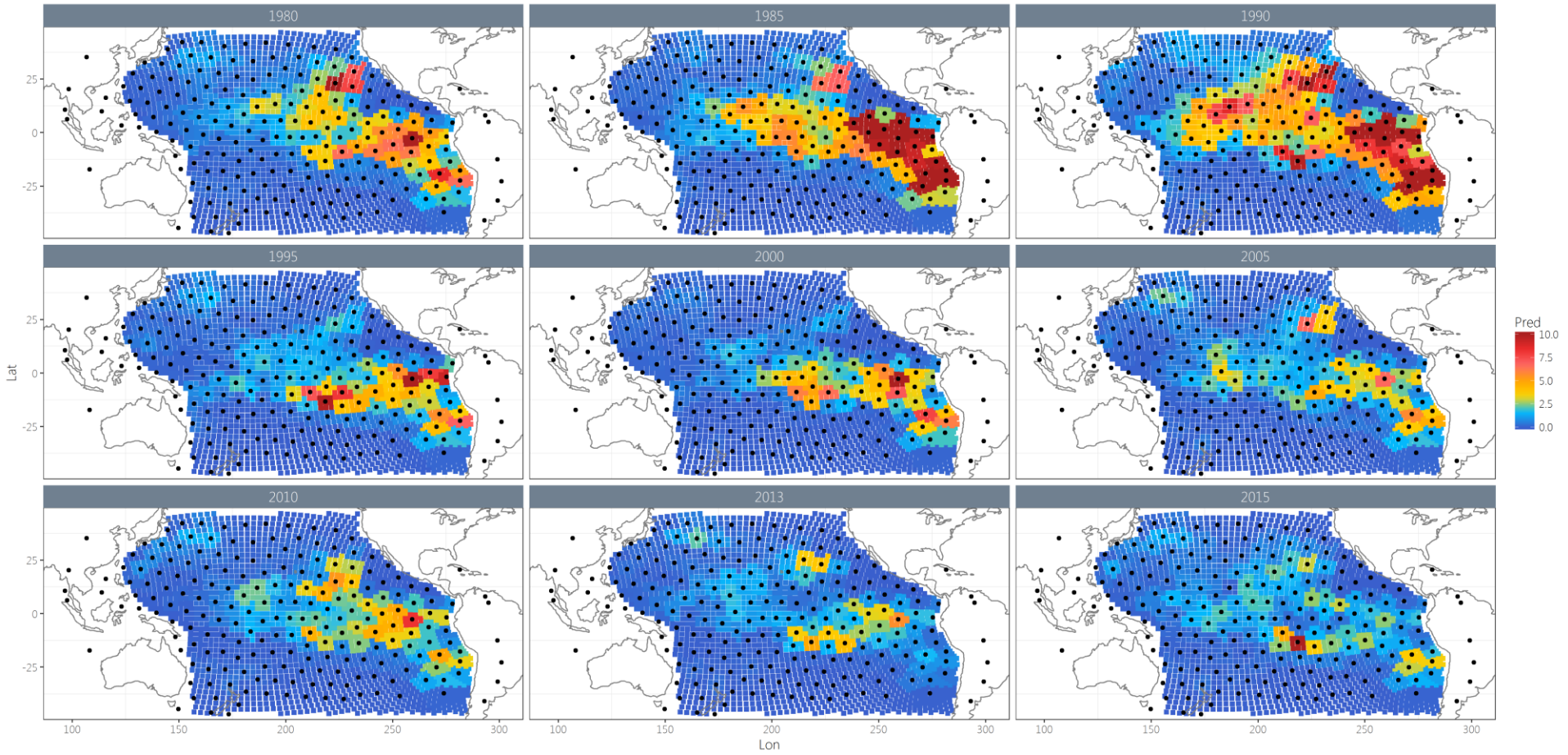
Subsampling by knot



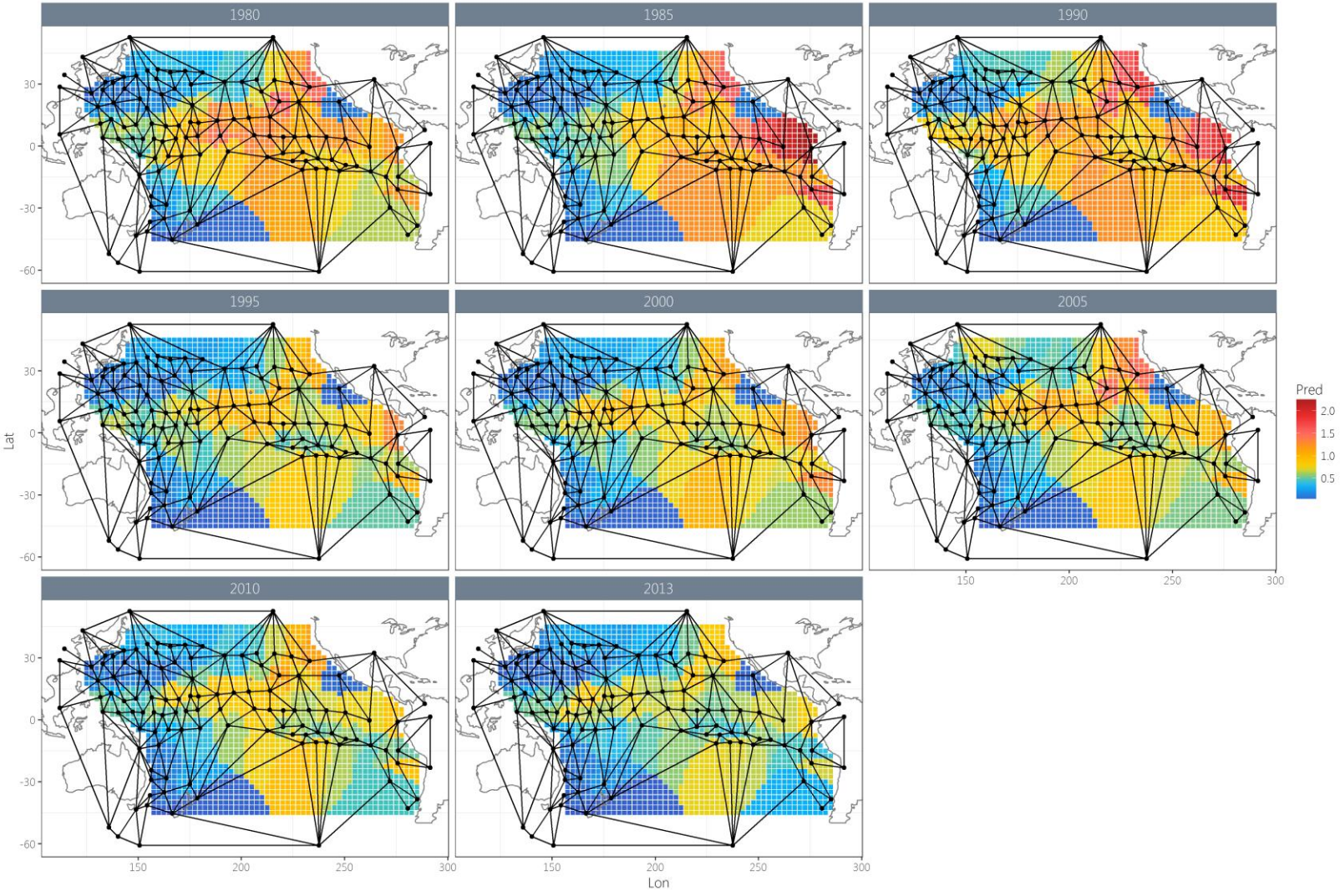
Subsampling by knot, with year interaction

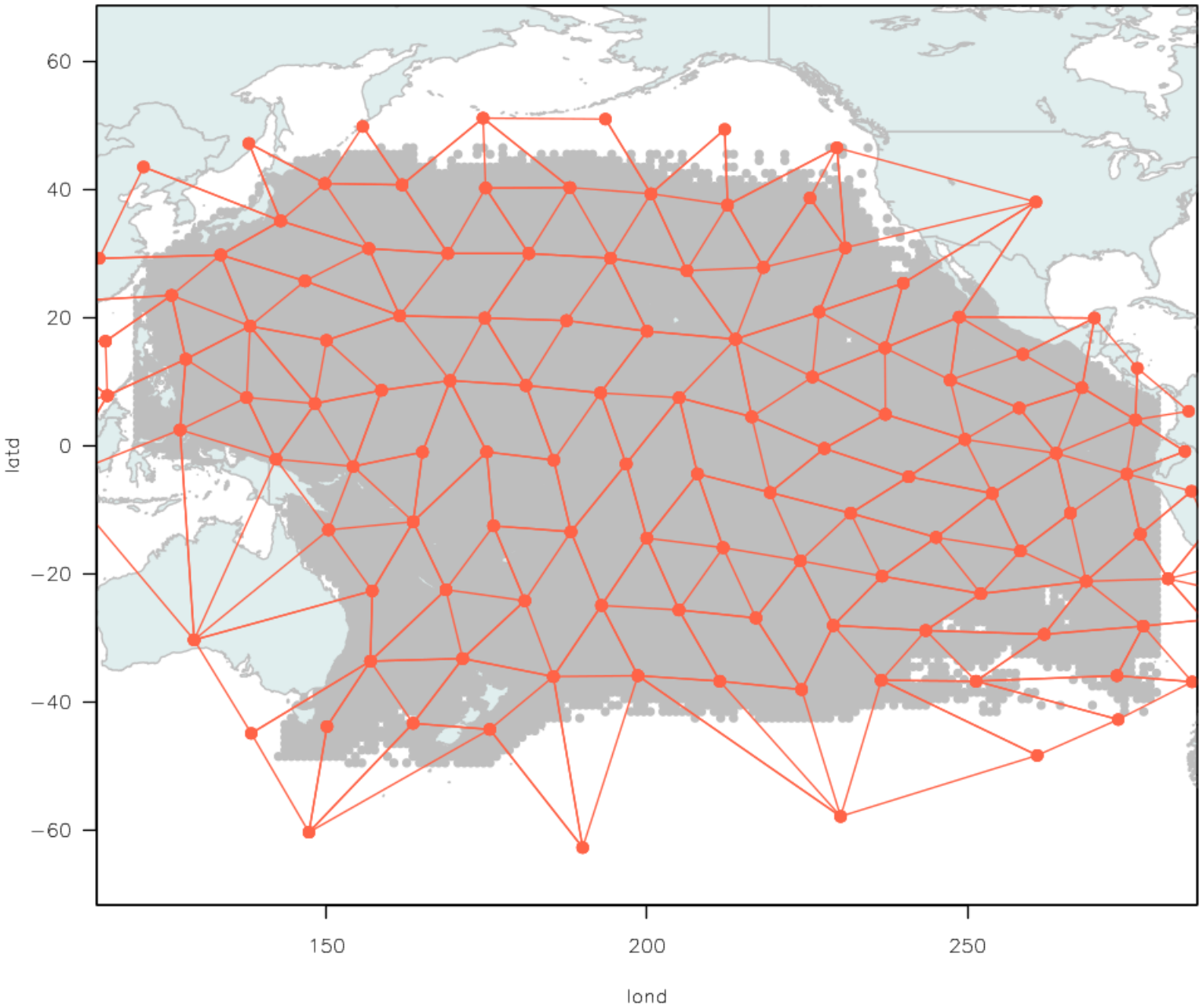


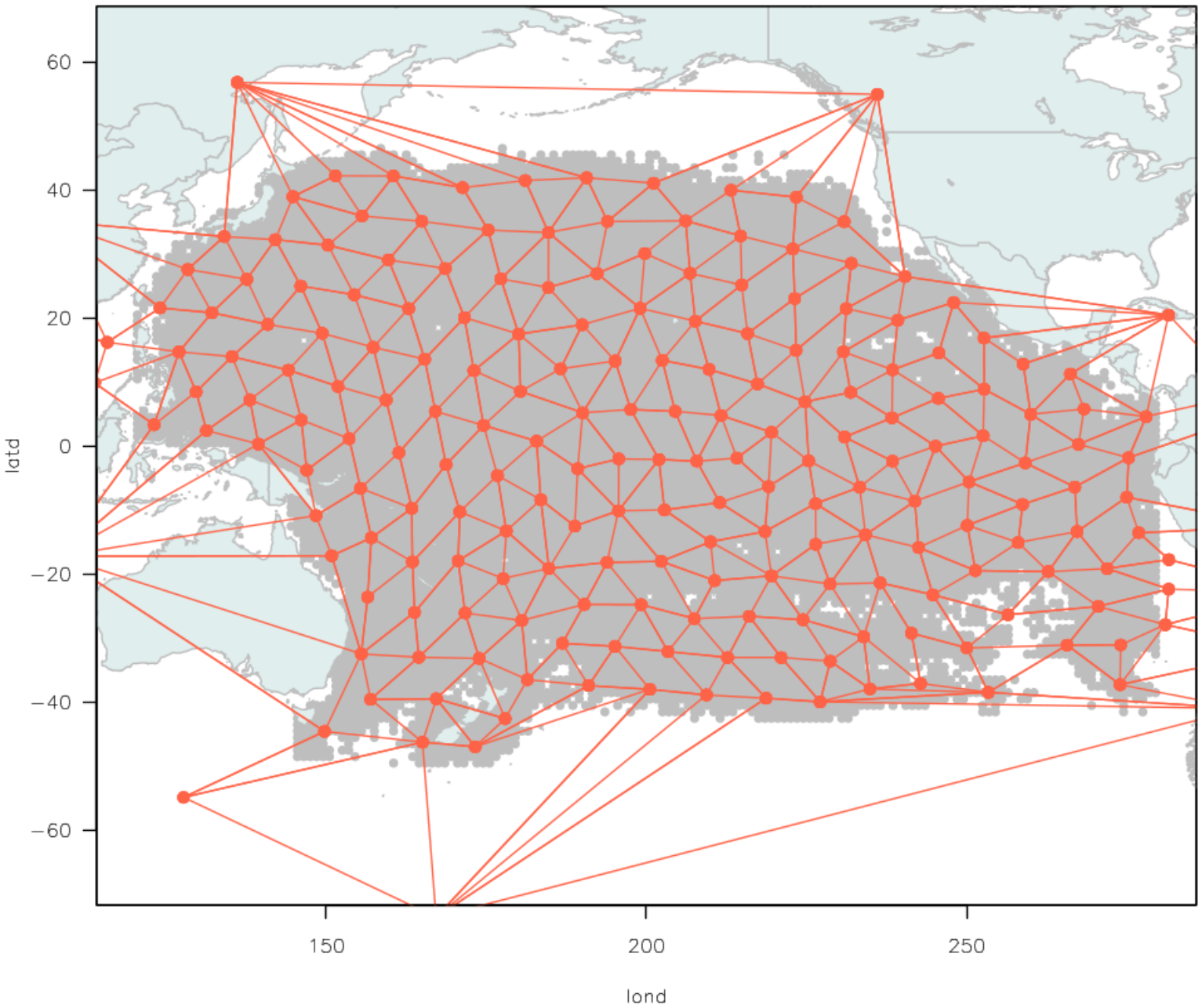
Subsampling by flag, with year interaction

