Development of a size-structured spatiotemporal model for invertebrates

Jie Cao, James T. Thorson, André Punt, Cody Szuwalski

CAPAM, La Jolla
Feb. 28, 2018

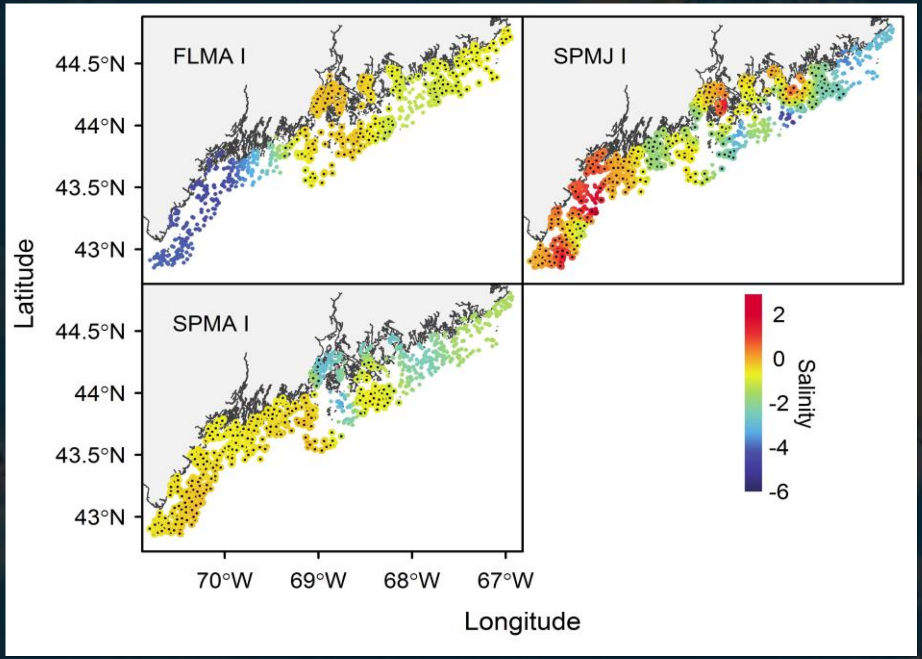
School of Aquatic and Fishery Sciences

Photo: https://thenavigatormagazine.com/dfo-wins-controlling-agreement-legal-challenge/

Spatial scale

The problem of scale is the central problem in ecology

- **Pattern & Process**
- **Statistical relationship**
- **Characteristic scale**



Population dynamic & stock assessment

- Spatial homogeneity

- Tracking total abundance across the entire stock
  - Survey counts/catches are aggregated spatially

- Consequences of ignoring spatial structure
  - Degrading stock assessment performance
  - Leading to overexploitation of weaker population units
  - Ineffective recovery plans

Spatial structured stock assessments

Incorporating Spatial Structure in Stock Assessment: Movement Modeling in Marine Fish Population Dynamics

DANIEL R. GOETHEL,[1] TERRA
[1]School for Marine Science and Technology
[2]Juneau Center, School of Fisheries and Oce

Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth and recruitment? A case study based on pink ling in Australia

André E. Punt[a,b,*], Malcolm Haddon[b], Geoffrey N. Tuck[b]

Journal of Animal Ecology 2017, **86**, 888–898 doi: 10.1111/1365-2656.12678

Fine-scale population dynamics in a marine fish species inferred from dynamic state-space models

Lauren A. Rogers[*,1,2], Geir O. Storvik[3], Halvor Knutsen[1,4,5], Esben M. Olsen[1,4,5] and Nils C. Stenseth[1,4,5]

Spatial structured stock assessments

- Spatial strata
  - few sub-stocks with connectivity

- Increasing the # of spatial strata?
  - very little data for each stratum
  - difficulties of estimating movement rates
  - Linkage among strata

Objectives

<u>Developing a spatiotemporal population model</u>

- *fine spatial scale*
- *geostatistical approach*
- *size-structured*

- *spatial variation*
  - *density*
  - *fishing mortality*
  - *catch*

❖ better interpret population dynamic

❖ improve spatial management

Spatiotemporal population model

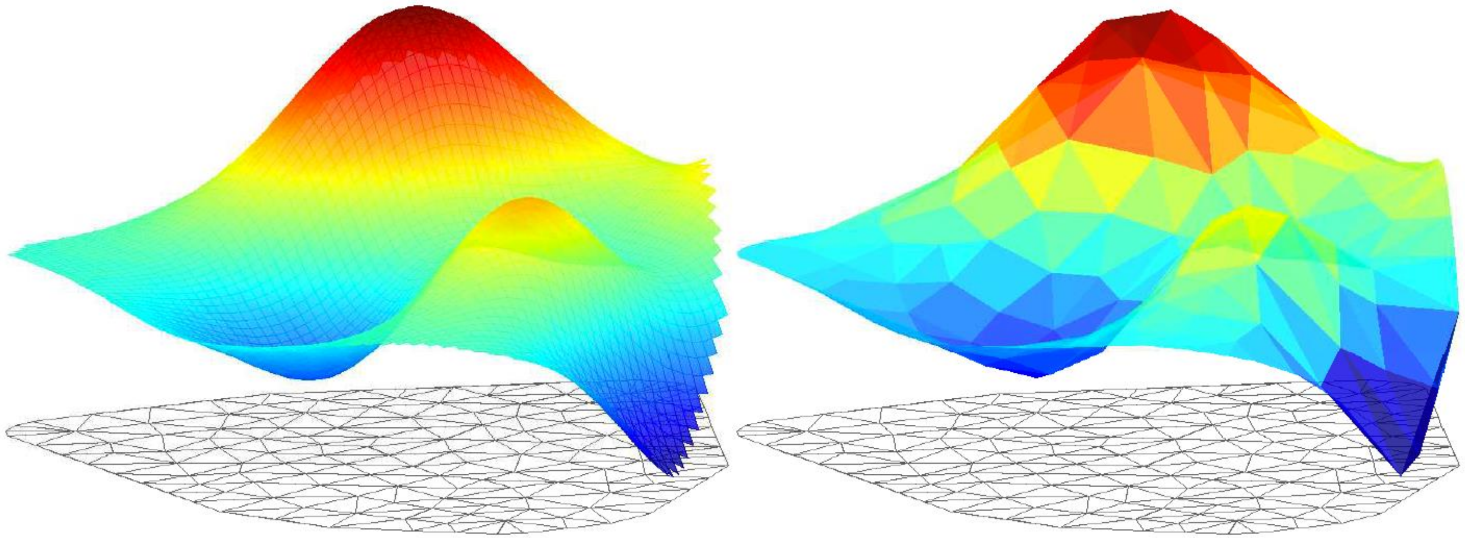
- Combines theory and methods from population dynamics and geostatistics
- Assume population density varies continuously across space

$$x(s_i) \sim N\left(\frac{1}{|n_i|}\sum_{j \epsilon n_i} x(s_j), \sigma^2\right)$$

- Joint distribution for density at all locations
- Expand to account for size-structured population dynamics

Gaussian Markov random field (GMRF)

- Continuous spatial process -> discretely indexed GMRF
- Matérn covariance function
- Mesh/knot



Thorson, J.T., Shelton, A.O., Ward, E.J. and Skaug, H.J., 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science*, *72*(5), pp.1297-1310.

Why size-structured models?

- Advantages:
  - Requires no ability to age animals (shrimps, crabs, lobsters)
  - Uses the data actually available
  - Vulnerability / maturity are often functions of size and not age

Abundance at size (n) for a given location s and time t

$$\boldsymbol{n}_{s,t+1} = f(\boldsymbol{n}_{s,t}) \circ e^{\boldsymbol{\varepsilon}_{s,t}}$$

$$\boldsymbol{\Sigma}_t \sim \text{MVN}(0, \mathbf{R}_{spatial} \otimes \boldsymbol{\Theta}_{\boldsymbol{L}})$$

$$f\left(\boldsymbol{n}_{s,t}^{\text{male}}\right) = \begin{cases} \boldsymbol{r}_{s,t} * p_{\text{male}} + \mathbf{G}(\boldsymbol{n}_{s,t-1}^{\text{immat}} e^{-\boldsymbol{m}_{s,t-1} - \boldsymbol{v} * f_{s,t-1}^{\text{male}}}) * (1 - \boldsymbol{w}), & \boldsymbol{n} = \boldsymbol{n}^{\text{immat}} \\ \mathbf{G}(\boldsymbol{n}_{s,t-1}^{\text{immat}} e^{-\boldsymbol{m}_{s,t-1} - \boldsymbol{v} * f_{s,t-1}^{\text{male}}}) * \boldsymbol{w} + \boldsymbol{n}_{s,t-1}^{\text{mat}} e^{-\boldsymbol{m}_{s,t-1} - \boldsymbol{v} * f_{s,t-1}^{\text{male}}}, & \boldsymbol{n} = \boldsymbol{n}^{\text{mat}} \end{cases}$$

$$f\left(\boldsymbol{n}_{s,t}^{\text{female}}\right) = \begin{cases} \boldsymbol{r}_{s,t} * p_{\text{female}} + \mathbf{G}(\boldsymbol{n}_{s,t-1}^{\text{immat}} e^{-\boldsymbol{m}_{s,t-1}}) * (1 - \boldsymbol{w}), & \boldsymbol{n} = \boldsymbol{n}^{\text{immat}} \\ \mathbf{G}(\boldsymbol{n}_{s,t-1}^{\text{immat}} e^{-\boldsymbol{m}_{s,t-1}}) * \boldsymbol{w} + \boldsymbol{n}_{s,t-1}^{\text{mat}} e^{-\boldsymbol{m}_{s,t-1}}, & \boldsymbol{n} = \boldsymbol{n}^{\text{mat}} \end{cases}$$

$$\boldsymbol{r}_{L,t} \sim \text{MVN}(r_{\mu}, \mathbf{R}_{spatial})$$

$$\boldsymbol{c}_{s,t} = \left(1 - e^{-\boldsymbol{v} * f_{s,t}^{\text{male}}}\right) * \boldsymbol{n}_{s,t} e^{-\boldsymbol{m}_{s,t}}$$

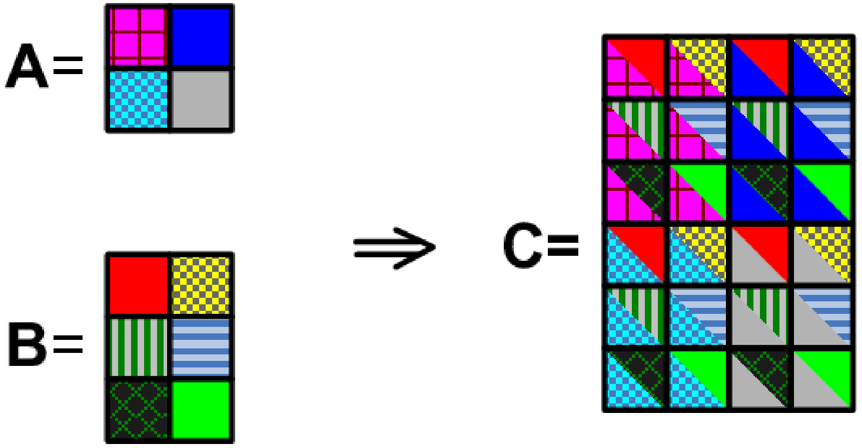
$$\boldsymbol{n}_{s,t+1} = f(\boldsymbol{n}_{s,t}) \circ e^{\boldsymbol{\varepsilon}_{s,t}}$$

$$\boldsymbol{\Sigma}_t \sim \mathrm{MVN}(0, \mathbf{R}_{spatial} \otimes \boldsymbol{\Theta}_L)$$

$\circ$	Hadamard product (entrywise product)
s	location
t	year
$\otimes$	Kronecker product

$\boldsymbol{n}_{s,t}$	vector of abundances for each of l size classes
$f()$	function representing population dynamic
$\boldsymbol{\varepsilon}_{s,t}$	vector of random effects (process error)
$\boldsymbol{\Theta}_L$	covariance among size classes (l by l matrix **L**)
$\mathbf{R}_{spatial}$	spatial covariance matrix (covariance between 2 locations follows a Matern function)

Kronecker product



Imagine 100 knots and 30 size classes !