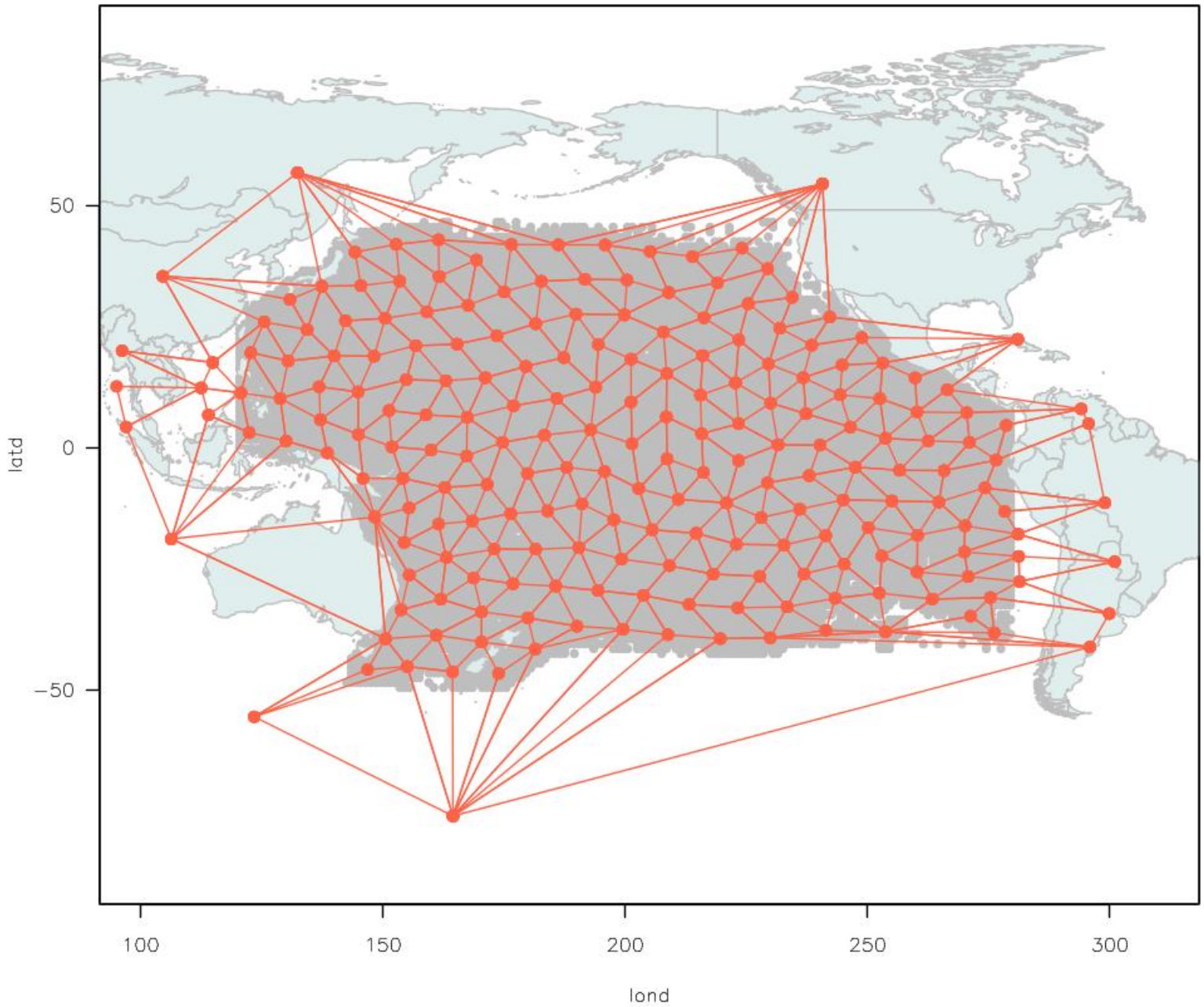


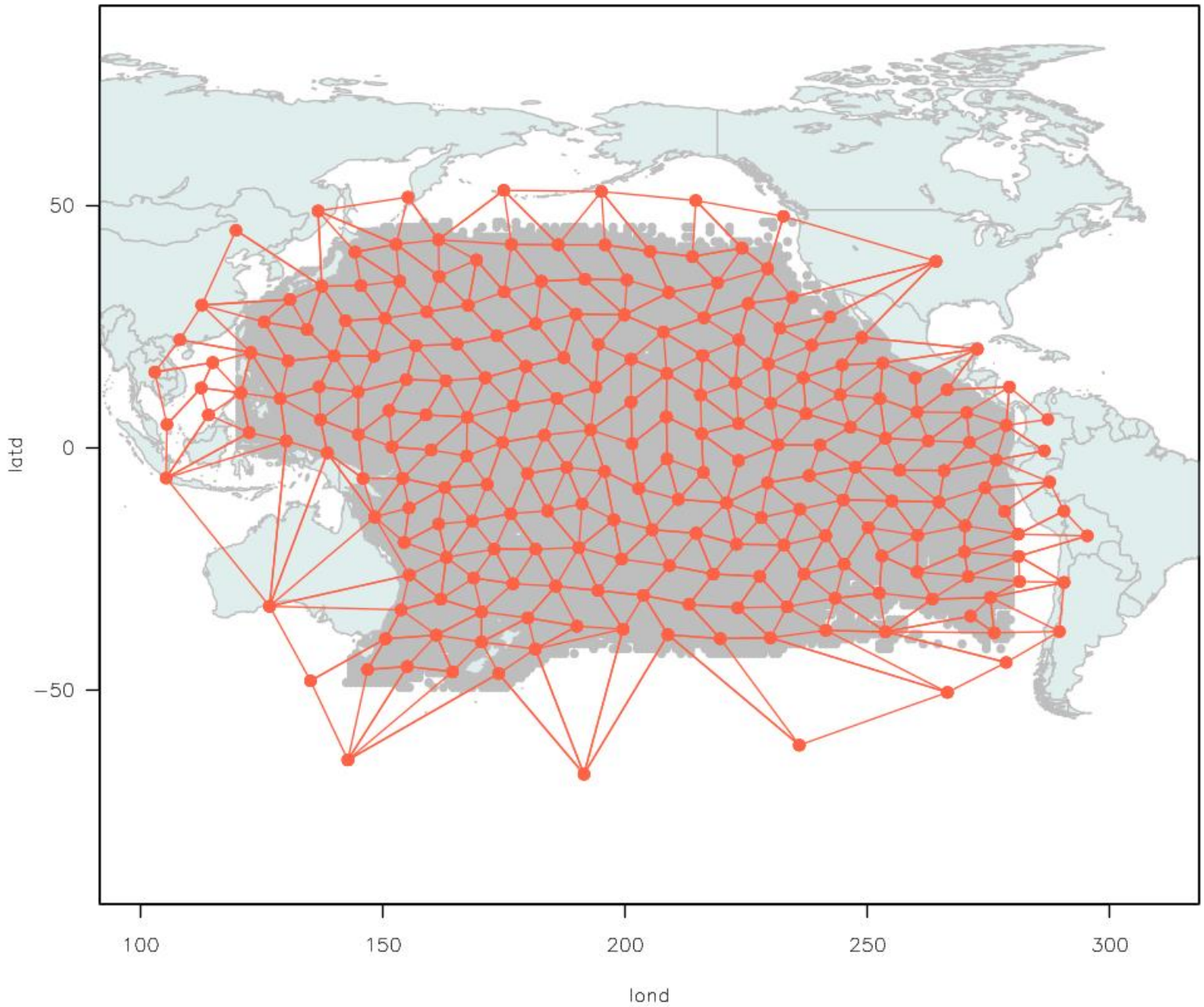
Mesh configuration?



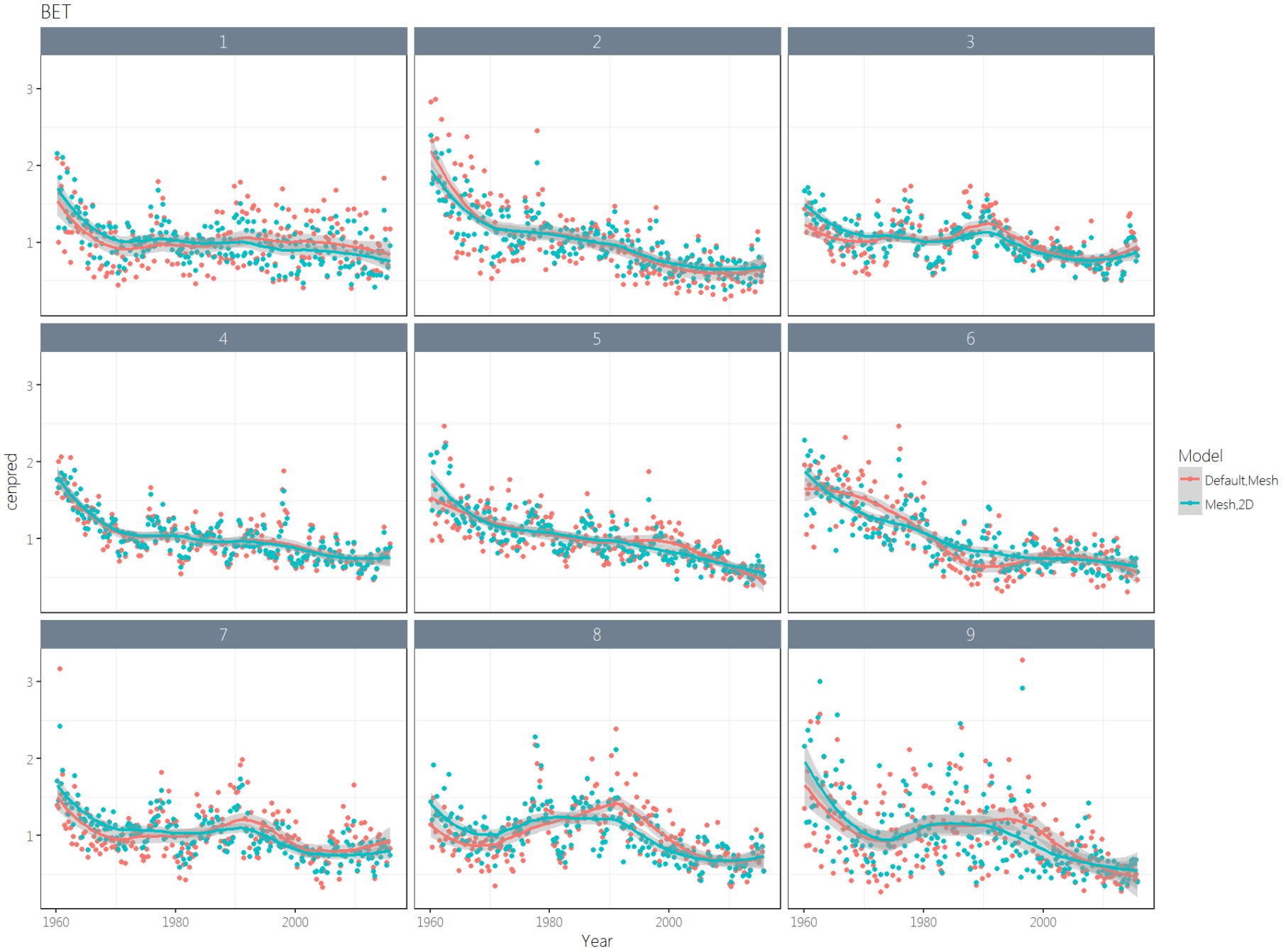

60
40
20
0
−20
−40
−60
latd
150
200
250
lond


60
40
20
0
−20
−40
−60
latd
150
200
250
lond

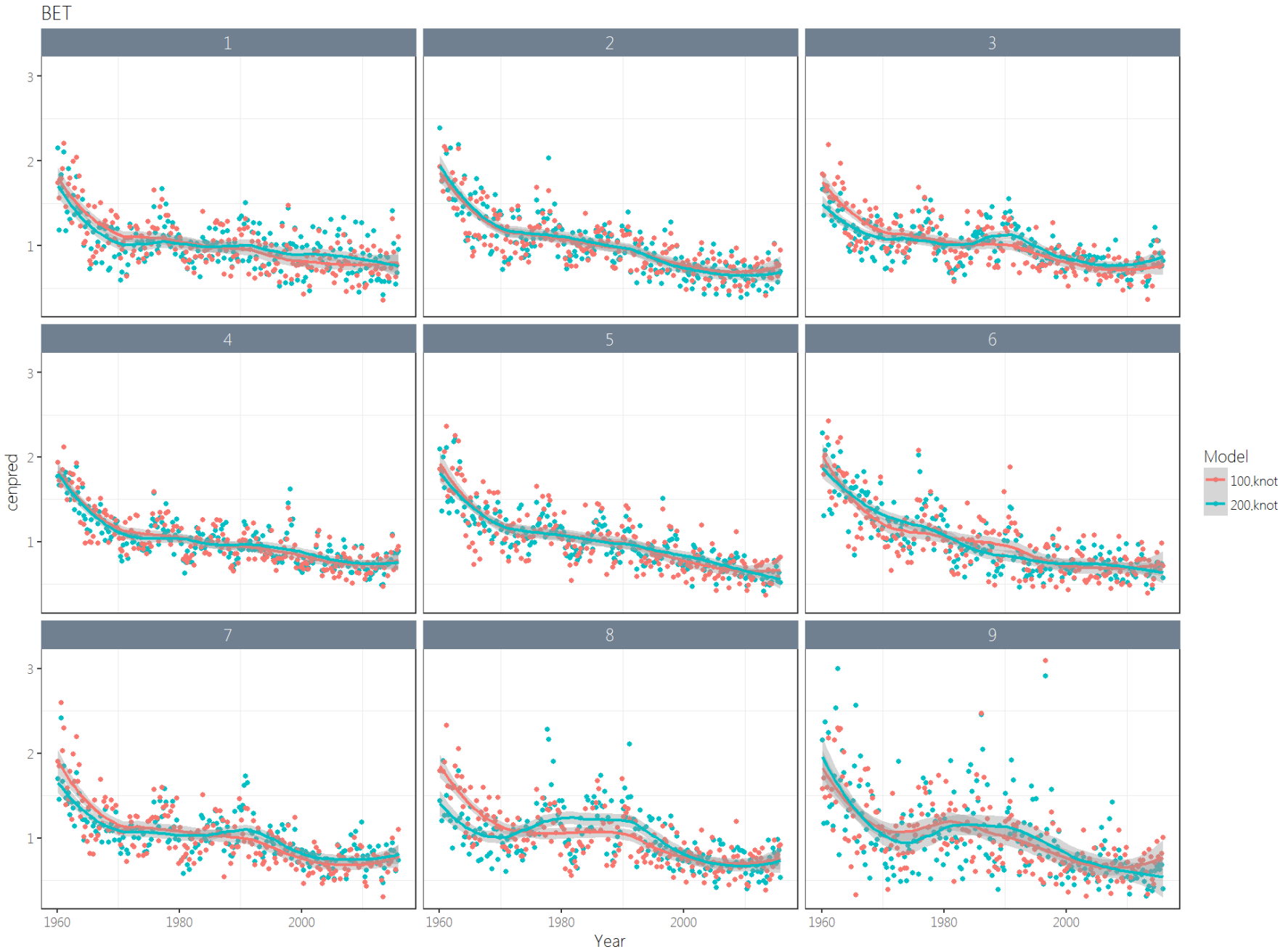

50
0
−50
latd
100
150
200
250
300
lond


50
0
−50
latd
100
150
200
250
300
lond

Mesh structure at the edges matters



knots matters



Inclusion of oceanography covariates in CPUE standardization

catchability vs. abundance

CPUE ~ YrQtr + […] + ocn-covar

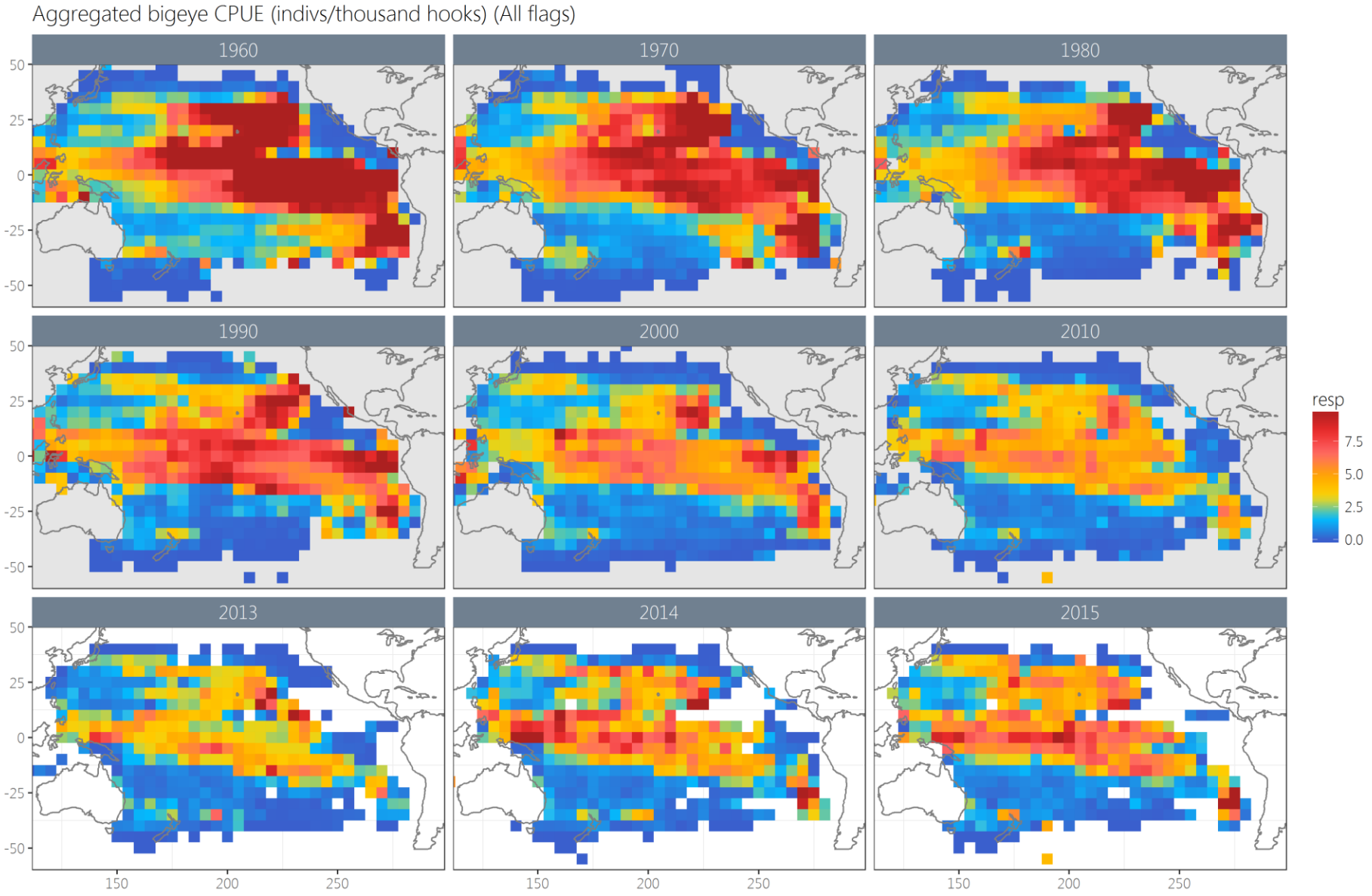
?

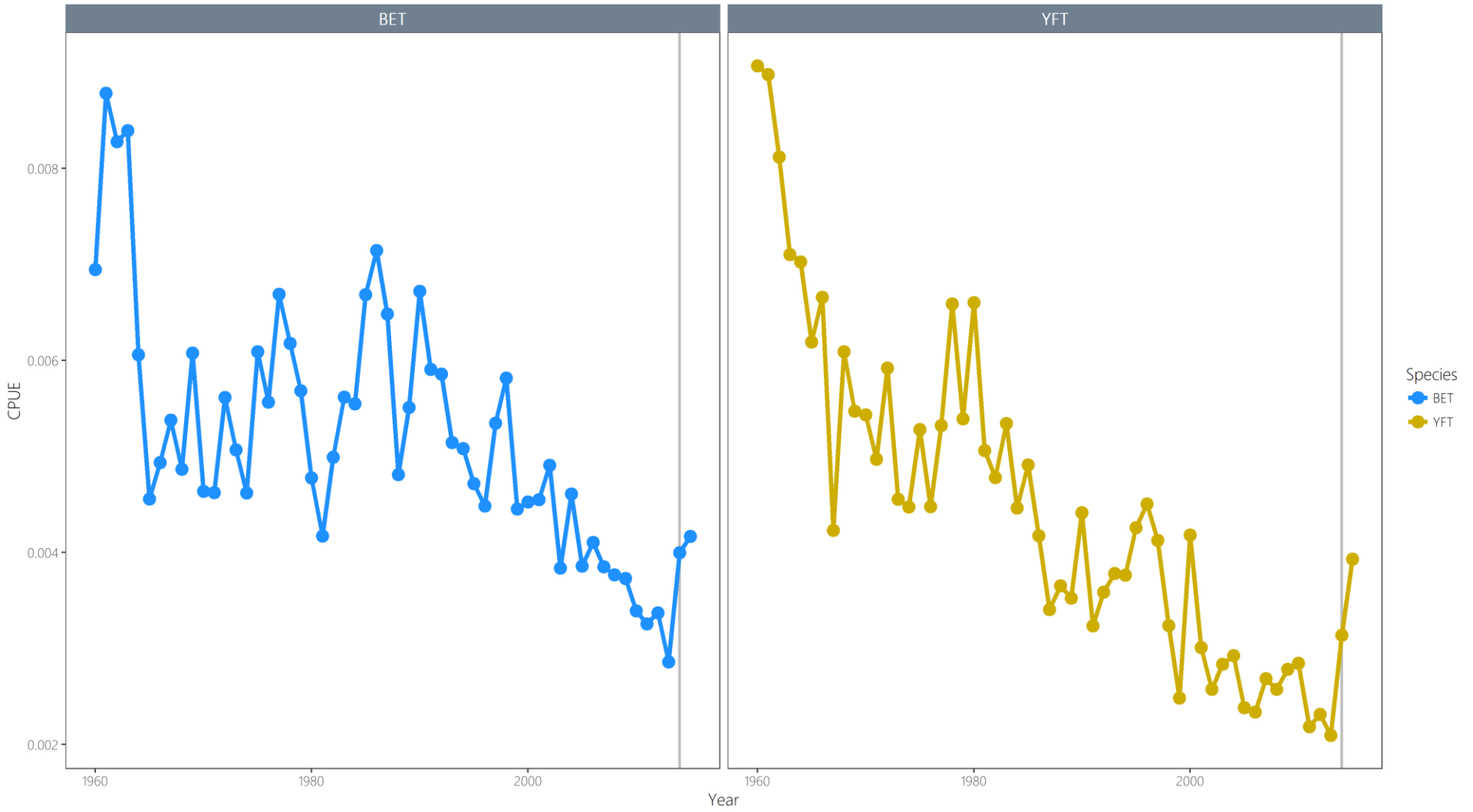
If the oceanography variable impacts abundance, we do not want it to be standardized against

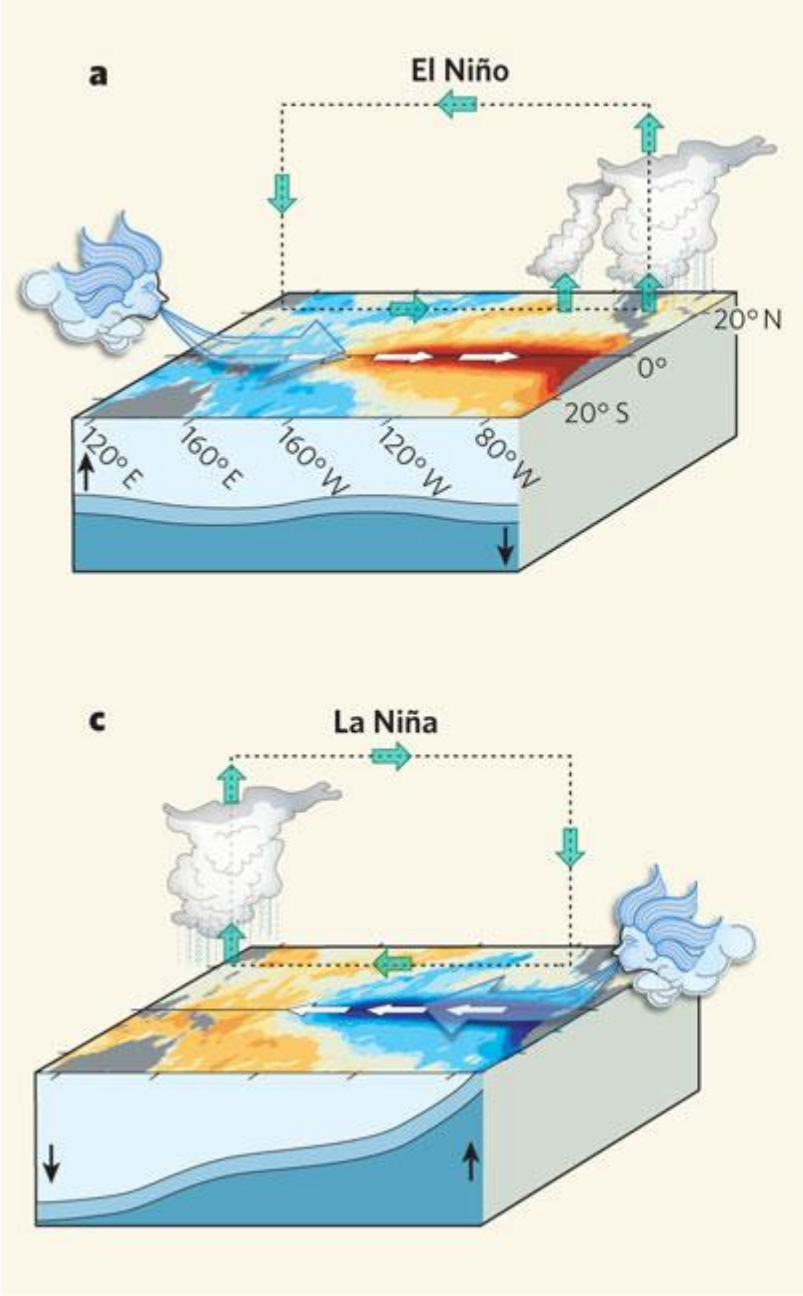
- Collinearity between oceanography variables?