$f(\boldsymbol{n}_{s,t})$ - population dynamic

Snow crab

- Male/Female
- Only males are retained in the fishery
- Split into maturity state
- Mature individuals do not molt

Population dynamic ($f()$)

$$f\left(\boldsymbol{n}_{s,t}^{\text{male}}\right) = \begin{cases} \boldsymbol{r}_{s,t} * p_{\text{male}} + \mathbf{G}(\boldsymbol{n}_{s,t-1}^{\text{immat}} e^{-\boldsymbol{m}_{s,t-1} - \boldsymbol{v} * f_{s,t-1}^{\text{male}}}) * (1 - \boldsymbol{w}), & \boldsymbol{n} = \boldsymbol{n}^{\text{immat}} \\ \mathbf{G}(\boldsymbol{n}_{s,t-1}^{\text{immat}} e^{-\boldsymbol{m}_{s,t-1} - \boldsymbol{v} * f_{s,t-1}^{\text{male}}}) * \boldsymbol{w} + \boldsymbol{n}_{s,t-1}^{\text{mat}} e^{-\boldsymbol{m}_{s,t-1} - \boldsymbol{v} * f_{s,t-1}^{\text{male}}}, & \boldsymbol{n} = \boldsymbol{n}^{\text{mat}} \end{cases}$$

$\boldsymbol{r}_{s,t}$	vector of recruitment for each of l size classes
p_{male}	proportion of male recruitment
$\mathbf{G}$	growth transition matrix
$\boldsymbol{m}_{s,t}$	vector of natural mortality at location s, year t
$f_{s,t}^{male}$	fishing mortality at location s, year t
$\boldsymbol{v}$	vector of selectivity at size
$\boldsymbol{n}_{s,t}^{\text{immat}}$	vector of immature abundance for each of l size classes
$\boldsymbol{n}_{s,t}^{\text{mat}}$	vector of mature abundance for each of l size classes
$\boldsymbol{w}$	vector of maturity of each size class

Population dynamic - *parameters*

Recruitment

$$\boldsymbol{r}_{s,t} = r^{u}_{s,t} * \boldsymbol{l}_{size}$$

$\boldsymbol{r}_{s,t}$	– vector of recruitment for each of p size classes
$\boldsymbol{l}_{size}$	– vector of proportion of recruitment
$r^{u}_{s,t}$	– recruitment at location *s* and year *t*

r^{u}_{t} follows a spatial process $\sim MVN\left(\boldsymbol{\mu}_r, \mathbf{Q}_r^{-1}\right)$

Fishing mortality

$$f_{l,s,t} = f_{s,t}\, v_l$$

$f_{s,t}$	– f at location *s* and year *t*
v_l	– selectivity of size class *l*

$$v_l = \frac{1}{1 + e^{\left(-k(L_p - L_{50})\right)}}$$

$f_{s,t} \mid f_{s,t-1} \sim N(f_{s,t-1}, \sigma_f^2)$ random effect

Growth transition matrix (**G**) and natural mortality (**m**) – input data

Summary of parameters

Fixed effects

Θ_L	process error covariance (among size classes)
κ	geostatistical range for correlations
μ_t	average offset of annual recruitment
φ	initial abundance of each size class
s	parameters of selectivity (logistic)
	Parameters of observation model

Random effects

r_t^u	spatial variation in recruitment
n_t	spatial variation in density for each size class and year
f	fishing mortality of location s over time

<u>treat density as random, rather than process errors (ε_t)</u>

Input data

survey data

Size_class	Year	Catch_N	AreaSwept_km2	Vessel	Lat	Lon
1	1	553	3.1	0	60	-174
1	1	629	3.1	0	63.5	-172
1	1	575	3.1	0	58	-170
1	1	618	3.1	0	61.5	-178
1	1	625	3.1	0	64.5	-170
1	1	634	3.1	0	61	-172

- used to create mesh/knots

commercial catch data

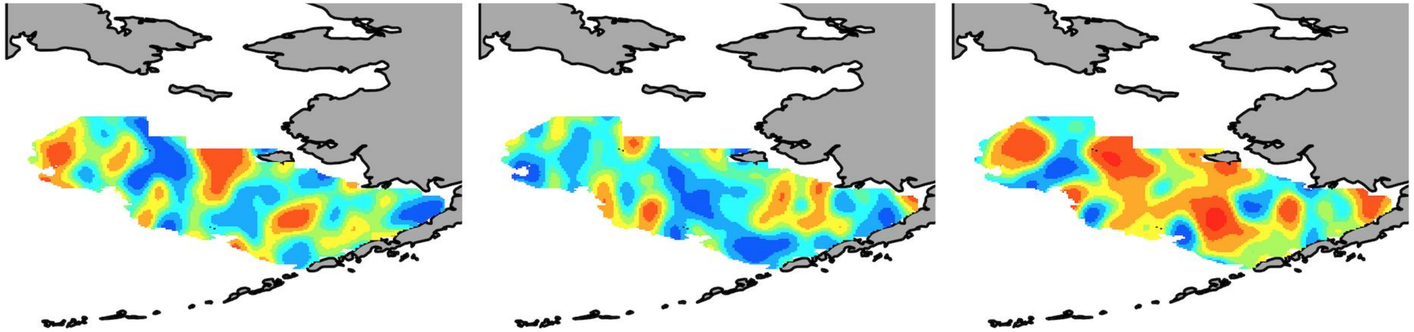
X	lat	lon	year	X.1	X.2	X.3	X.4	X.5
1	55.6	-169	1	0.802	2.82	2.32	3.18	7.18
2	56.3	-170	1	0.657	1.83	1.54	2.15	4.94
3	56.3	-170	1	0.662	1.82	1.54	2.16	4.96
4	56.2	-171	1	0.64	1.78	1.5	2.1	4.81
5	56.2	-170	1	0.645	1.8	1.51	2.12	4.85
6	56.2	-170	1	0.646	1.82	1.52	2.13	4.88

- fine scale
- aggregated to knot-level

Model outputs

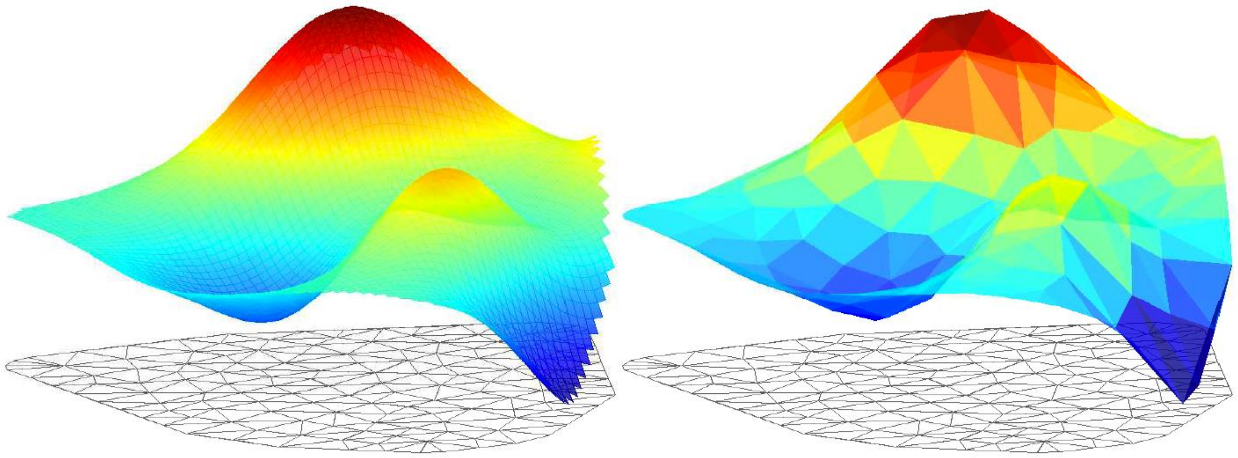
- Predicted population density map
- Estimated fishing mortality map
- Predicted catch map
- Estimated covariance of process error

0.219	0.063	0.056	-0.007
0.063	0.105	0.034	-0.002
0.056	0.034	0.178	-0.003
-0.007	-0.002	-0.003	0.021



Estimation

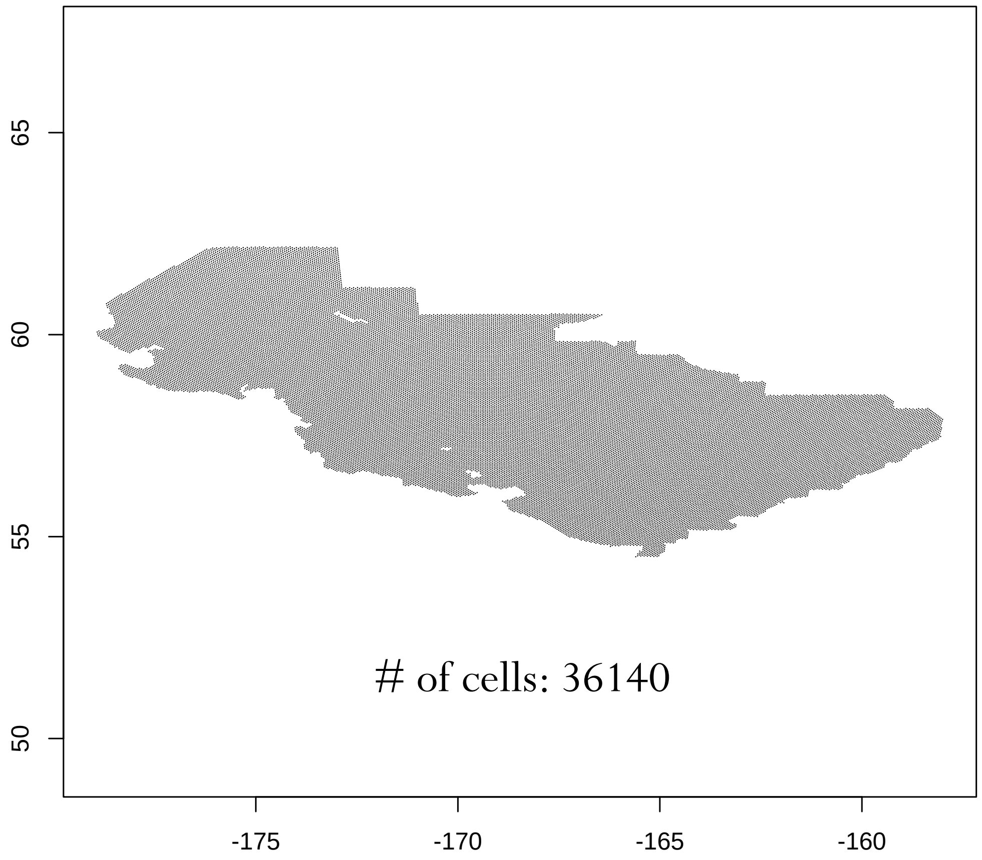
- SPDE – MVN
- Piecewise constant
- Catch – lognormal
- Survey – lognormal/Poisson-link



Template Model Builder (TMB)

Operating model – overview

- Dynamics occur at fine scale
- Population dynamics (non-spatial) formulated identically to EM
- Cell-specific parameters (spatially correlated)
- No movement
- Annual time step



Operating model – recruitment

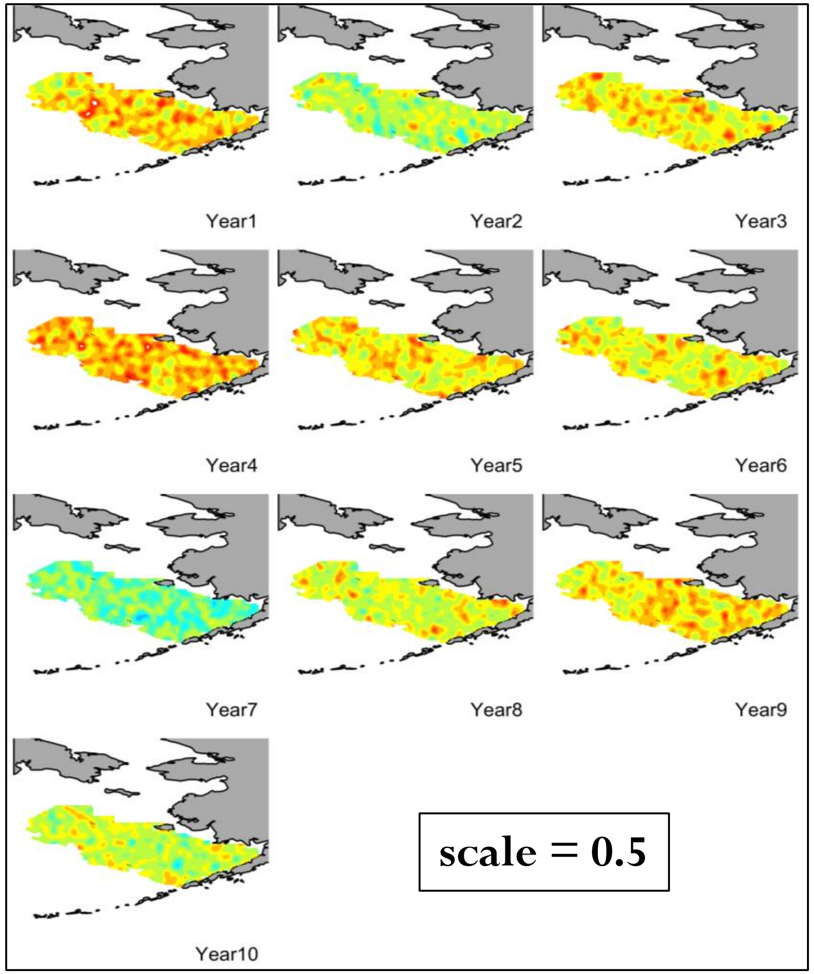
1. Draw average annual recruitments (μ_t) from a Poisson distribution
2. Define spatial variance and scale (σ_t^2, κ_t) for each year