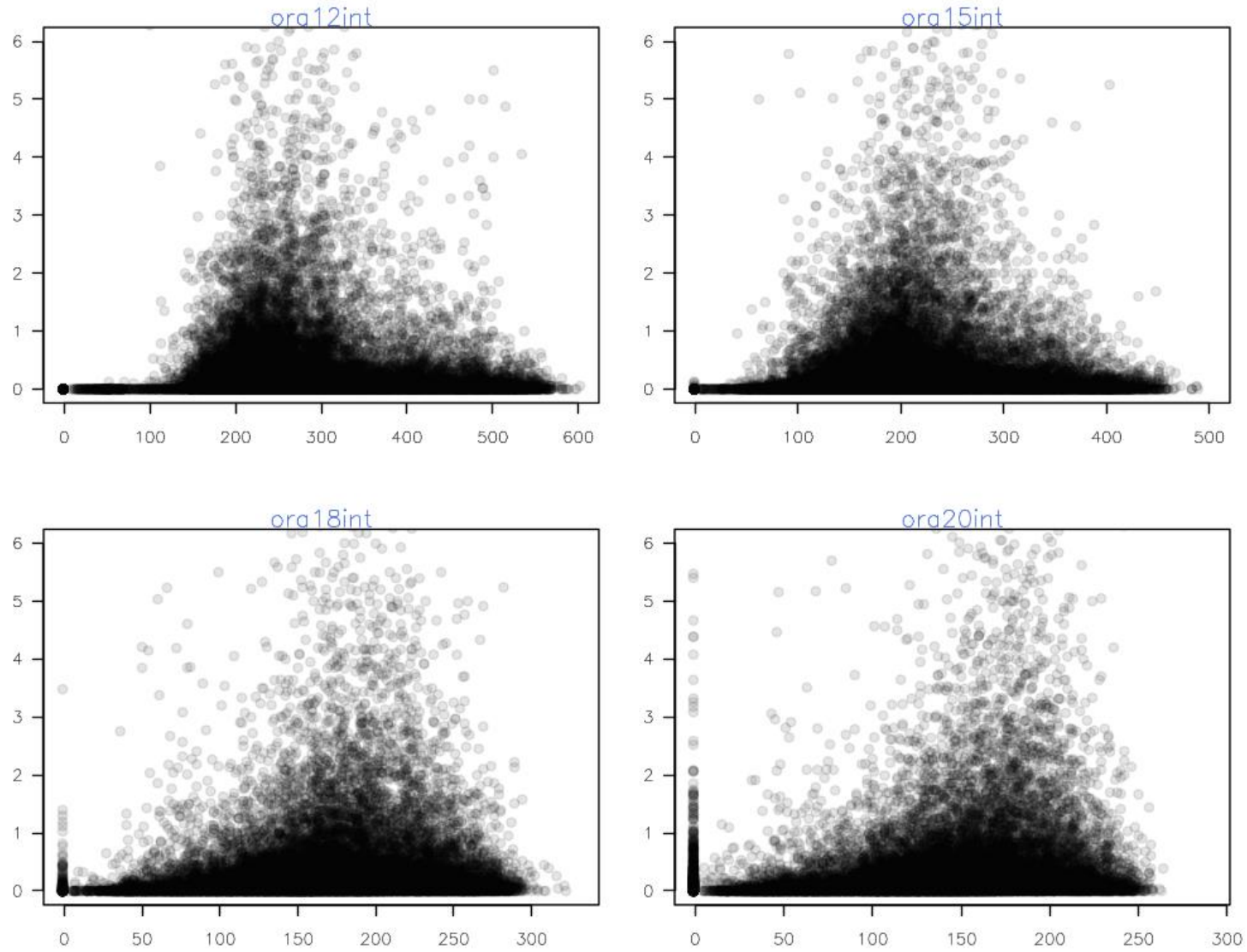
Other considerations for depth proxy:
Non-linear relationship

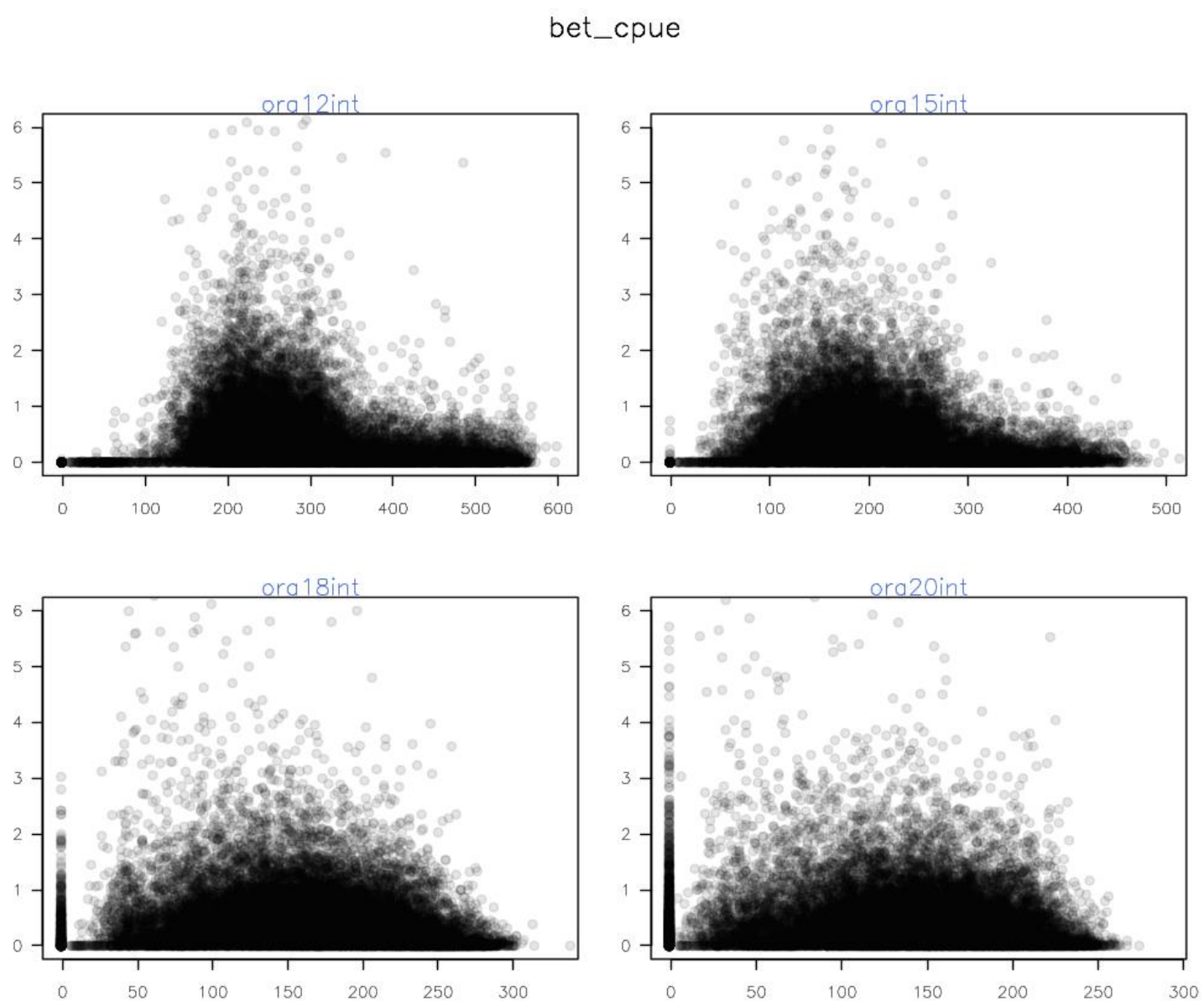
Inclusion of oceanography covariates



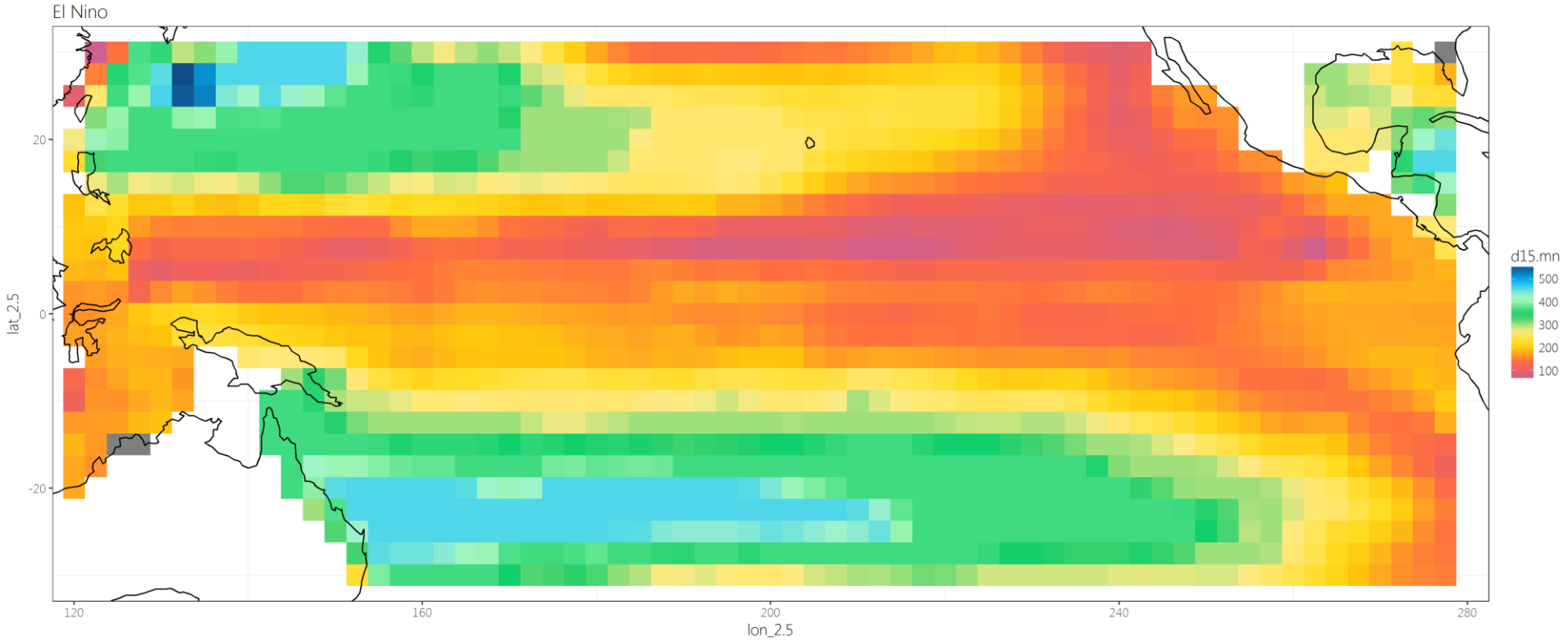

BET
YFT
CPUE
0.008
0.006
0.004
0.002
1960
1980
2000
Year
Species
BET
YFT


a
El Niño
20° N
0°
20° S
120° E
160° E
160° W
120° W
80° W
c
La Niña

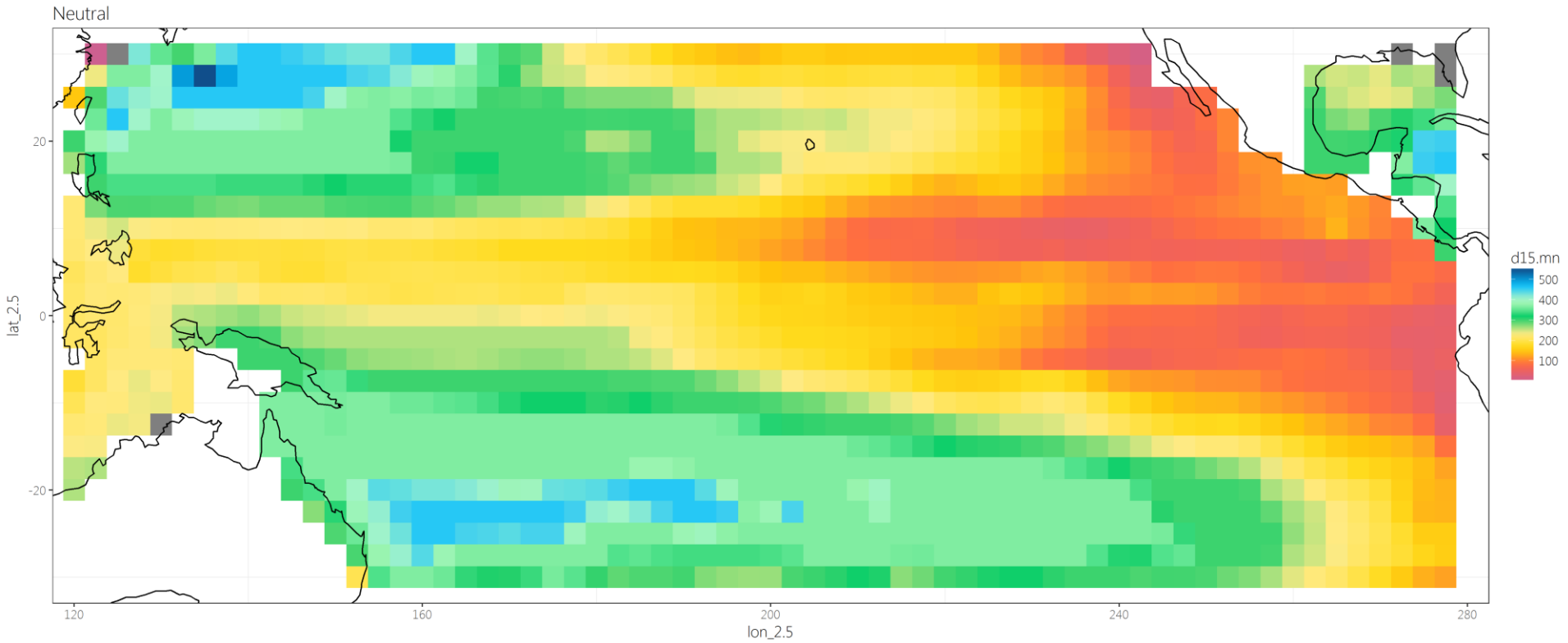

yft_cpue
org12int
org15int
org18int
org20int
0
1
2
3
4
5
6
0 100 200 300 400 500 600
0 100 200 300 400 500
0 50 100 150 200 250 300
0 50 100 150 200 250 300


bet_cpue
org12int
org15int
org18int
org20int
0
1
2
3
4
5
6
0
100
200
300
400
500
600
0
50
100
150
200
250
300

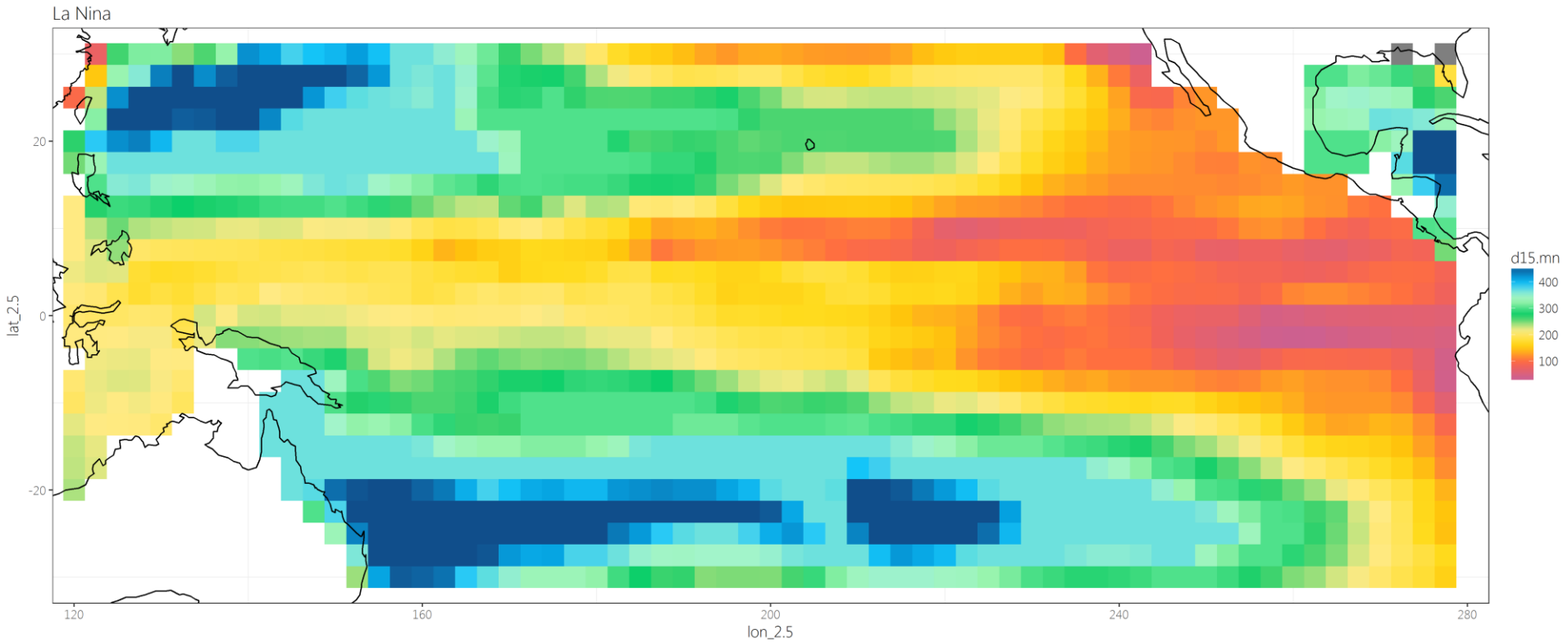
Depth of 15th degree layer
El Nino



Neutral


Neutral
20
0
-20
lat_2.5
120
160
200
240
280
lon_2.5
d15.mn
500
400
300
200
100

La Nina


La Nina
20
0
-20
lat_2.5
120
160
200
240
280
lon_2.5
d15.mn
400
300
200
100

DIY custom VAST non-linear relationship

Steal design matrix from mgcv::gam (zero-centered)

```
if(add.iso) { ## create zero-centered model matrix for thermocline spline
    if(therm.knots<0) stop('therm.knots should be greater than 0')
    thermospl_i <- gam(dg[,'Catch_KG'] ~ s(isoval, k=therm.knots+1), fit=FALSE)$X[,-1]
    # (...$X[,-1] = removing intercept term)
} else { ## model matrix filled with zero
    thermospl_i <- matrix(rep(0, length(r2k)), ncol=1)
}
```

Add your new covariate in the .cpp:

```
// Covariates
vector<Type> eta1_x = X_xj * gamma1_j.matrix();
vector<Type> zeta1_i = Q_ik * lambda1_k.matrix();
vector<Type> eta2_x = X_xj * gamma2_j.matrix();
vector<Type> zeta2_i = Q_ik * lambda2_k.matrix();
// Oceanography spline
vector<Type> therm1_i = thermospl_i * ocnspl1_i.matrix();
vector<Type> therm2_i = thermospl_i * ocnspl2_i.matrix();
vector<Type> delta1_i = thermodelta_i * deltaspl1_i.matrix();
vector<Type> delta2_i = thermodelta_i * deltaspl2_i.matrix();
```

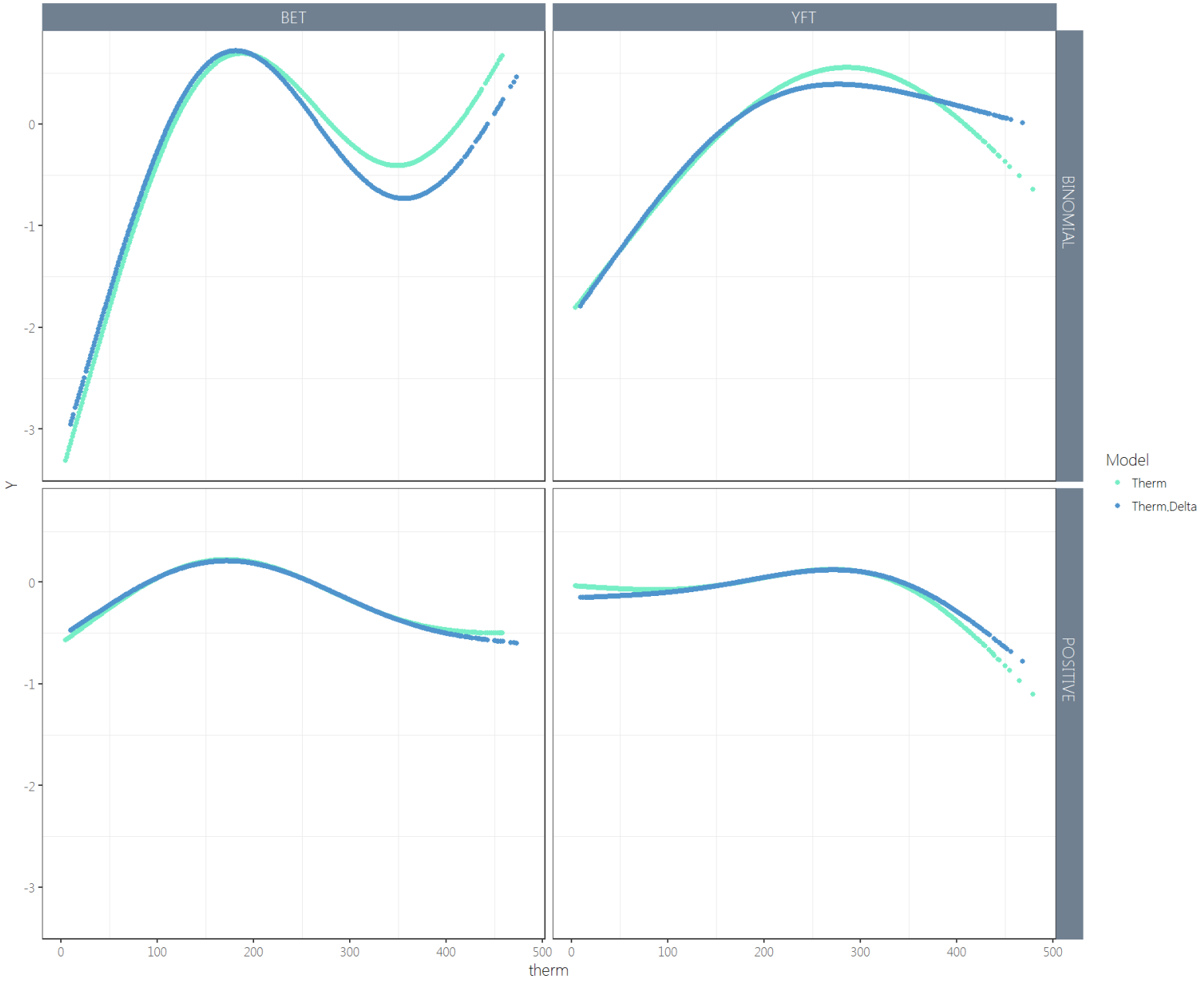
All the covariates go in the likelihood:

```
// Likelihood contribution from observations
for (int i=0;i<n_i;i++){
  // Presence-absence prediction
  P1_i(i) =  beta1_t(t_i(i)) + Omega1_s(s_i(i)) + Epsilon1_st(s_i(i
  R1_i(i) = invlogit( P1_i(i) );
  // Positive density prediction
  if( b_i(i)>0 | ObsModel(0)==5 | ObsModel(0)==6 ){    // 1e-500 ca
    P2_i(i) =  beta2_t(t_i(i)) + Omega2_s(s_i(i)) + Epsilon2_st(s_i
    R2_i(i) = exp( P2_i(i) );
  }else{
    P2_i(i) = 0;
    R2_i(i) = 0;
  }
```

Only abundance covariates go in the predictions:

```
 for(int t=0; t<n_t; t++){
 for(int x=0; x<n_x; x++){
   P1_xt(x,t) = beta1_t(t) + Omega1_s(x) + Epsilon1_st(x,t) + eta1_
   R1_xt(x,t) = invlogit( P1_xt(x,t) );
   P2_xt(x,t) =  beta2_t(t) + Omega2_s(x) + Epsilon2_st(x,t) + eta2
   if(ObsModel(0)==0 | ObsModel(0)==1 | ObsModel(0)==2 | ObsModel(0
   if(ObsModel(0)==11 | ObsModel(0)==12) R2_xt(x,t) = SigmaM(1)*exp
```

Thermocline as a catchability covariate only = Hmmm.



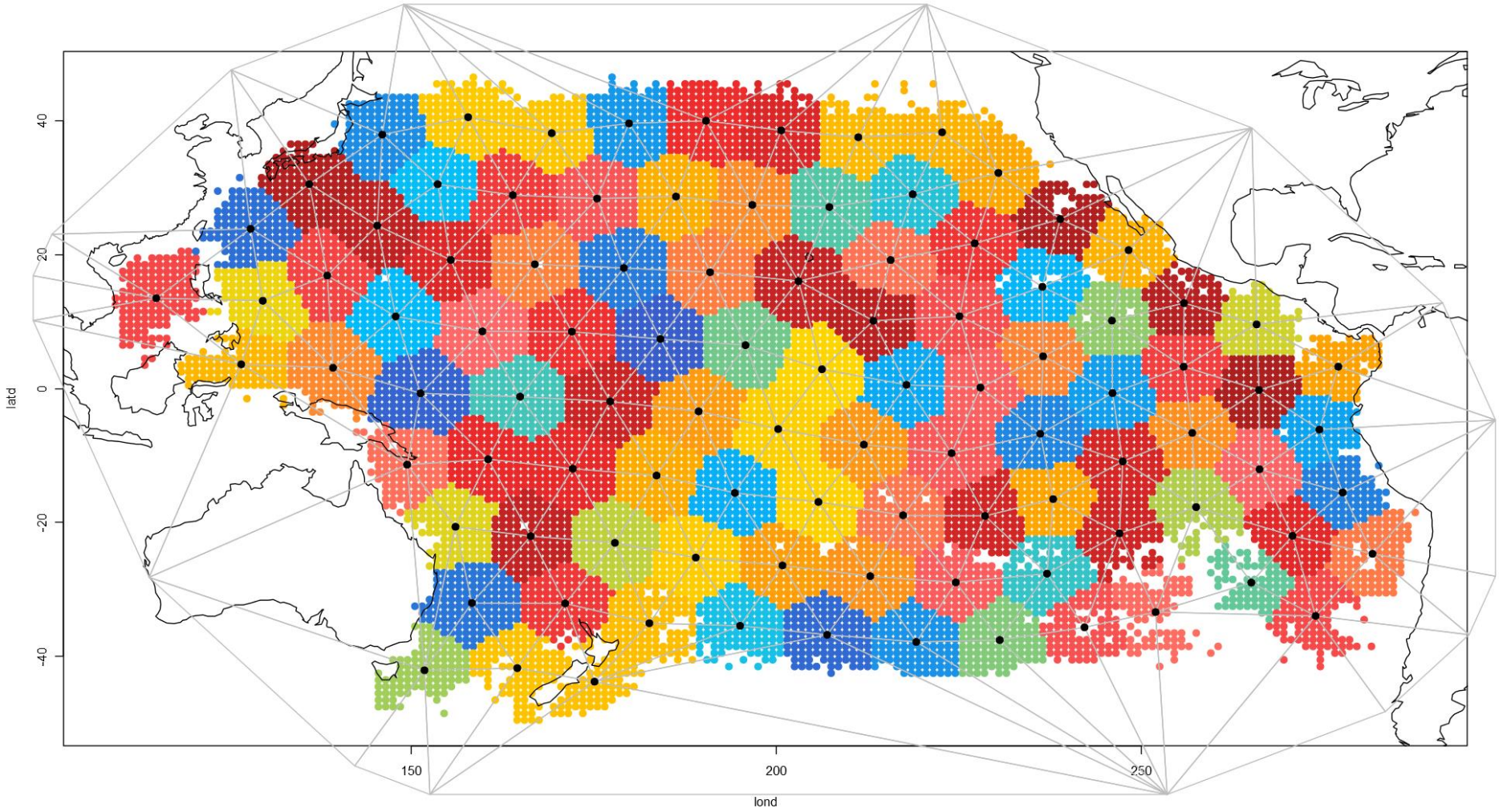
from resp ~ YrQtr + … + cell (*time)

to

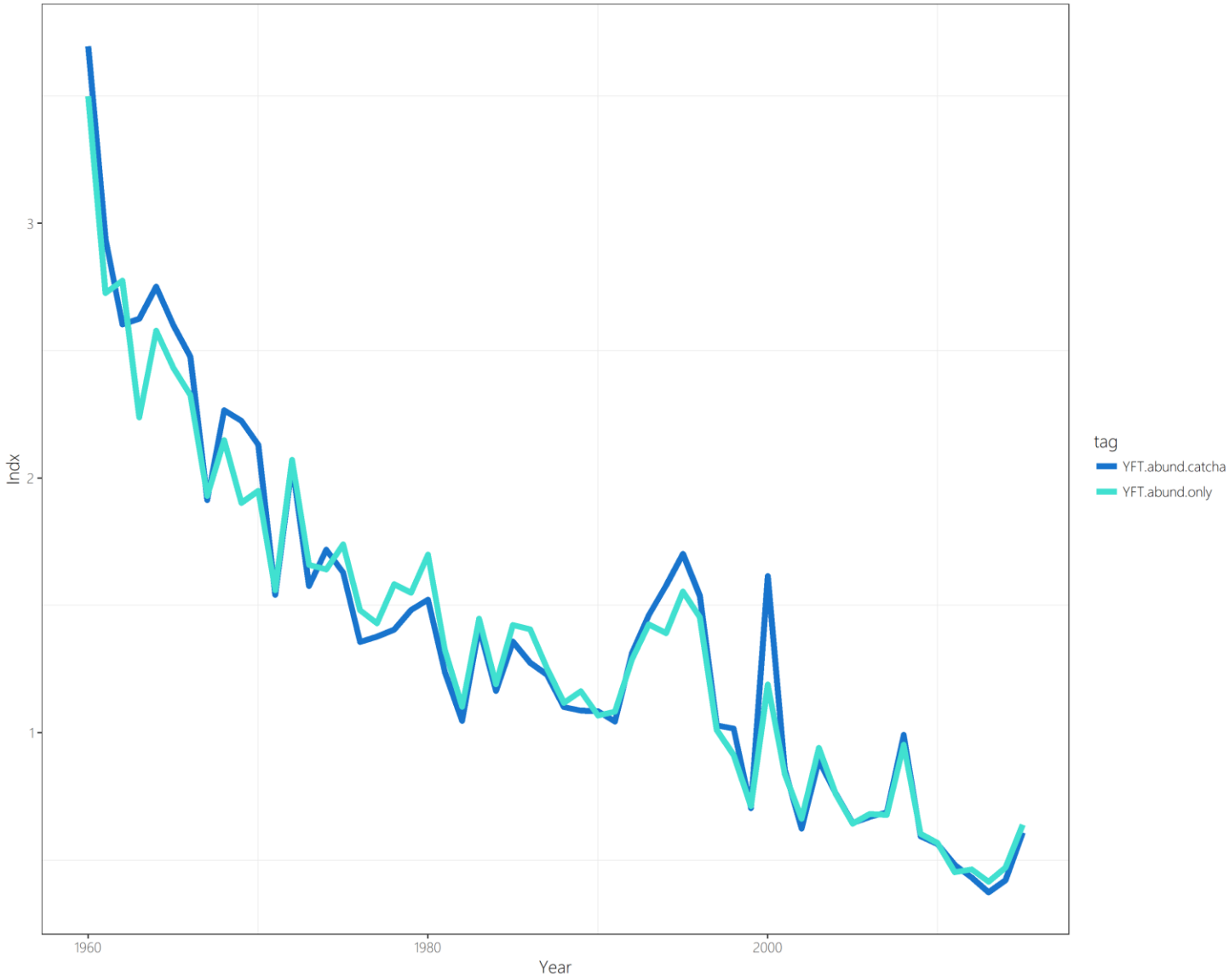
resp ~ YrQtr + … + f(knot_i) + f(knot_i,YrQtr)