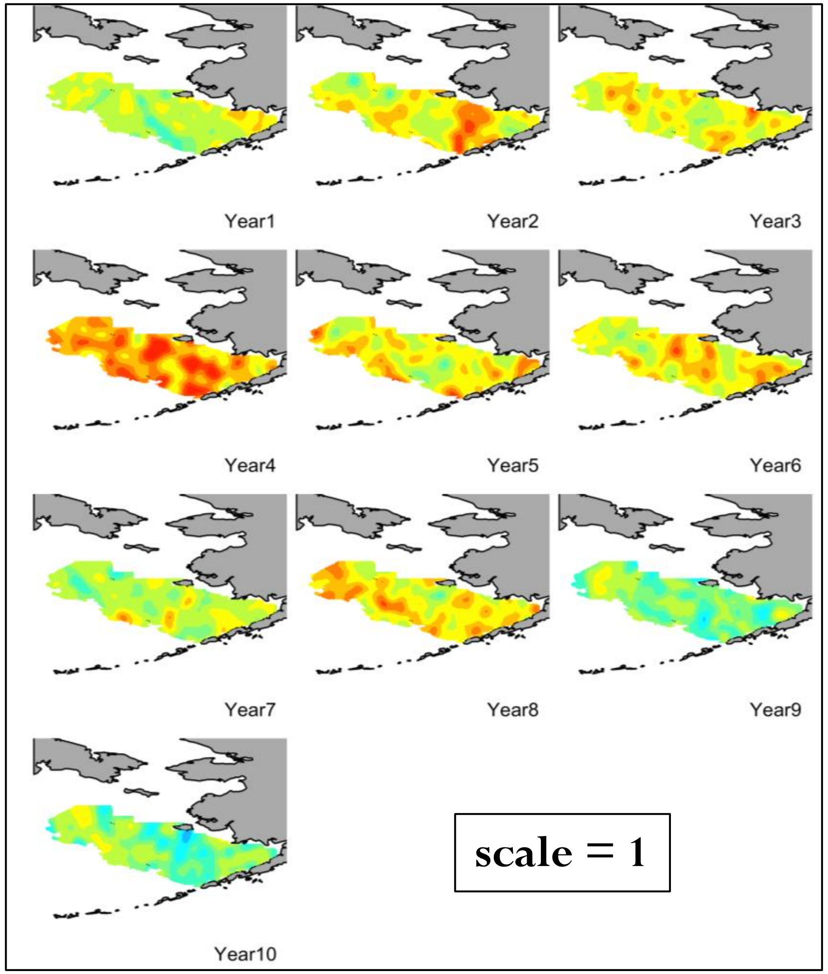
   model_R <- RMgauss(σ_t^2, κ_t)
3. Simulate a Gaussian random field for each year on the grid (ε_t)

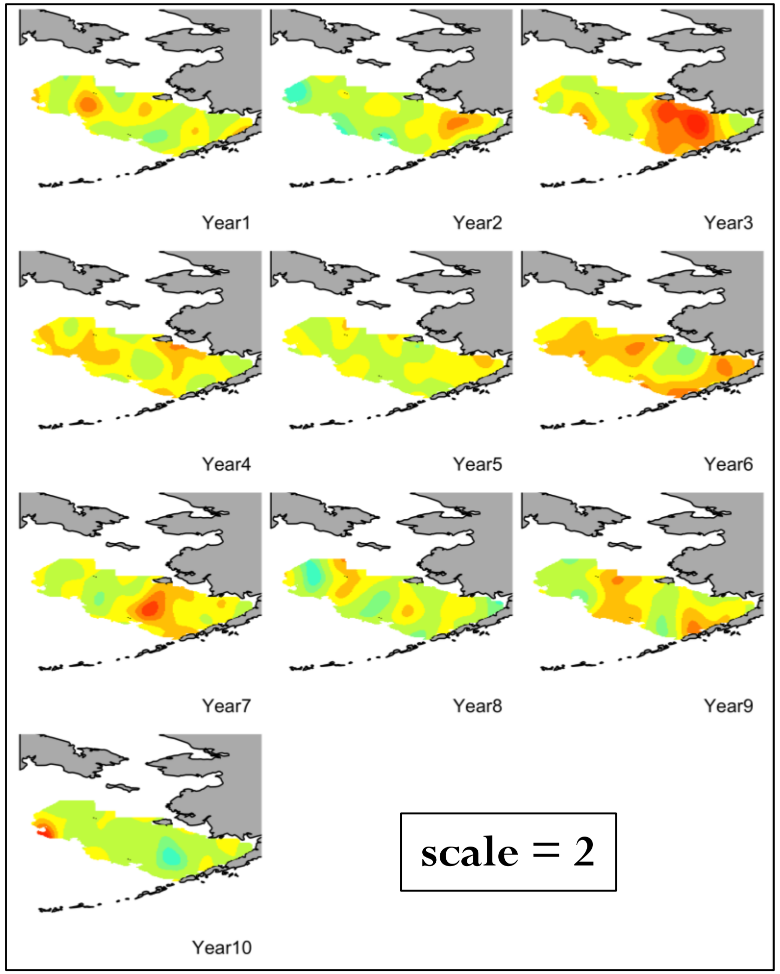
   ε_t<- RFsimulate(model = smodel_R, x=loc_x[,1], y=loc_x[,2])
4. Calculate recruitment of each cell *s* and year *t*, $R_{s,t} = \mu_t e^{\varepsilon_t}$
5. Allocate recruitment $R_{s,t}$ to each size class

Operating model
– recruitment examples

# of size classes (population):	n_p = 5
# of years:	n_t = 10
# of size bins (recruitment):	n_r = 1







Operating model – fishing mortality

- Similar way as simulating recruitment ($f_{s,t} = f_t e^{\varepsilon_t}$)
- Selectivity (*s*) - Logistic function (2 parameters)
- Fishing mortality $f_{p,s,t} = f_{s,t}\ v_l$
- Flexibility in f_t and ε_t
- Different parameterization in EM

Operating model – growth

- EM uses growth transition matrix (GTM) directly
- Two options of calculating GTM
    1. 5-parameter VBGF (Chen et al. 2003)
    2. Linear relationship between pre- and post-molt length, gamma function (snow crab stock assessment report)
- Spatial dependence – parameters of growth function

Calculating GTM – VBGF

The distribution of the growth increment is assumed to be normal with mean, $E(\Delta L_k)$, and variance, $Var(\Delta L_k)$, calculated as

$$E(\Delta L_k) = (L_\infty - L_k)(1.0 - e^{-K})$$

$$Var(\Delta L_k) = \sigma_{L_\infty}^2(1 - e^{-K})^2 + (L_\infty - L_k)^2\sigma_K^2 e^{-2K} + 2\rho_b\sigma_{L_\infty}\sigma_K(1 - e^{-K_b})(L_\infty - L_k)e^{-K}$$

L_∞, K, σ_{L_∞}, σ_K, and the correlation between L_∞ and K (ρ_b) are the parameters

The probability of growing from length class *k* to length class *k+1*, $Pp_{k\rightarrow k+1}$, is calculated as:

$$Pp_{k\rightarrow k+1} = \int_{low}^{up} norm\big(E(\Delta L_k), Var(\Delta L_k)\big)$$

Calculating GTM – linear relationship

For crab that do molt, growth is modeled as a linear function to estimate the mean width after molting given the mean width before molting:

$$L_{k+1} = \text{int} + \text{slope} * L_k$$

The probability of growing from length class *k* to length class *k+1*, $Pp_{k \to k+1}$, is calculated as:

$$Pp_{k \to k+1} = \int_{low}^{up} gamma\left(\frac{L_{k+1}}{\alpha}, \beta\right)$$

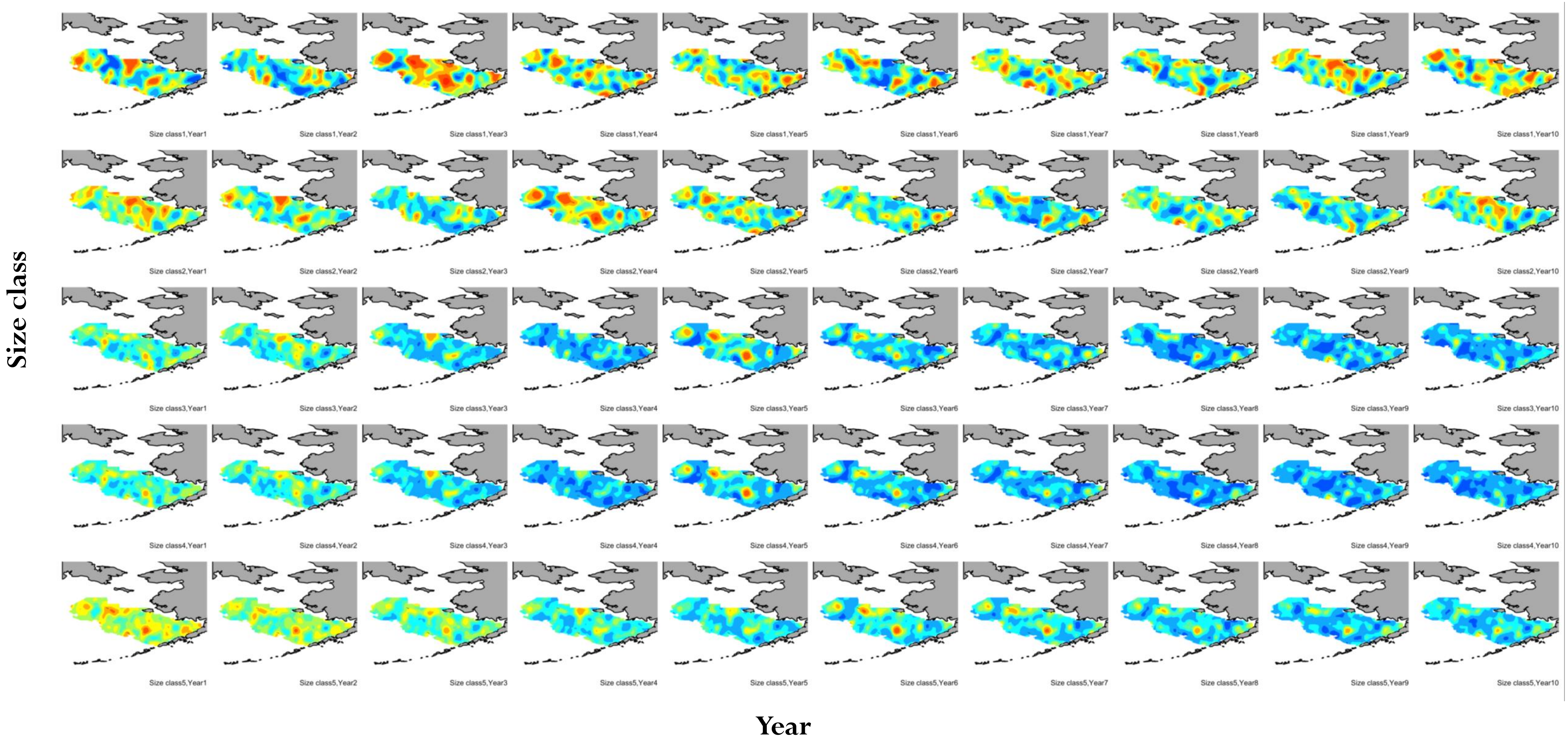
Operating model – a simulated population of snow crab

Item	Descriptor	Note
Years covered	10	
Number sexes	2	Female/Male
Lengths	25-125 mm	
Length bins	20 mm	5 size classes
Recruitment length bin	size class 1	sex ratio = 0.5
Natural mortality	0.23	constant across space over time
Growth	intercept = 1; slope = 1.5; beta = 0.5	constant across space over time
Commercial selectivity	Logistic	logistic (k=0.05; L_{50}=70)
Survey	1	beginning of the year; catchability = 1 selectivity =1 for all size

Spatial variations

Item	Descriptor	Note
Initial condition		50-year burn-in period
Fishing mortality	mean F_t = 0.5 SD F_t = 0.1 var ε_t = 0.1 scale ε_t = 2	
Recruitment dynamics	mean μ_t= 1e6/n_s Var ε_t = 0.1 Scale ε_t = 2	n_s = 36140 No functional relationship with SSB

Simulated population density



Model testing – link OM with EM

- no sampling error
- it took 12 hours !
- not converged !!

What we found…

- too many parameters
  - covariance of process error (1000×1000 matrix)

- estimated recruitment had almost no spatial variation
  - $\boldsymbol{l}_{size} = \mathrm{c}(1, 0, 0, 0, 0)$
  - r_t^u follows a spatial process $\sim MVN\left(\boldsymbol{\mu}_r, \mathbf{Q}_r^{-1}\right)$
  - $\varepsilon_t \sim MVN(0, \boldsymbol{R}_{spatial} \otimes \boldsymbol{\Theta}_L)$

What we changed…

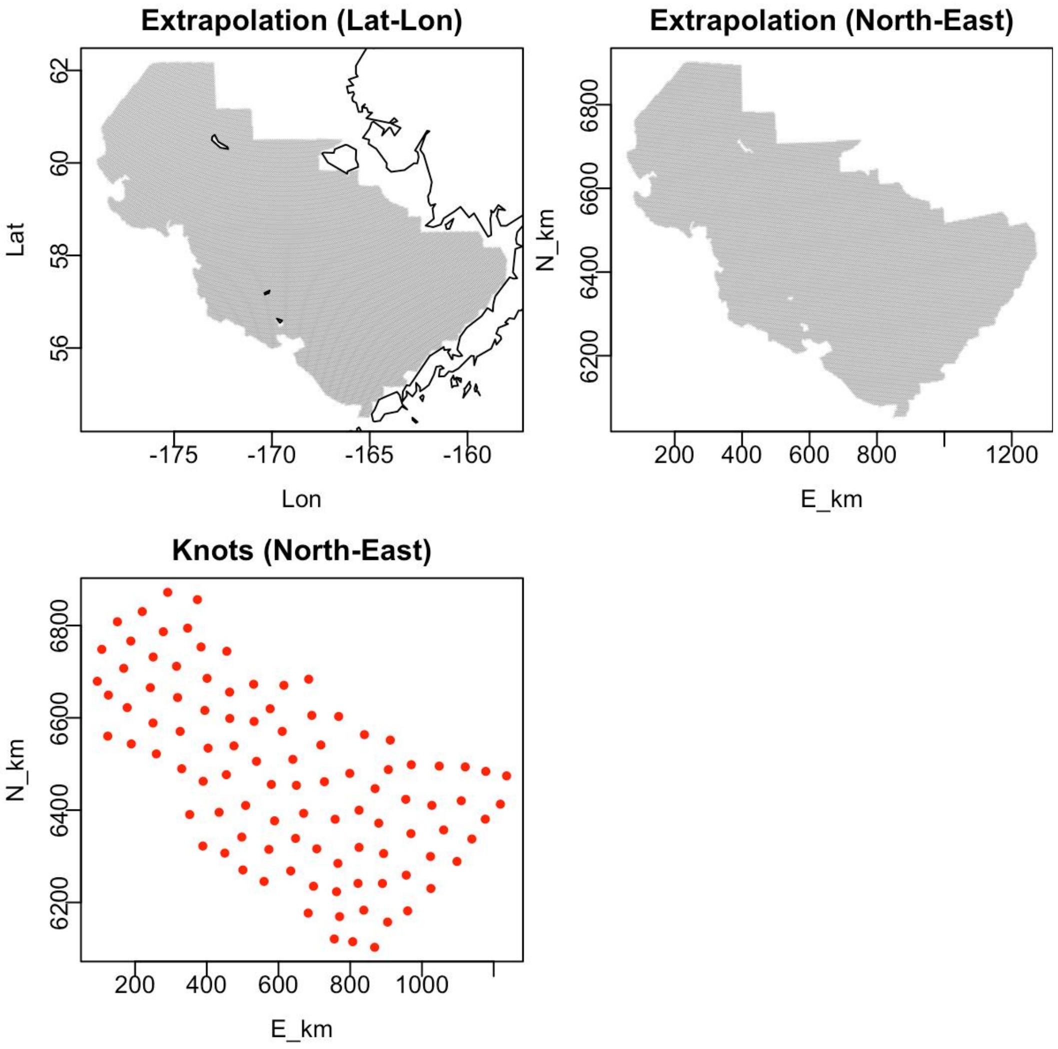
- reduce the dimension - only tracking males
  - only males are retained in the fishery