[…] can be any other covariate

Error distribution: delta-log-normal

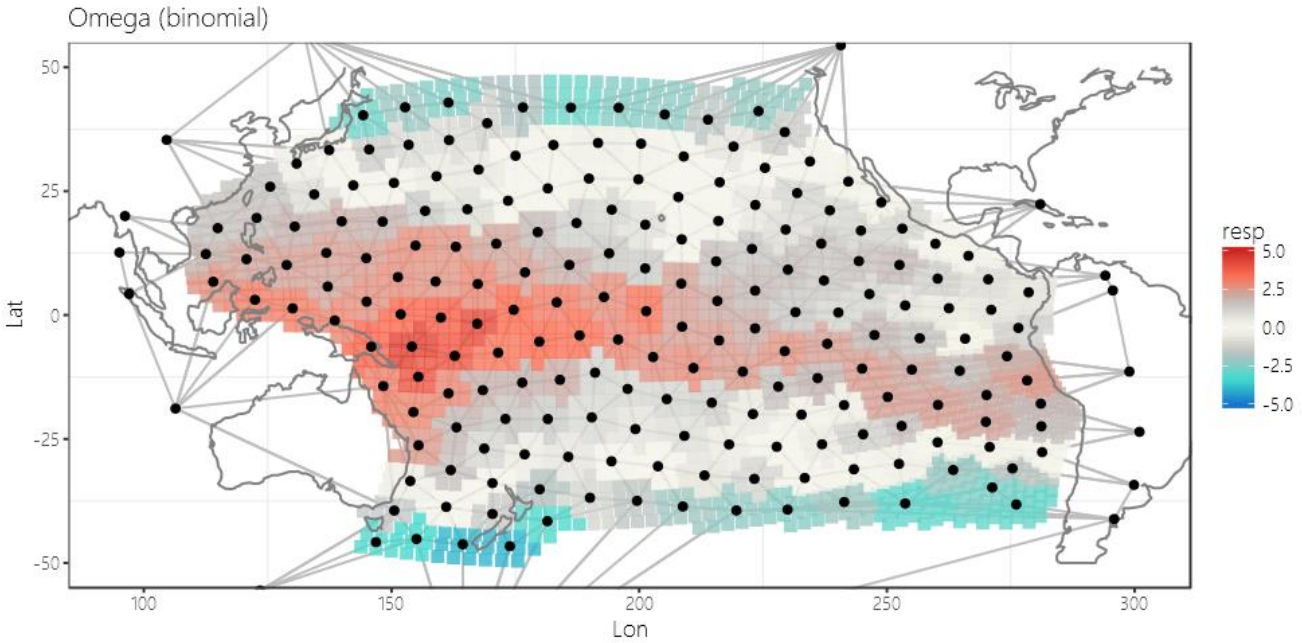


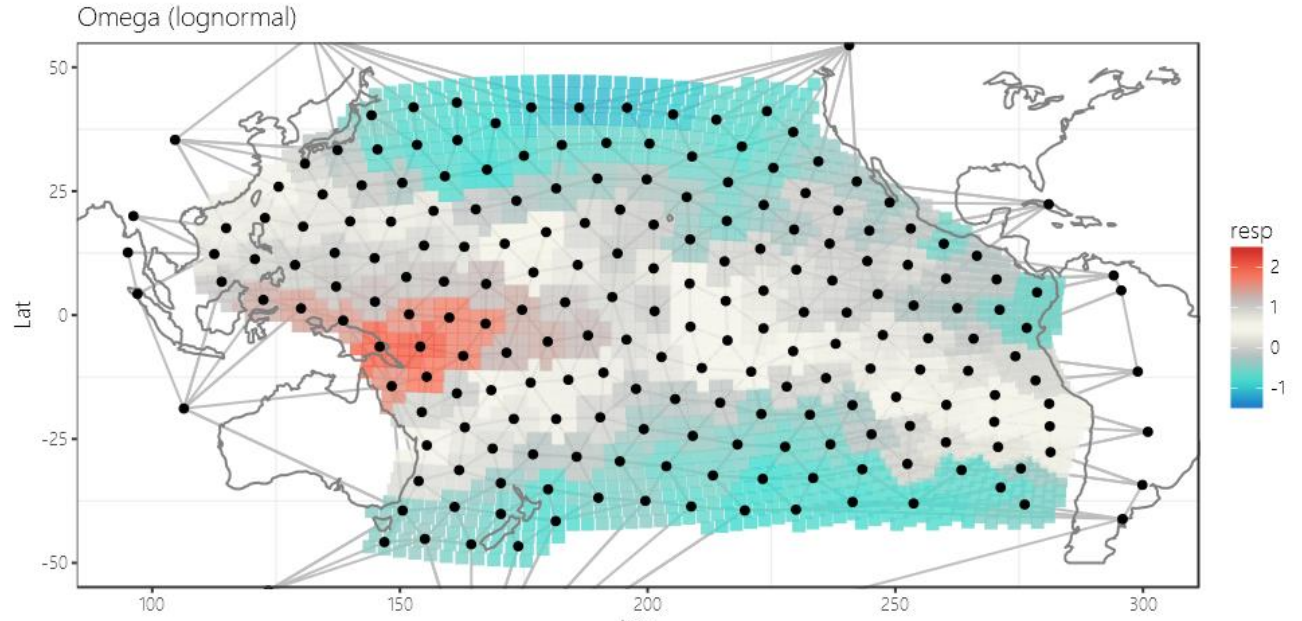
YFT: Thermocline as a catchability covariate + SST.m1 as an abundance covariate



YFT:

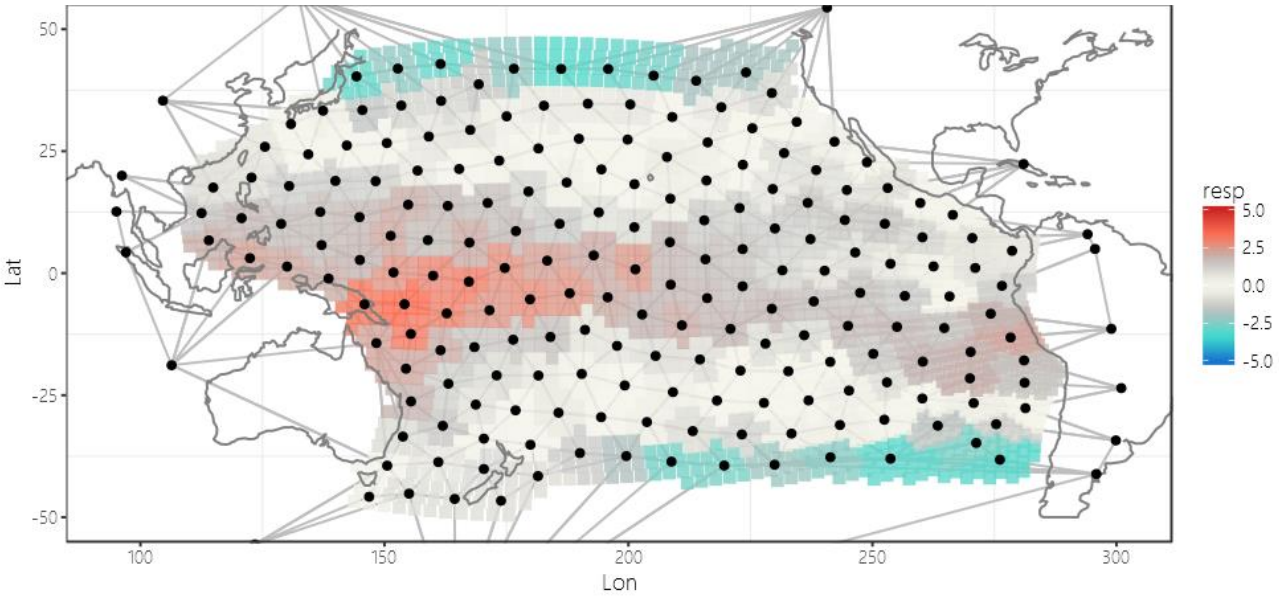
SST.m1 as abundance covariate, no catchability covariate

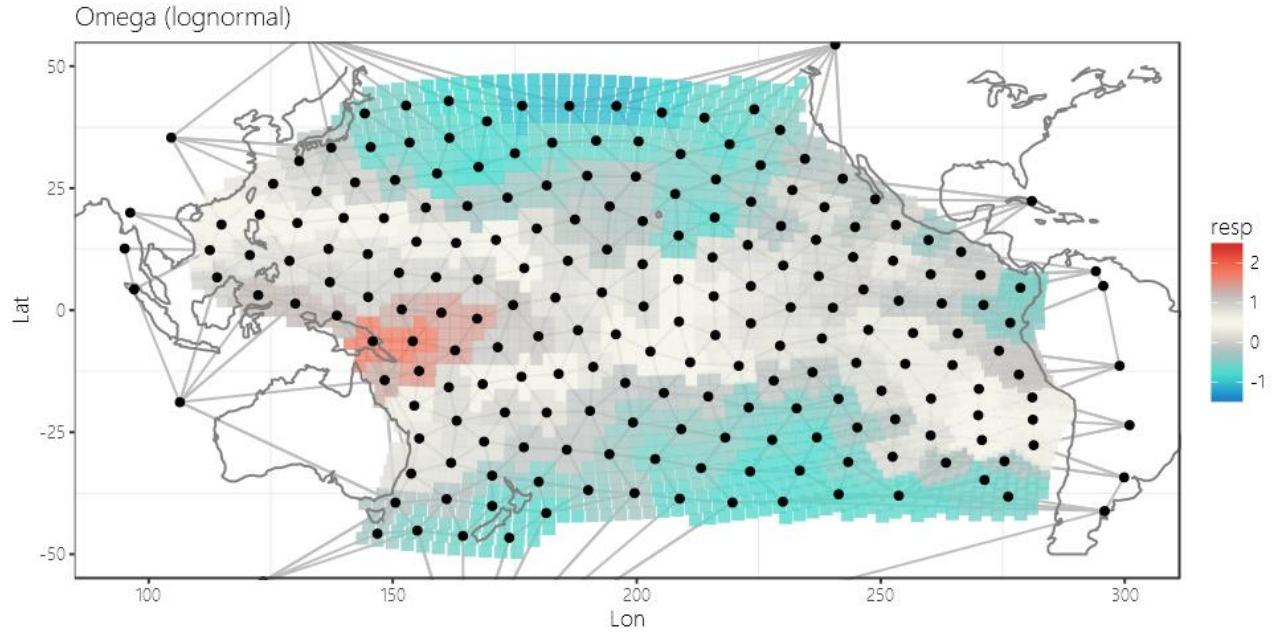




Yellowfin tuna:

SST.m1 as abundance cov., with depth of 20C as catchability covariate



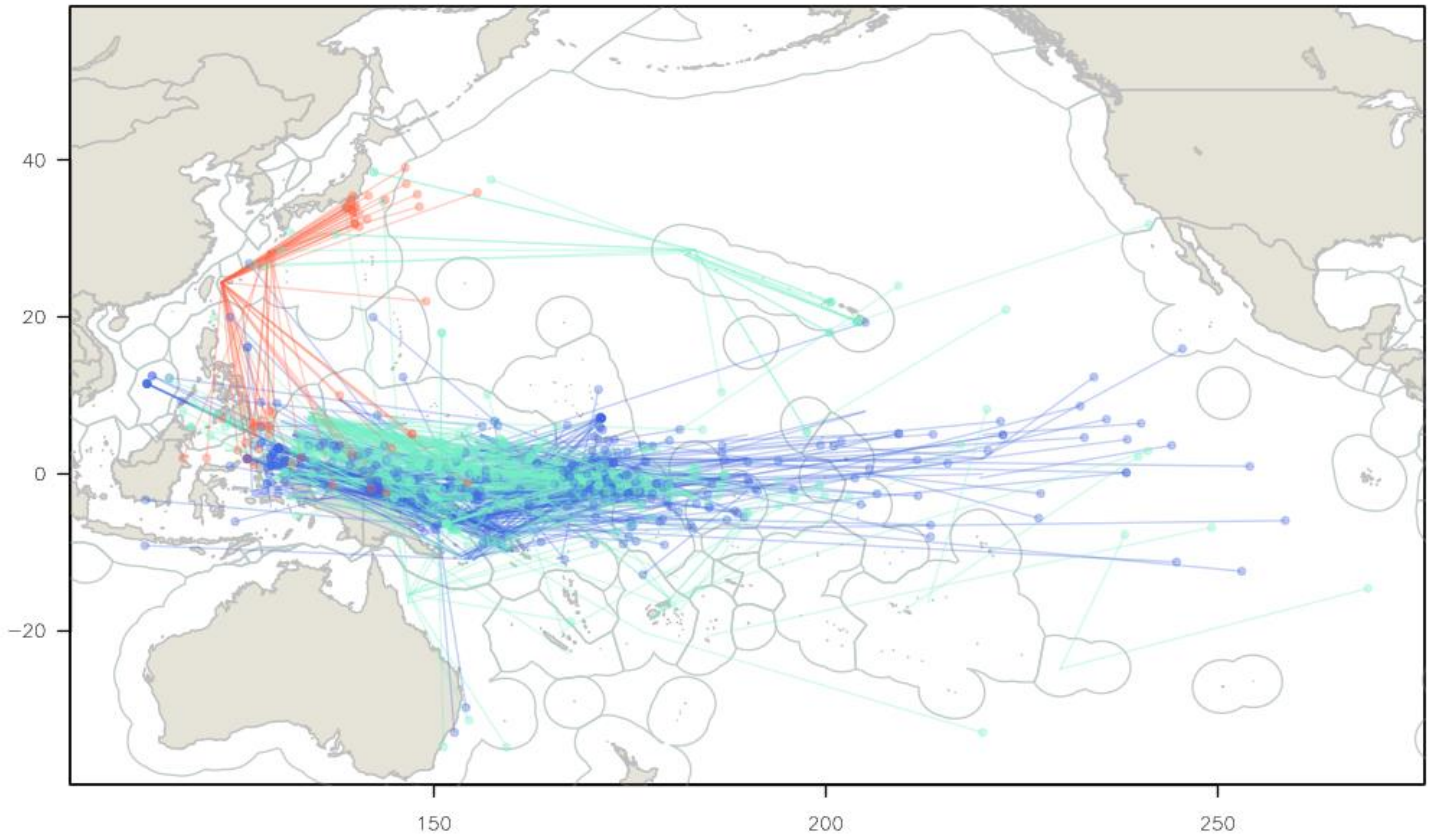


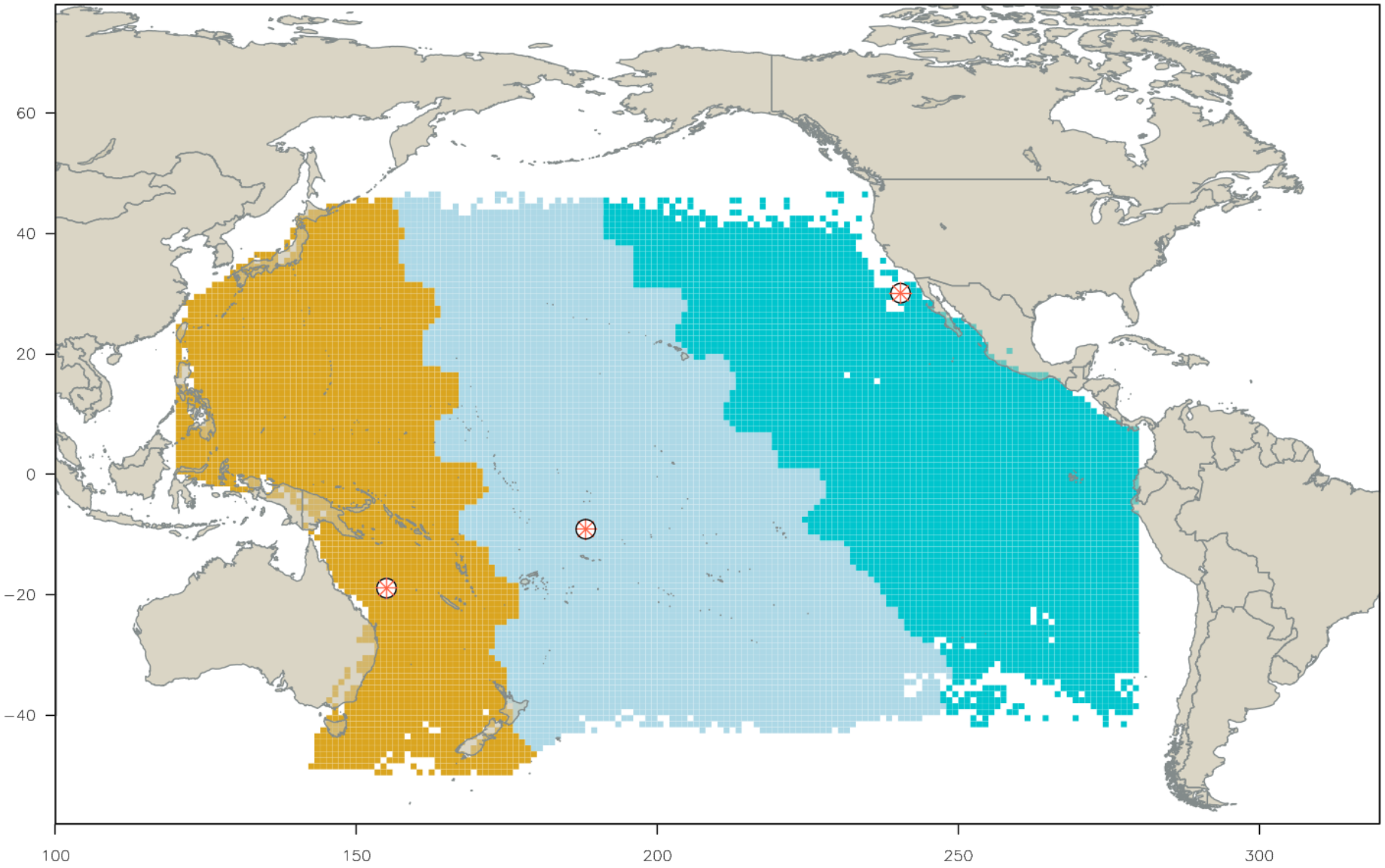
Plot twist…!

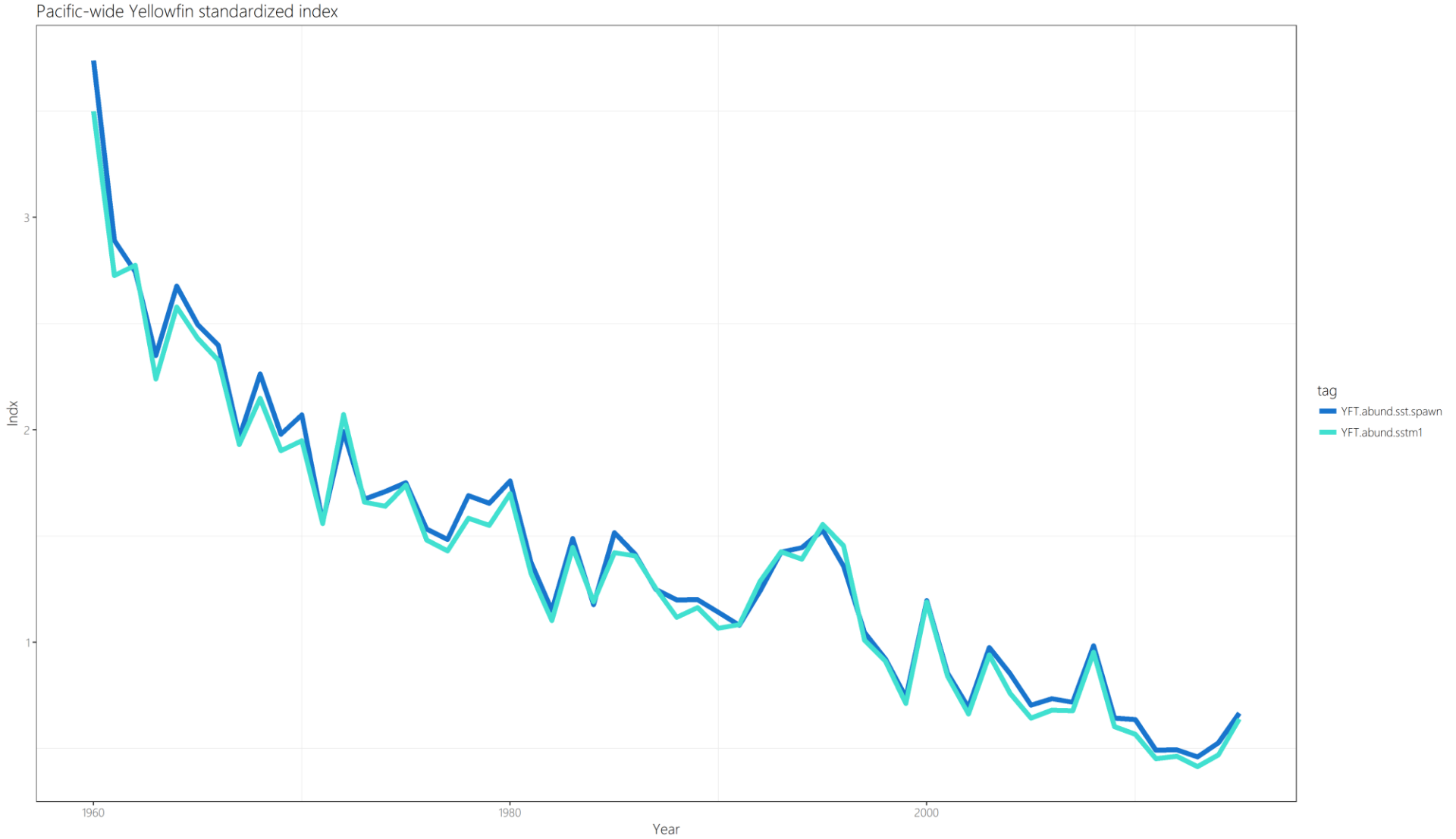
Regional or local abundance covariate?

SST from set location vs. regional trends in SST (recruitment)?

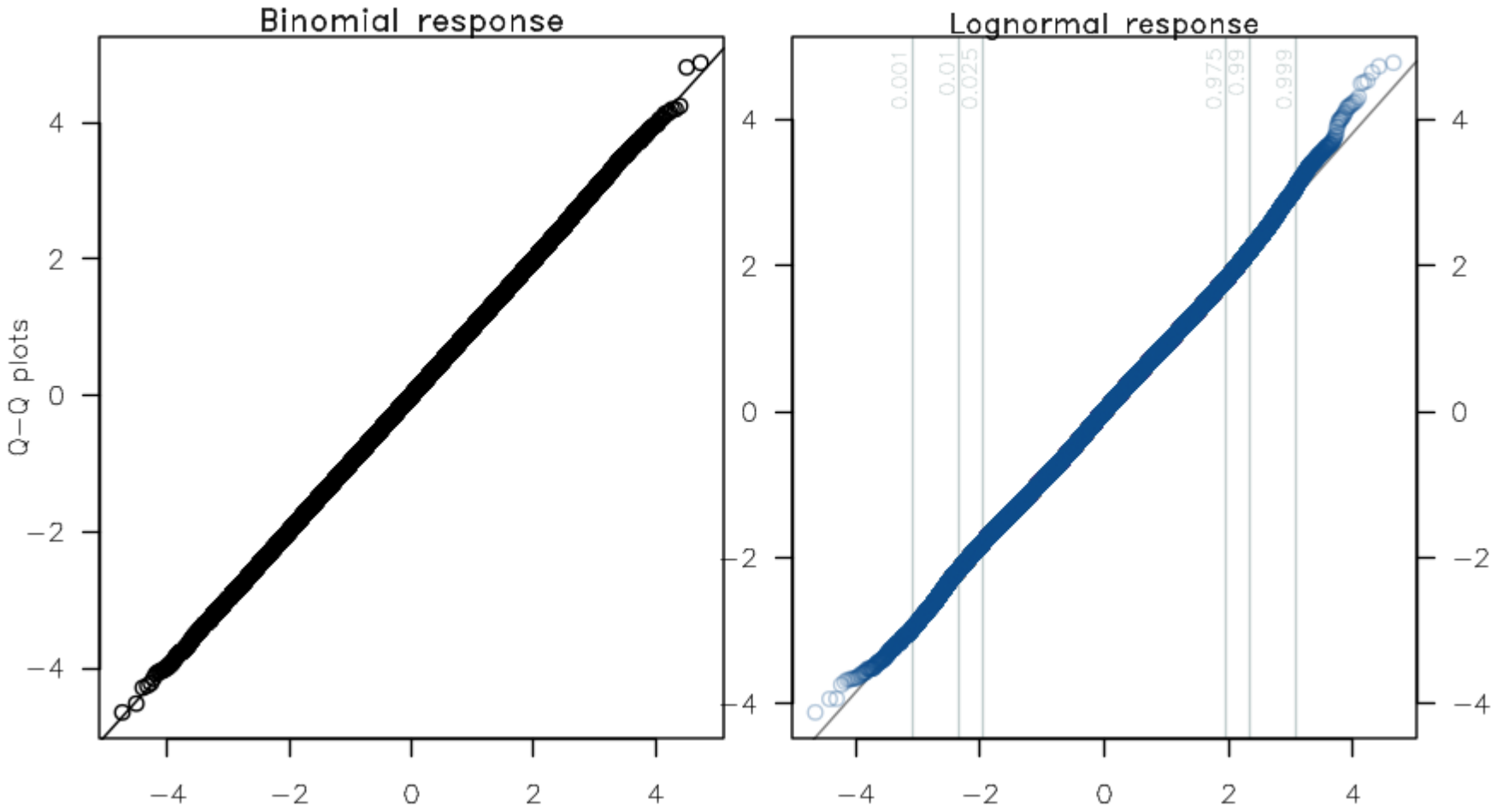
YFT tag recaptures > 1000nm:

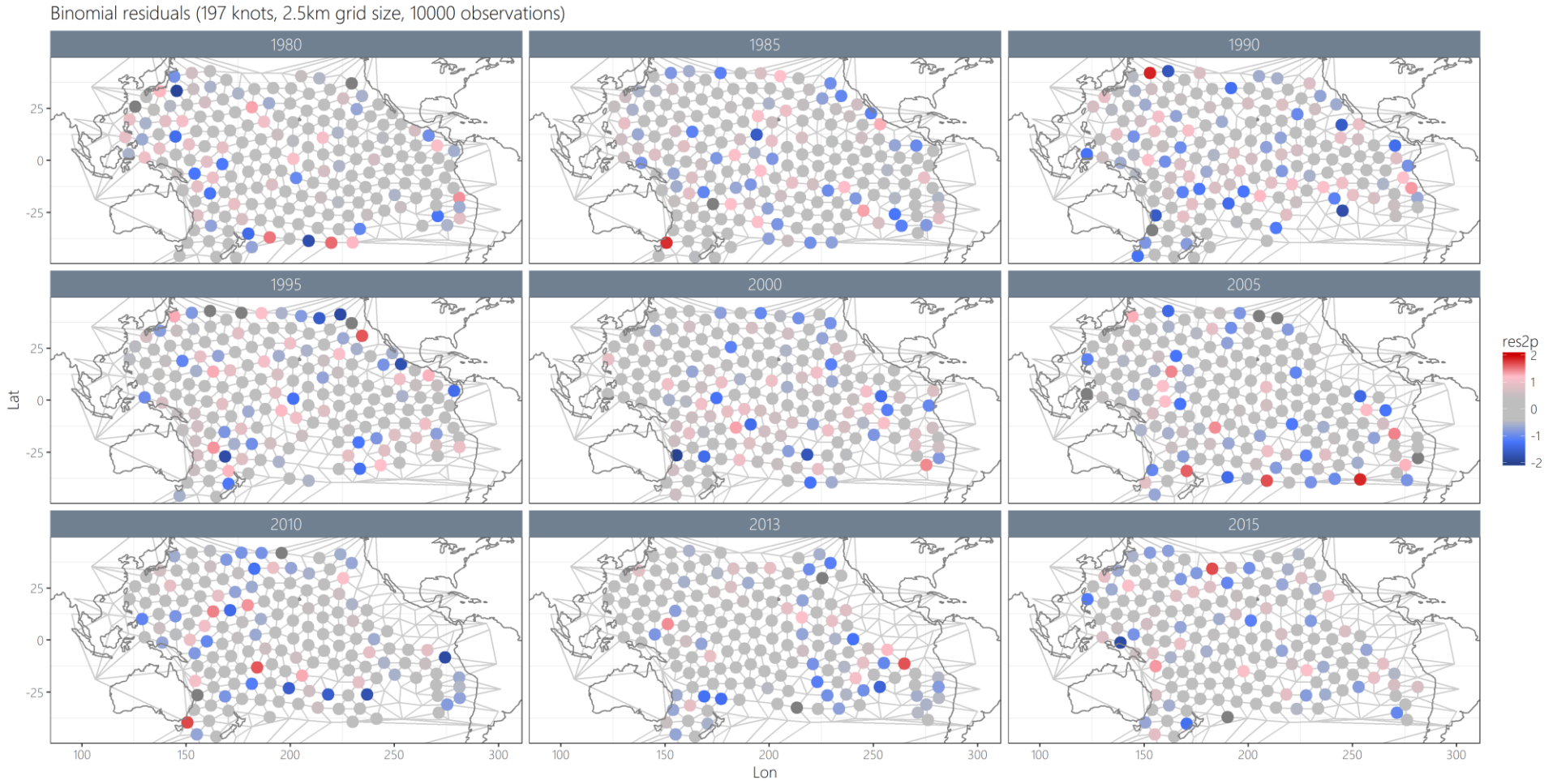


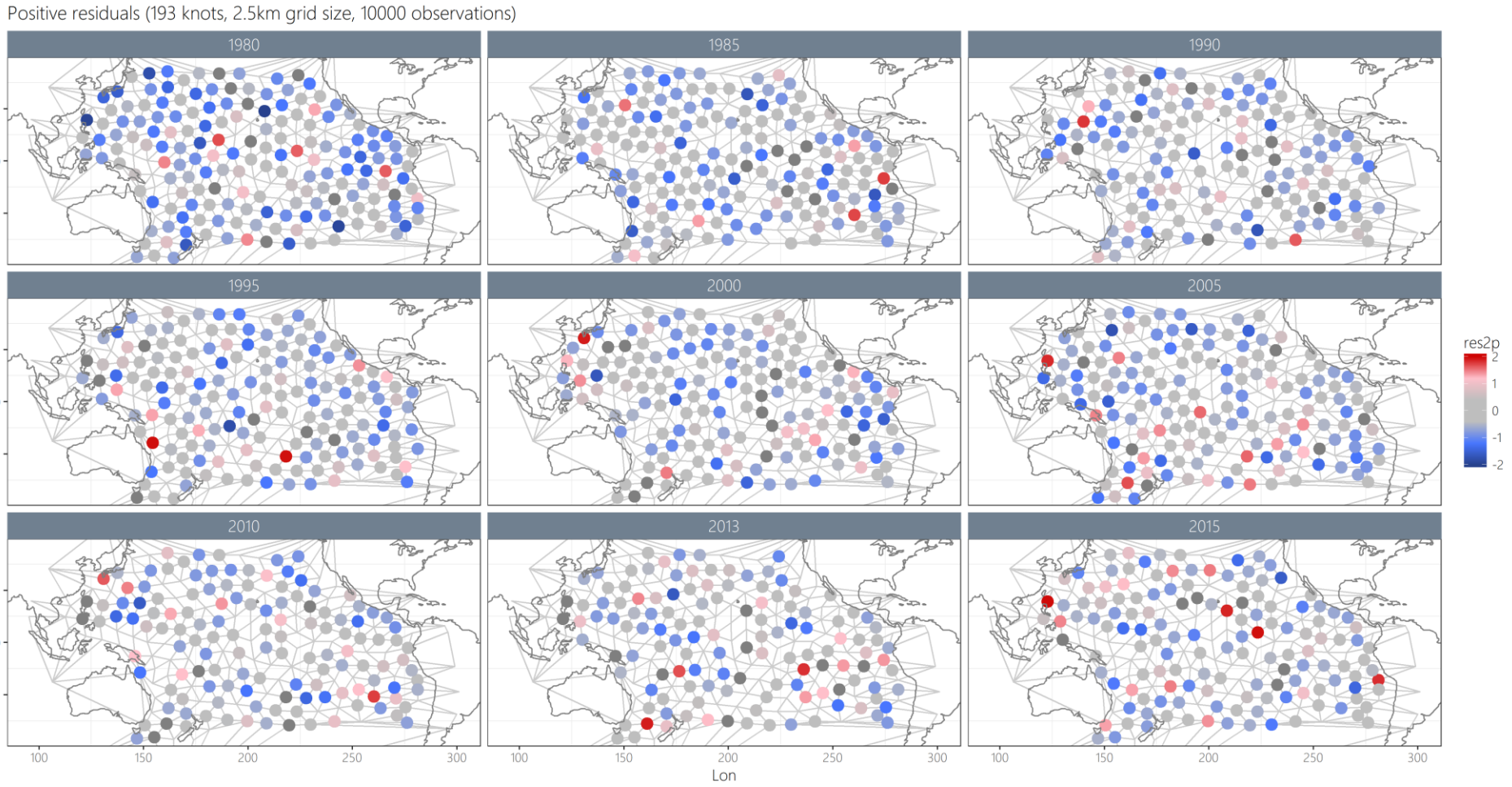

60
40
20
0
−20
−40
100
150
200
250
300

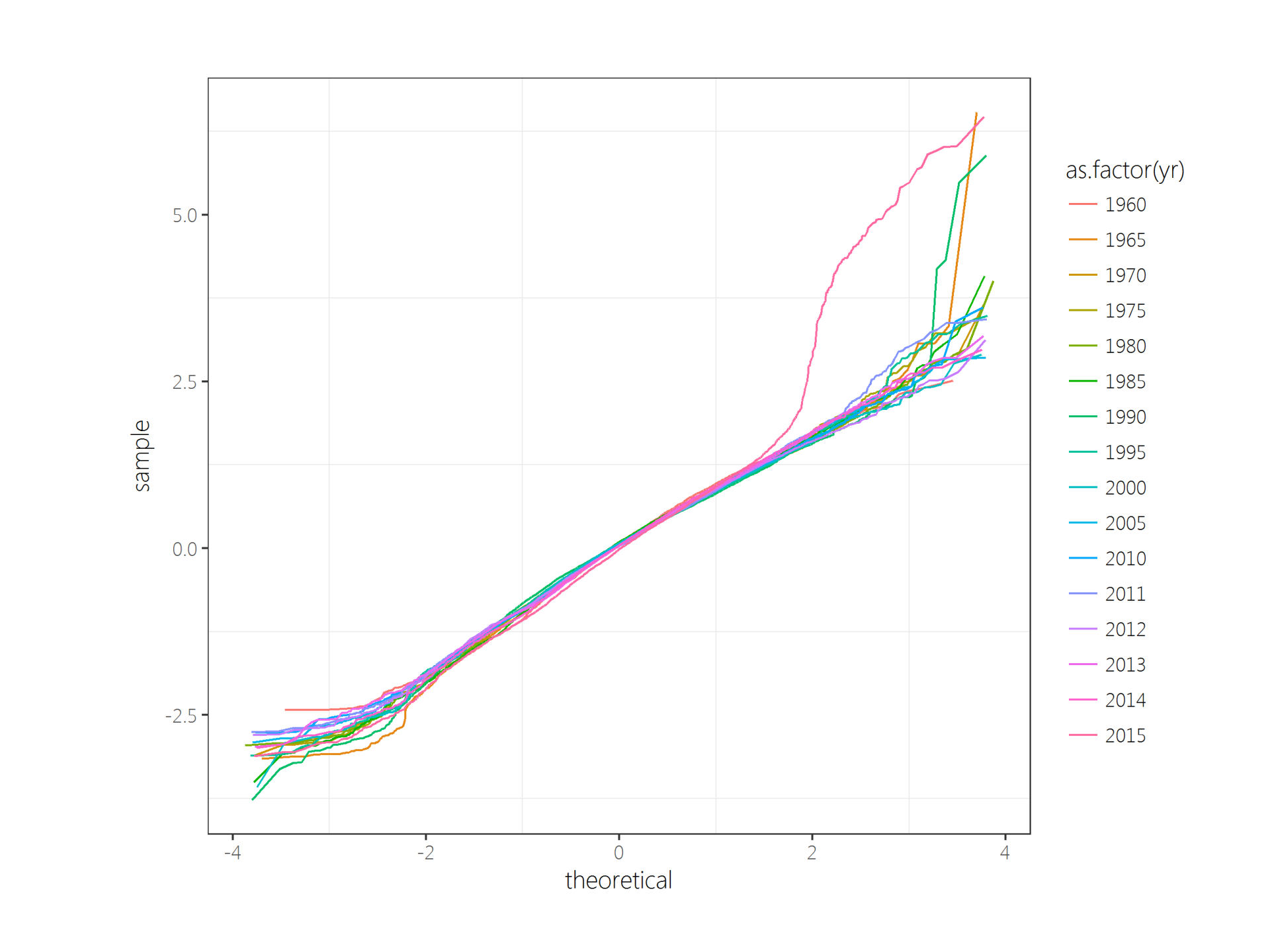

Pacific-wide Yellowfin standardized index
Indx
3
2
1
1960
1980
2000
Year
tag
YFT.abund.sst.spawn
YFT.abund.sstm1

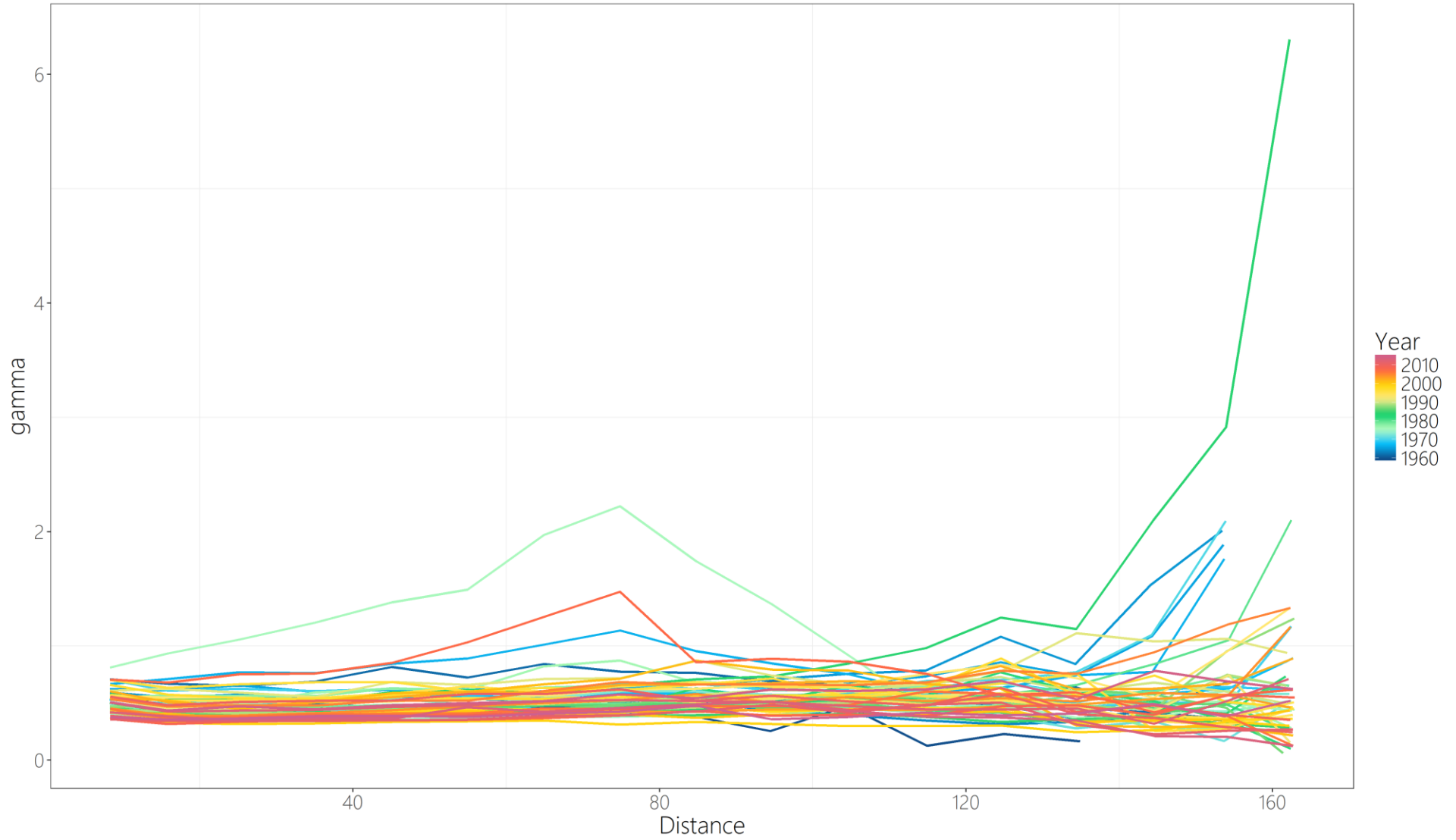
Diagnostics!




Binomial residuals (197 knots, 2.5km grid size, 10000 observations)
1980
1985
1990
1995
2000
2005
2010
2013
2015
Lat
Lon
25
0
-25
100
150
200
250
300
res2p
2
1
0
-1
-2


Positive residuals (193 knots, 2.5km grid size, 10000 observations)
1980
1985
1990
1995
2000
2005
2010
2013
2015
Lat
25
0
-25
Lon
100
150
200
250
300
res2p
2
1
0
-1
-2


sample
5.0
2.5
0.0
-2.5
-4
-2
0
2
4
theoretical
as.factor(yr)
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2011
2012
2013
2014
2015


gamma
6
4
2
0
40
80
120
160
Distance
Year
2010
2000
1990
1980
1970
1960

Spatio-temporal CPUE indices:

Moving from surveys to fisheries dependent: where to next?

Preferential sampling (…over covariates?)

Density vs. catchability covariates (confounding of spatial variables)

Diagnostics and model selection

Mesh/correlation relationships:

rules of thumb via simulation models?