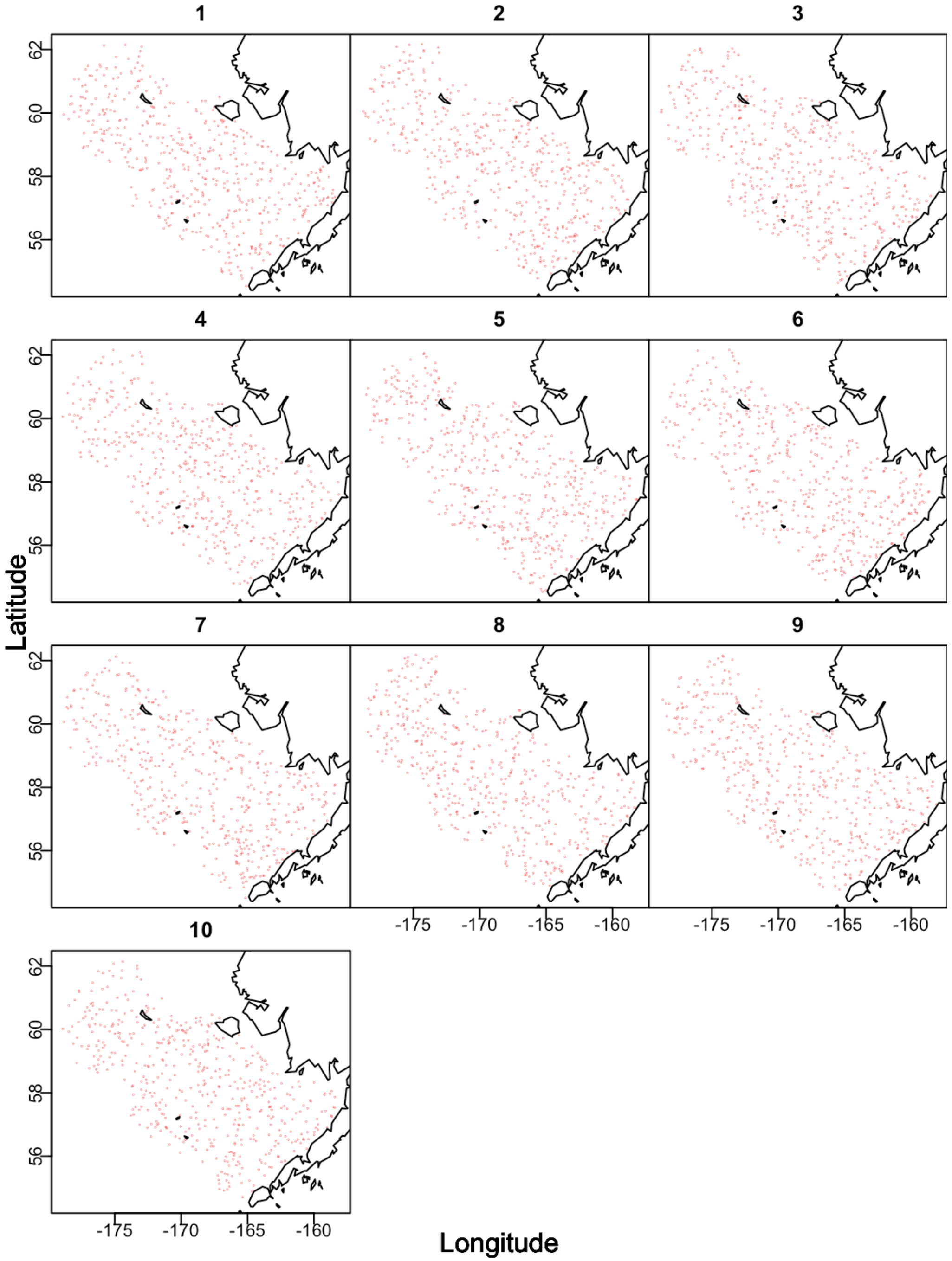
- turn off ε_t (recruitment)
  - only estimate a set of total recruitments (i.e., r_t)

Model testing – link OM with EM

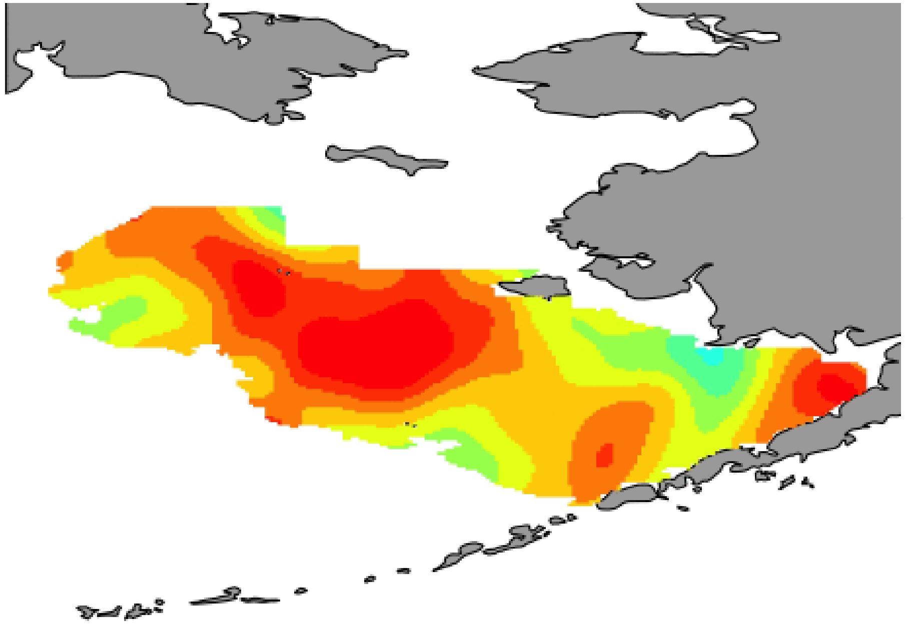
- no sampling error
- it only took 1 hour !
- converged !!

Simulated data

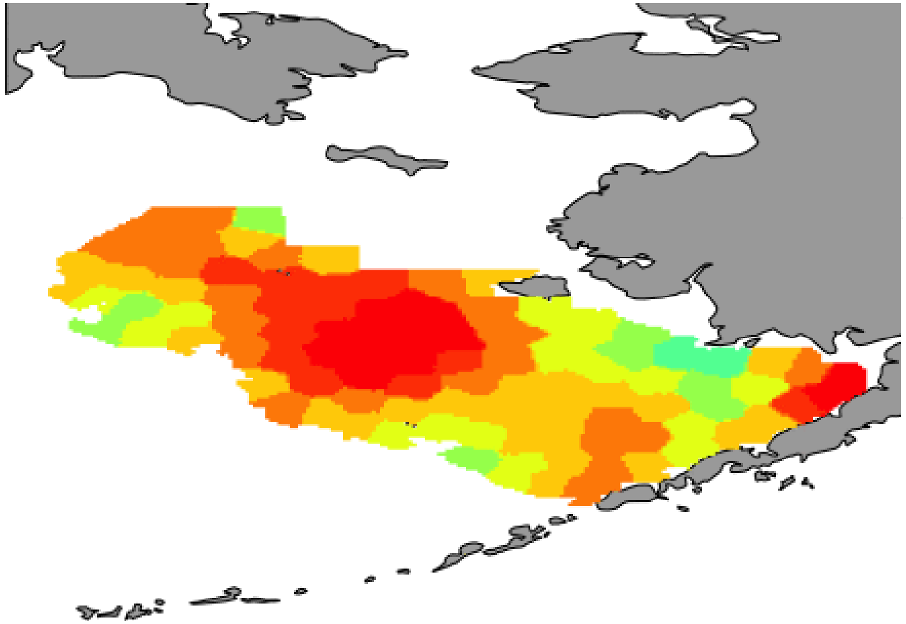




Simulation vs. Estimation

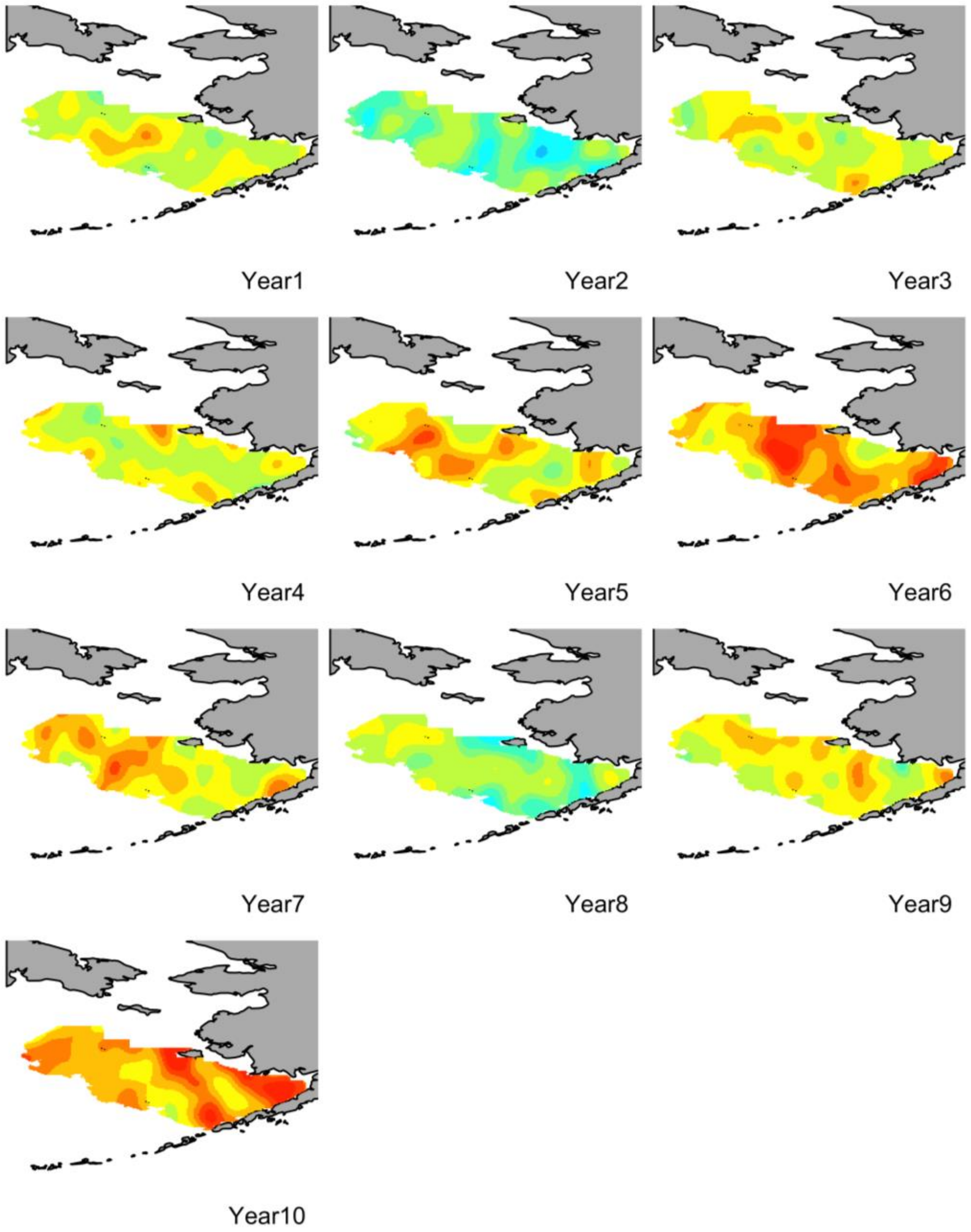


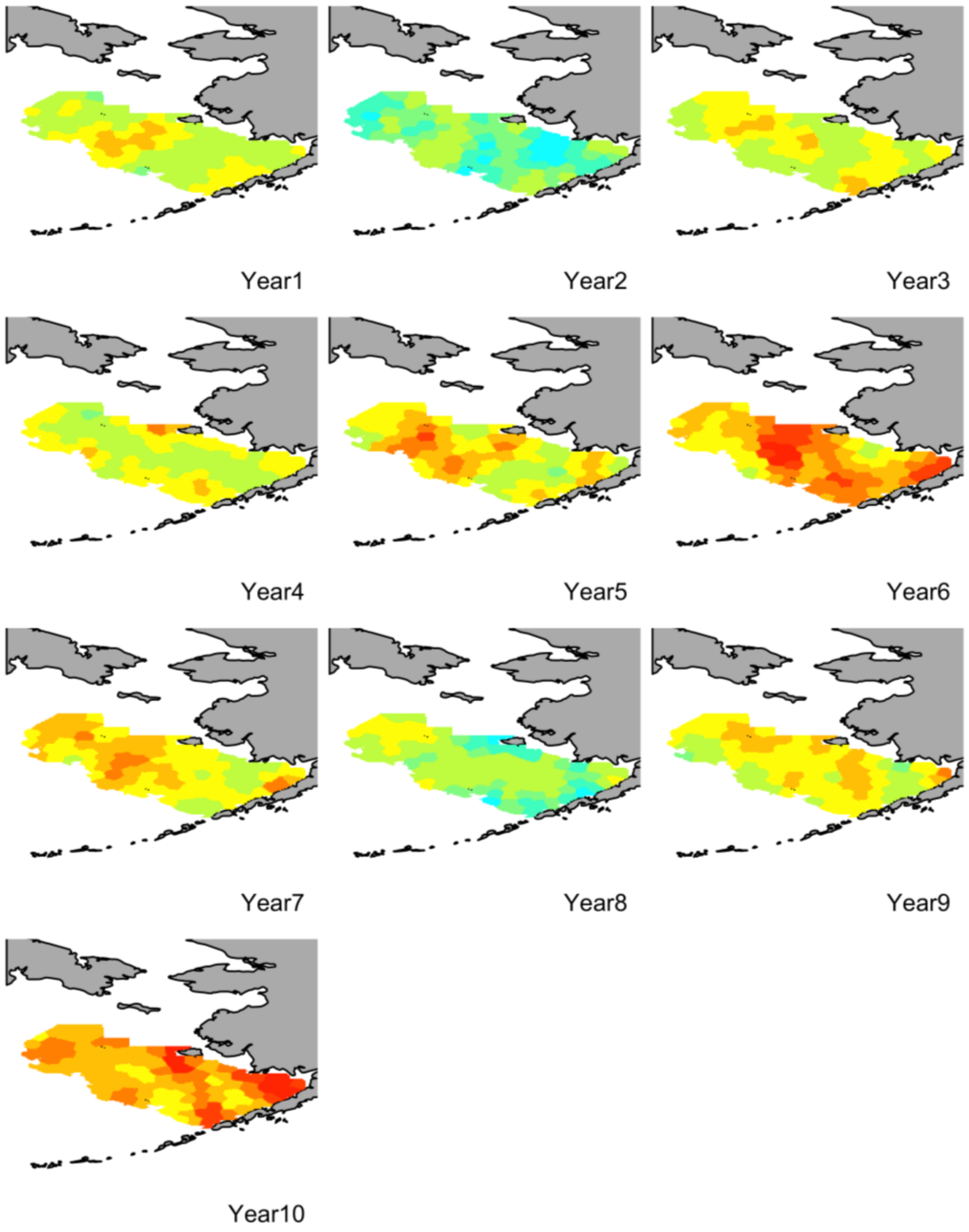
Density of all size classes-Year1



Density of all size classes-Year1

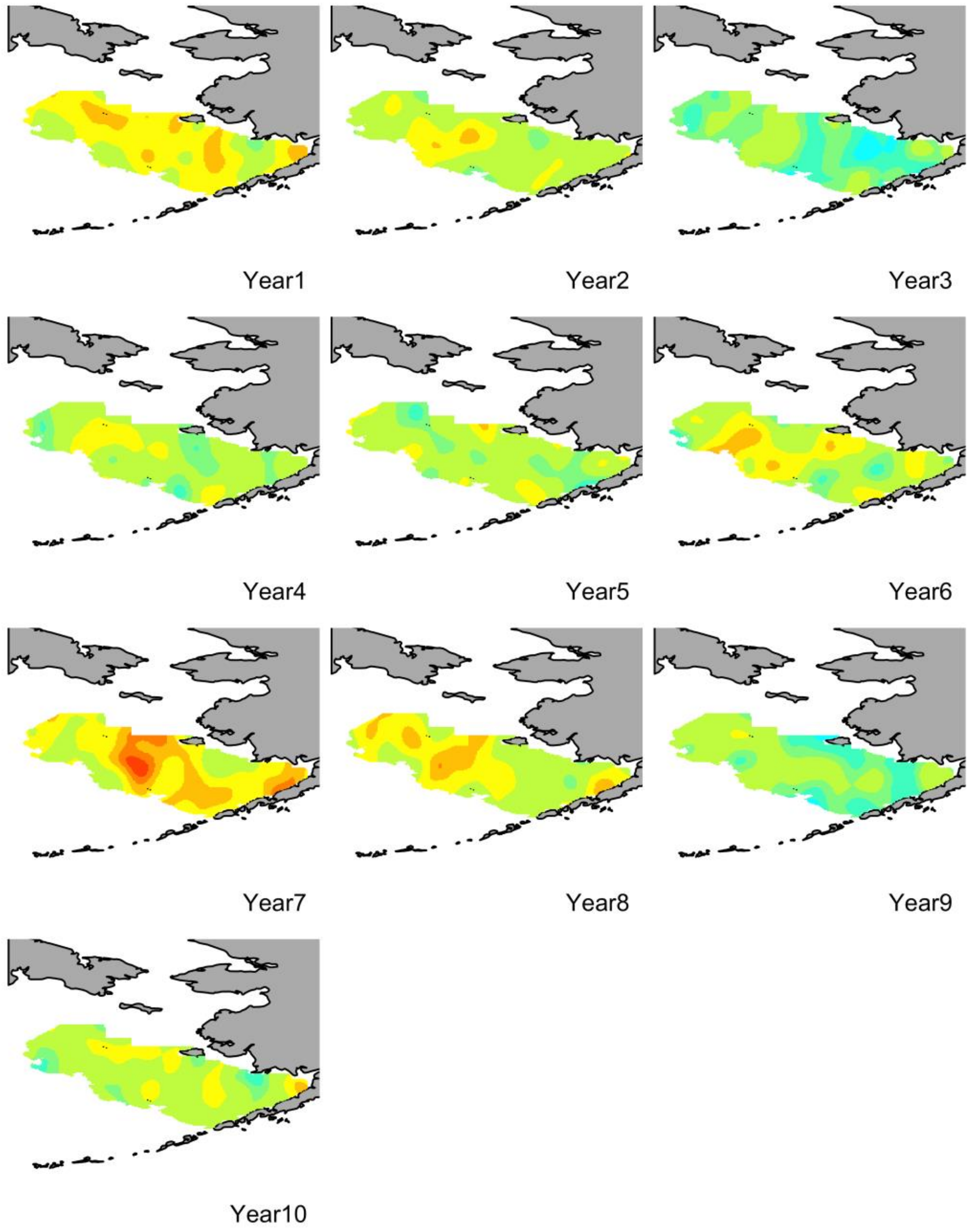
Simulation – size class 1 (recruitment)



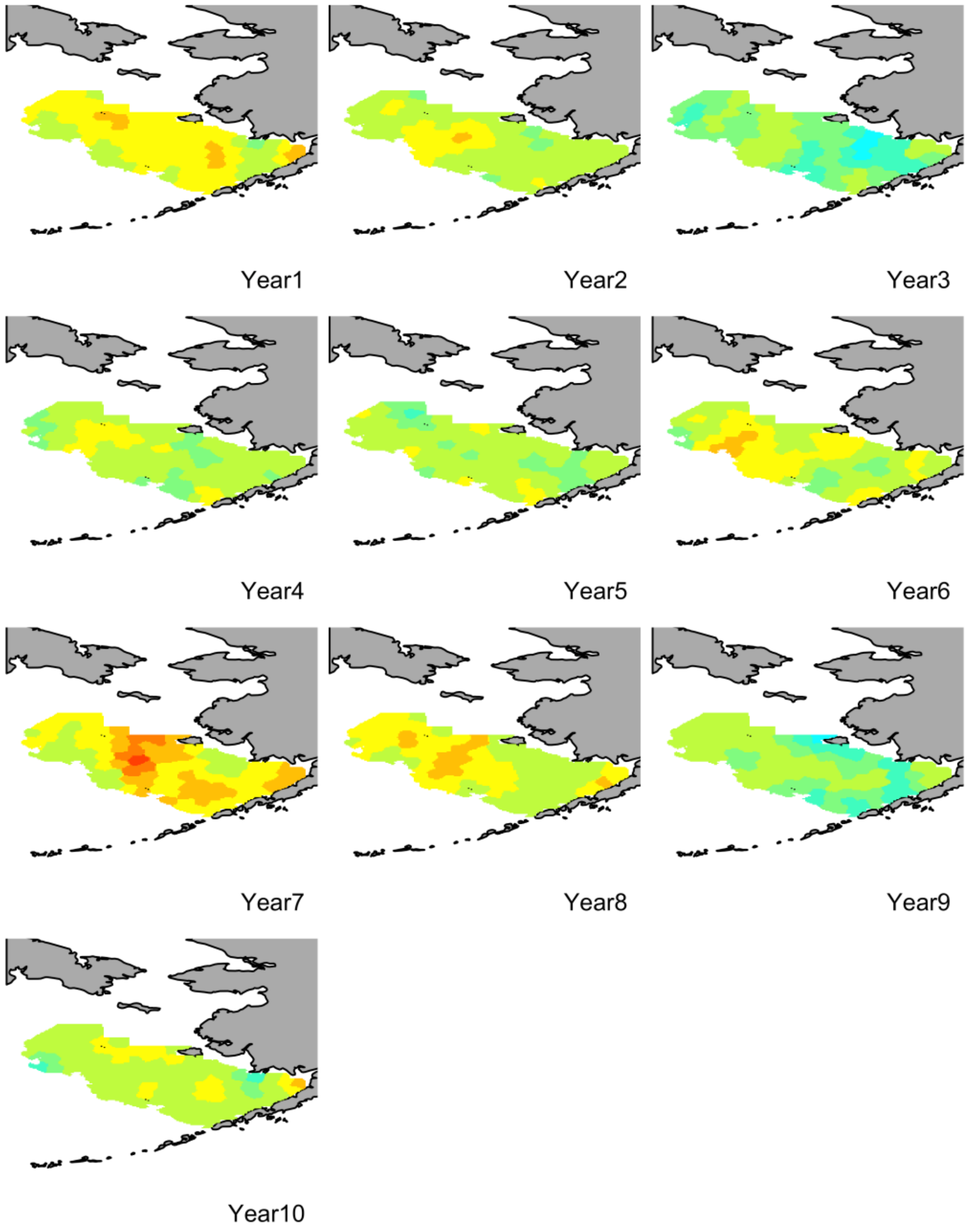


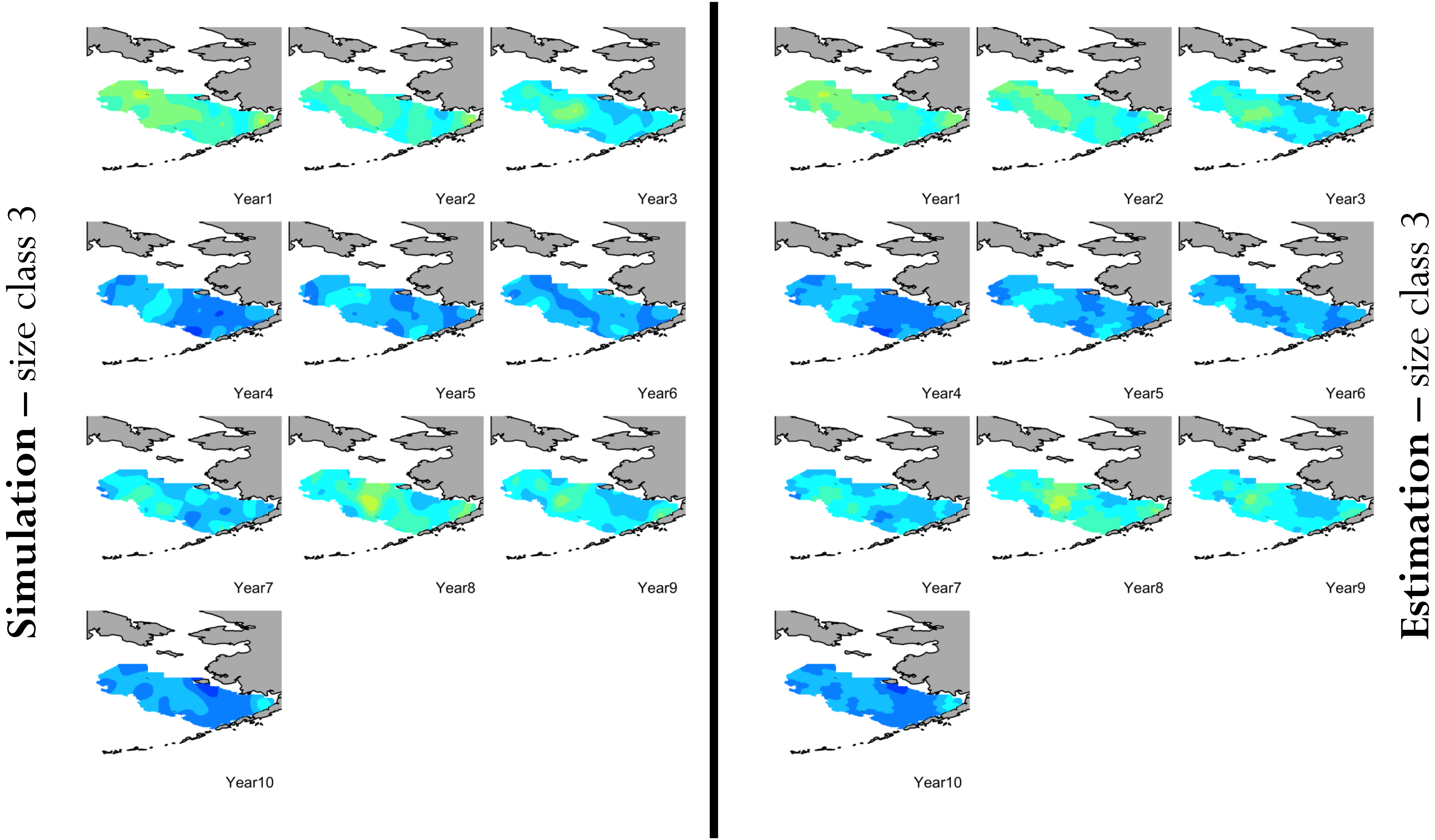
Estimation – size class 1 (recruitment)

Simulation – size class 2

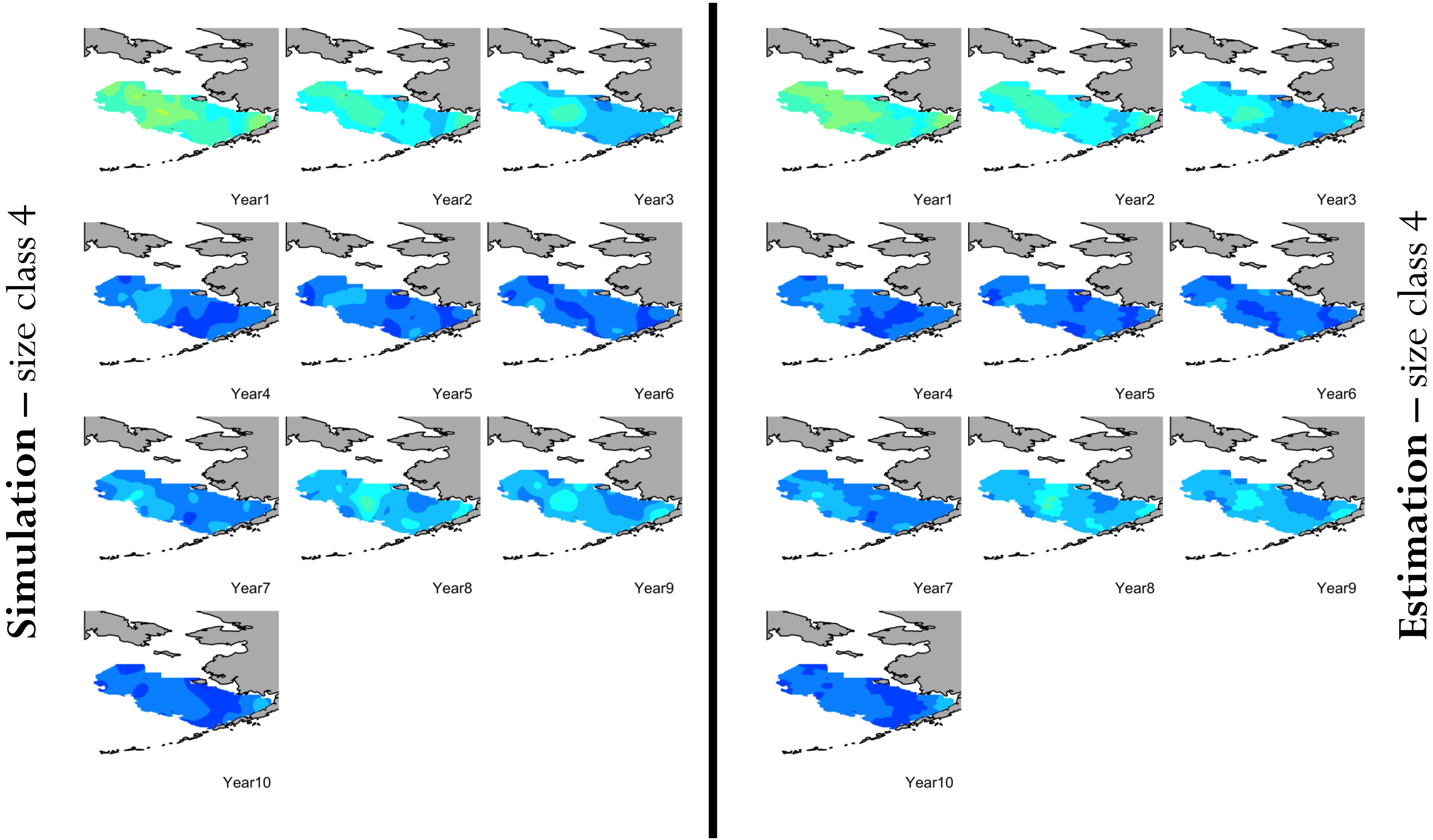


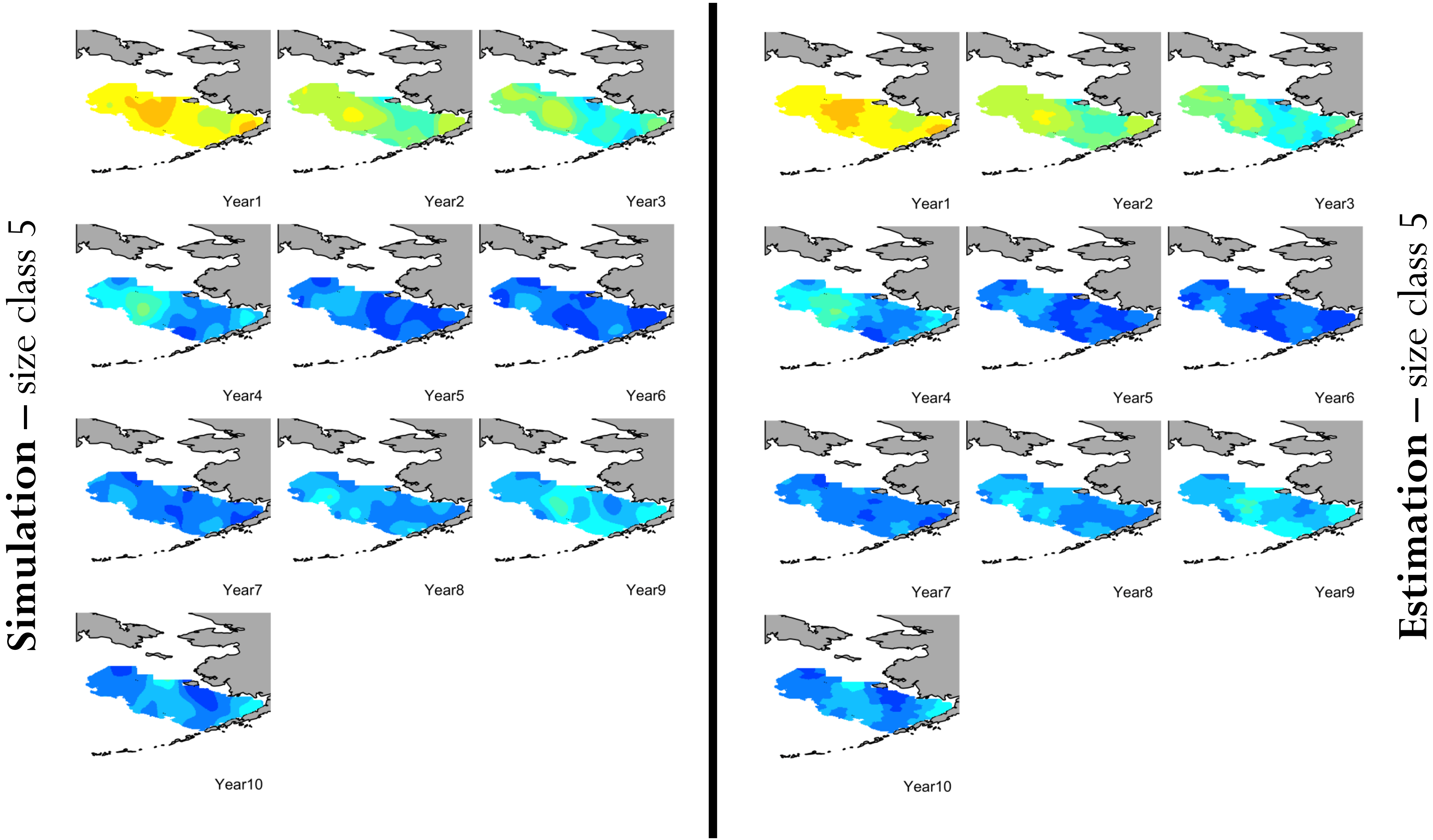
Estimation – size class 2





Simulation – size class 3
Year1
Year2
Year3
Year4
Year5
Year6
Year7
Year8
Year9
Year10
Estimation – size class 3
Year1
Year2
Year3
Year4
Year5
Year6
Year7
Year8
Year9
Year10


Simulation – size class 4
Year1
Year2
Year3
Year4
Year5
Year6
Year7
Year8
Year9
Year10
Estimation – size class 4
Year1
Year2
Year3
Year4
Year5
Year6
Year7
Year8
Year9
Year10

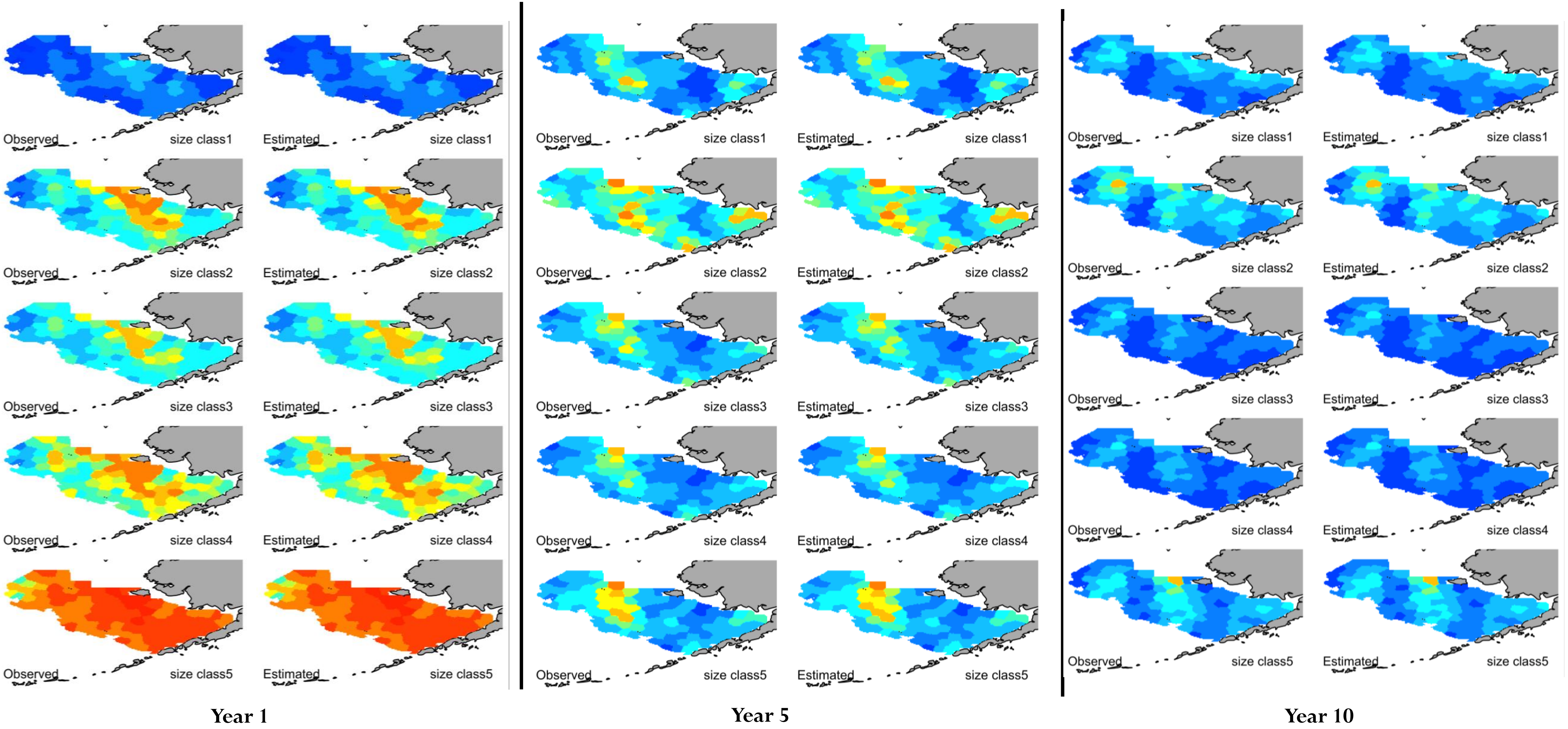

Simulation – size class 5
Year1
Year2
Year3
Year4
Year5
Year6
Year7
Year8
Year9
Year10
Estimation – size class 5
Year1
Year2
Year3
Year4
Year5
Year6
Year7
Year8
Year9
Year10

Simulation vs. Estimation

(spatially aggregated abundance)

Simulation vs. Estimation

(catch at size)

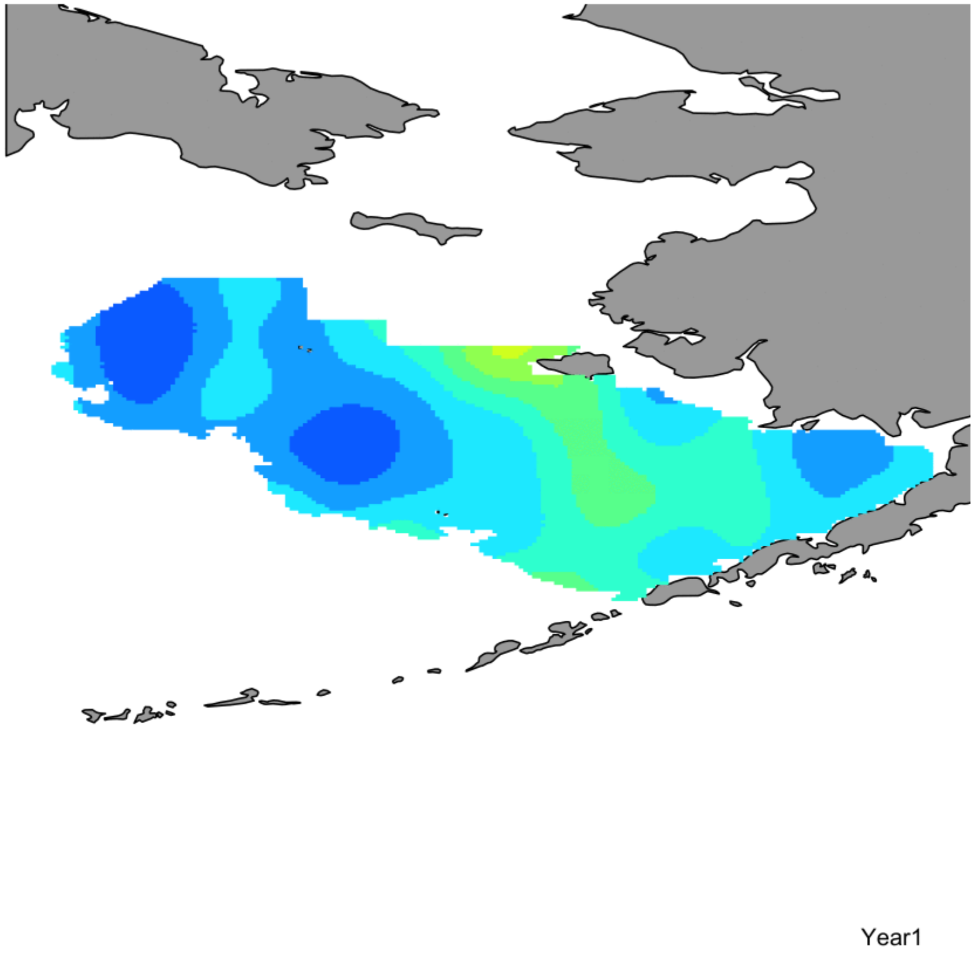


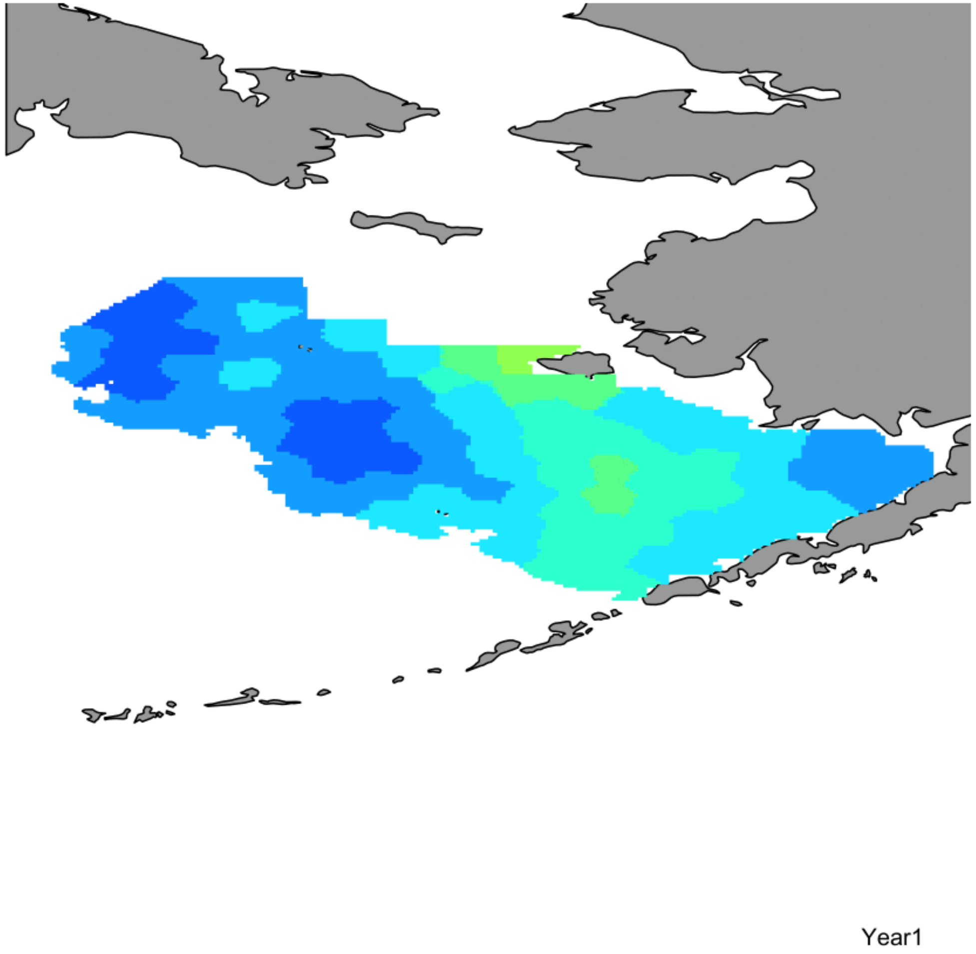
Simulation vs. Estimation

(spatially aggregated catch)

Simulation vs. Estimation

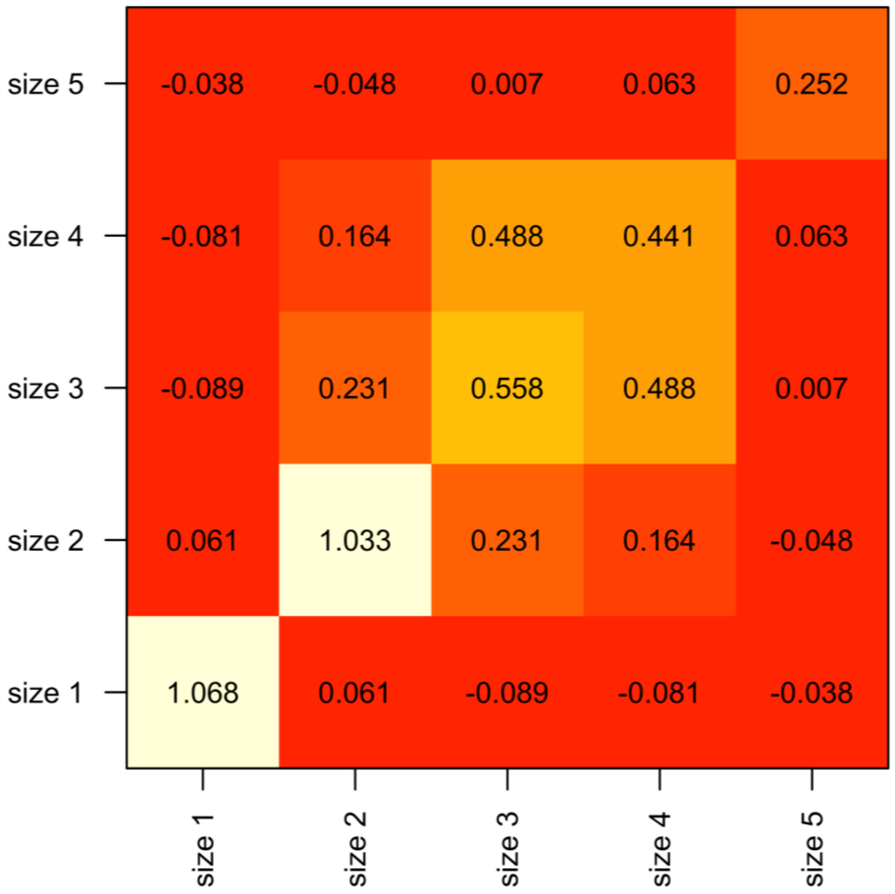
(fishing mortality)



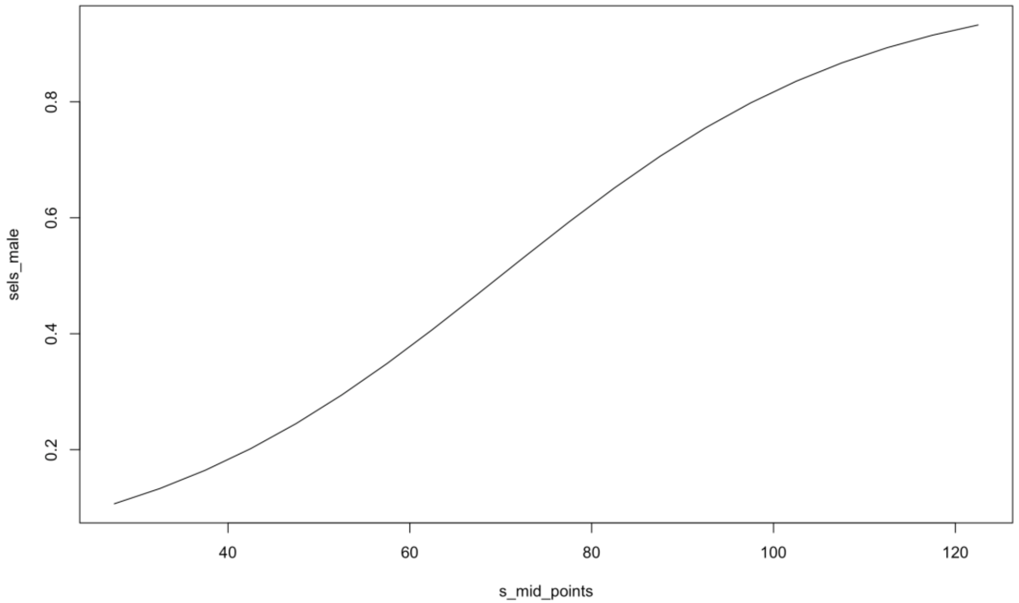


Simulation vs. Estimation
(parameters)

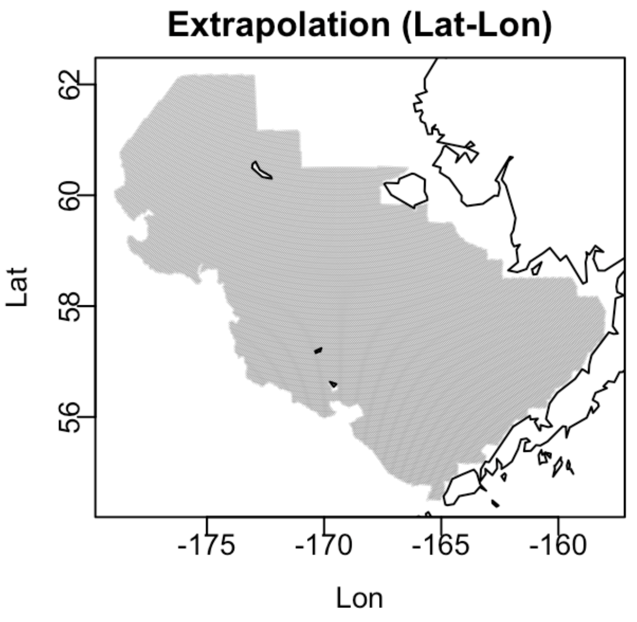
Estimated covariance of process errors (Σ_L)

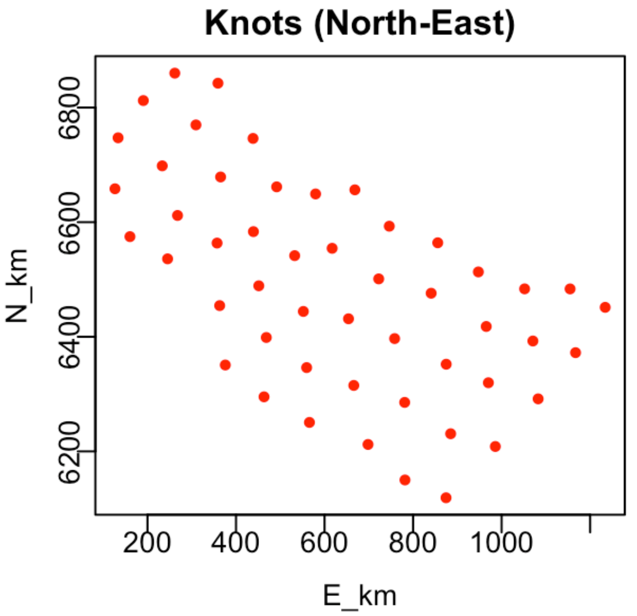


Simulation		Estimation	
select	70.000	select	69.7971
select	0.05	select	0.0499

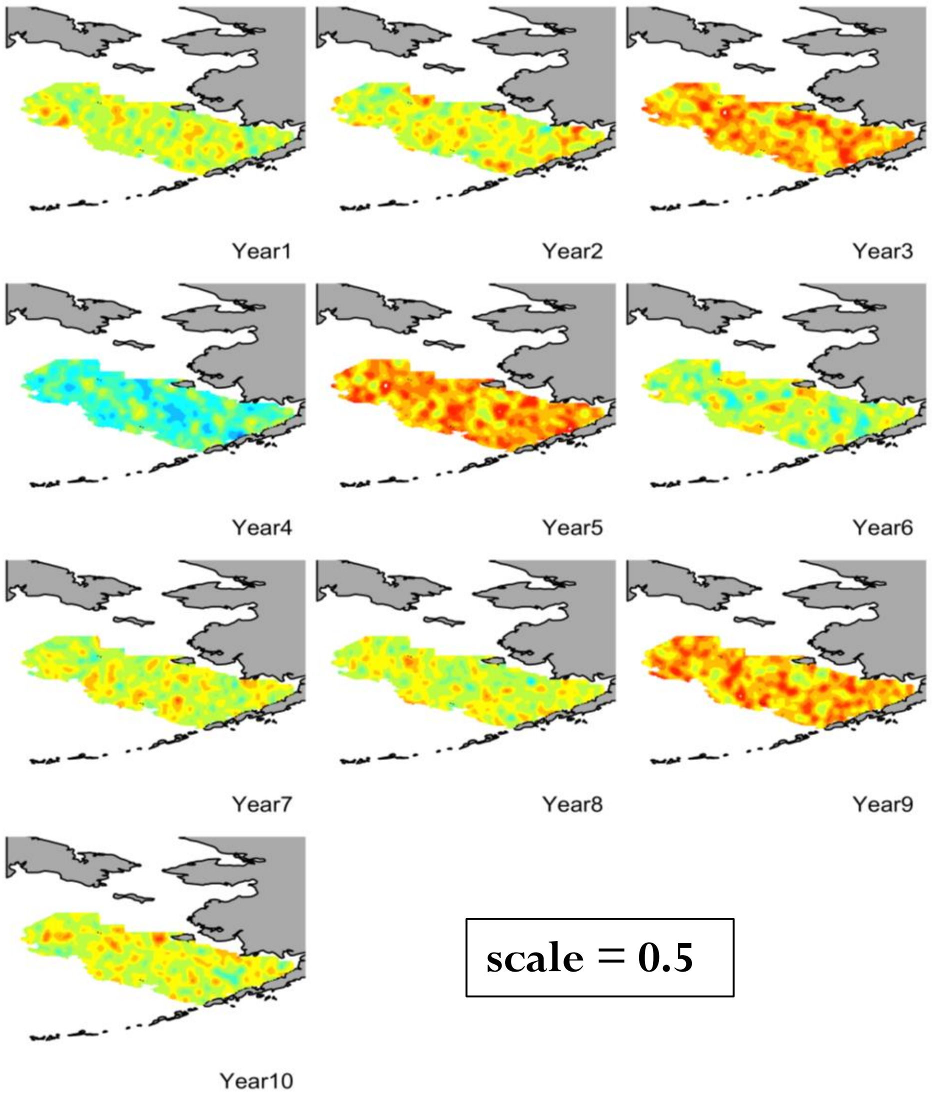


Effects of spatial scale



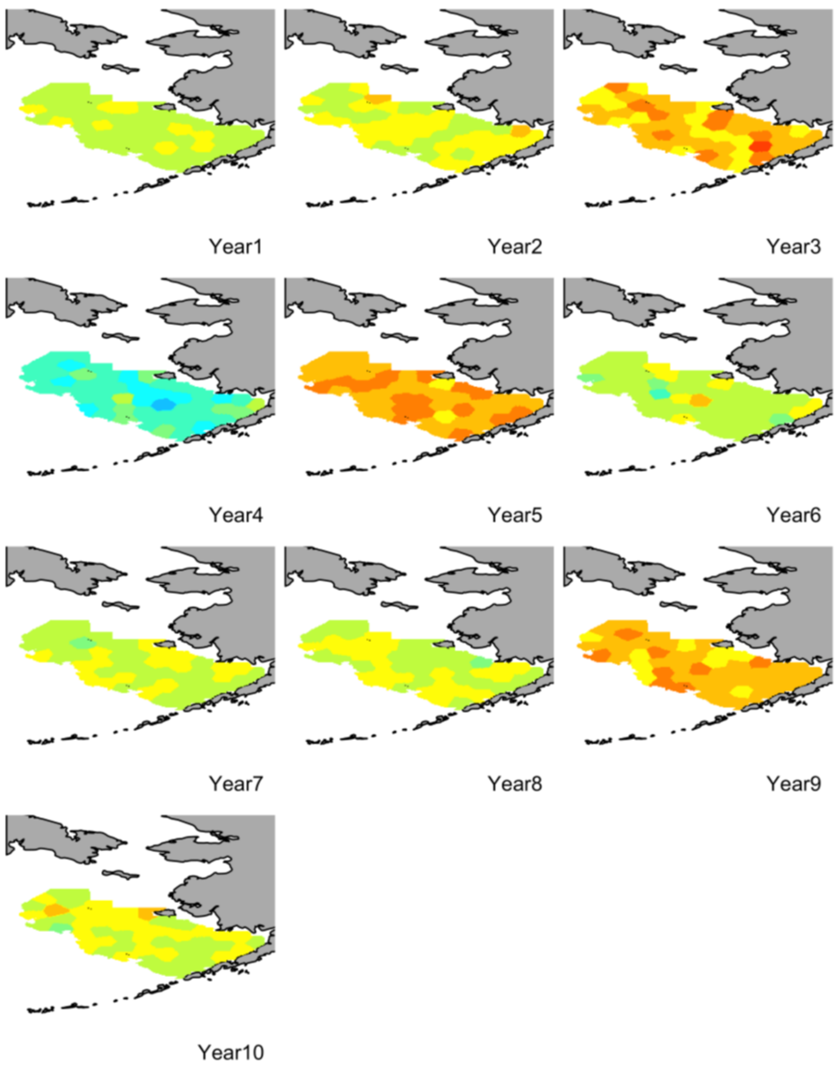


Simulated Recruitment

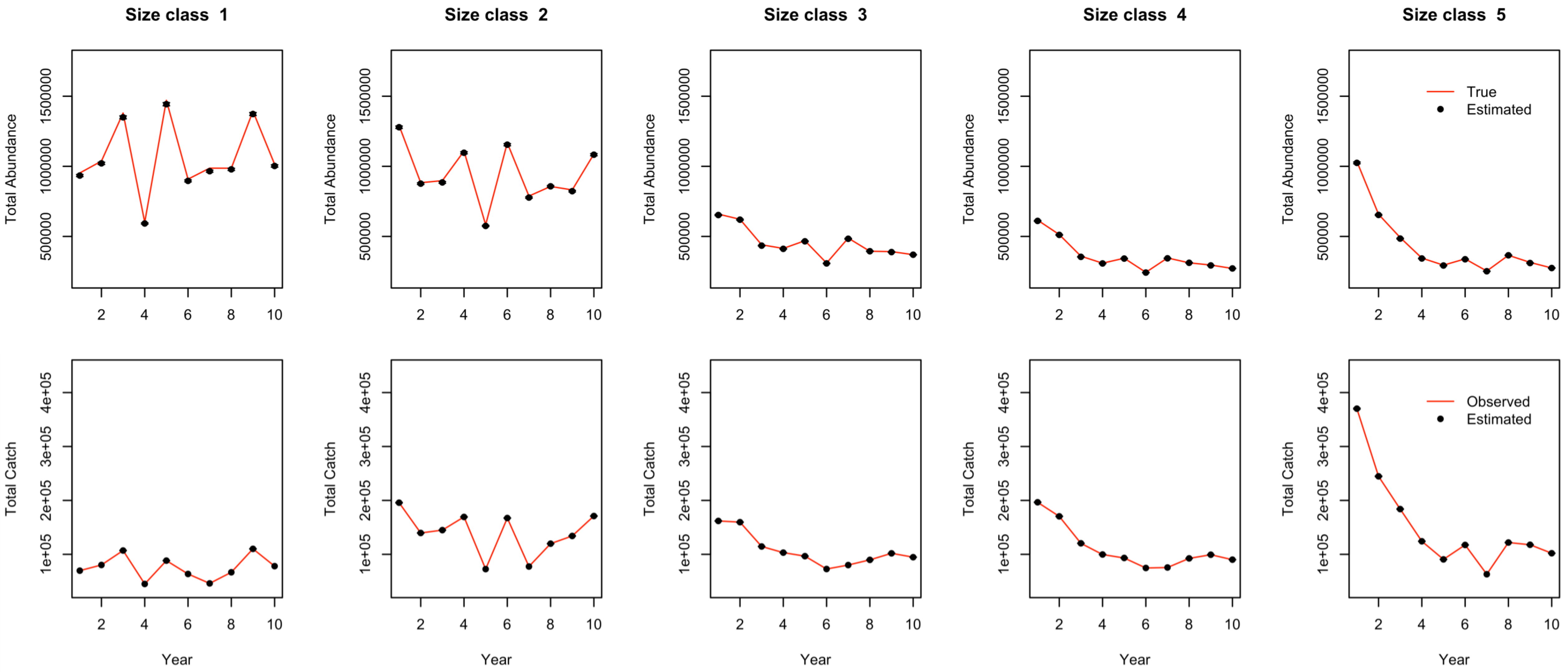


scale = 0.5

Estimated Recruitment



Effects of spatial scale



Adding sampling error

Poisson-link delta-model

Predicted encounter probability: $p_i = 1 - \exp(-n_i)$

Predicted positive: $r_i = n_i / p_i$

Likelihood function

$$\Pr(C = c_i) = \begin{cases} 1 - p_i & \text{if } c_i = 0 \\ p_i \times f(C;\, r_i, \sigma^2) & \text{if } c_i > 0 \end{cases}$$

Process to simulate data

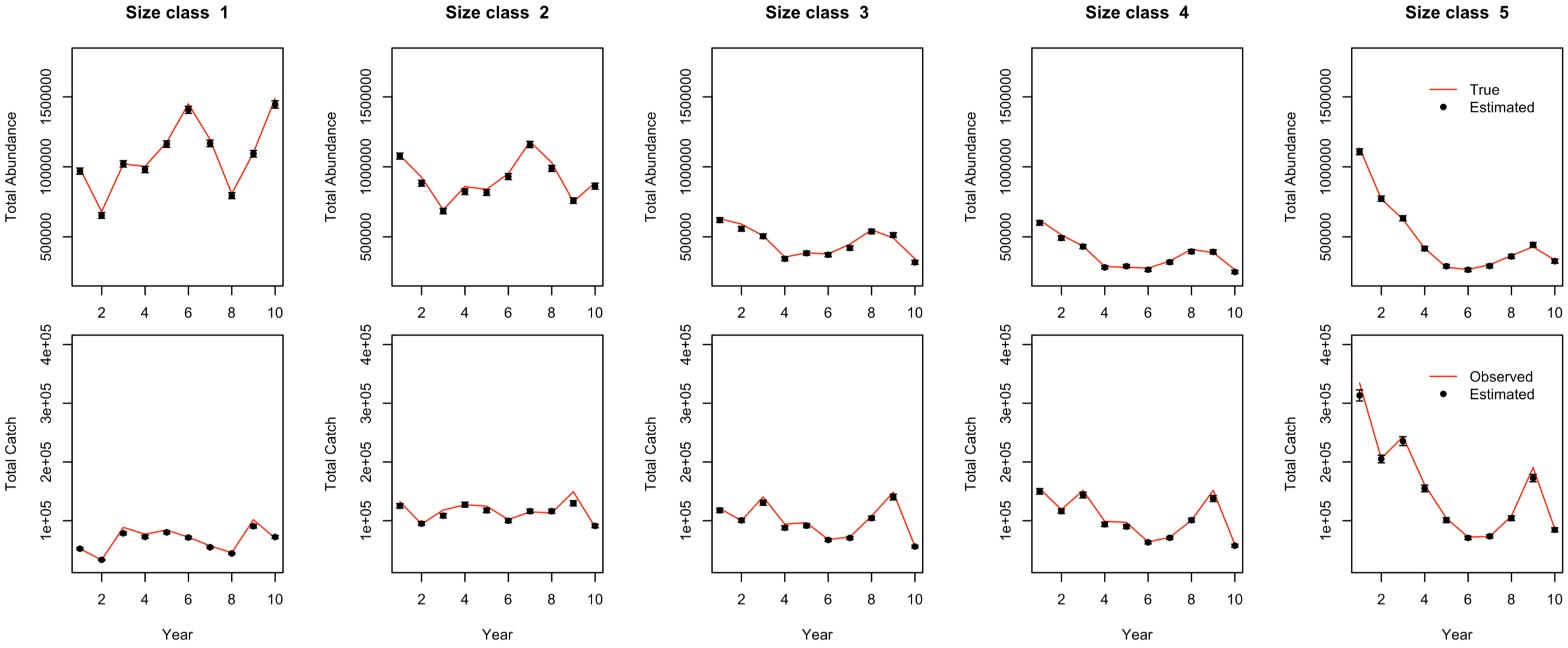
$$P \sim Bernoulli(p_i)$$

$$c_i = \begin{cases} 0 & \text{if } P = 0 \\ LN\left(\log(r_i) - \frac{\sigma^2}{2}, \sigma^2\right) & \text{if } P > 0 \end{cases}$$

Adding sampling error

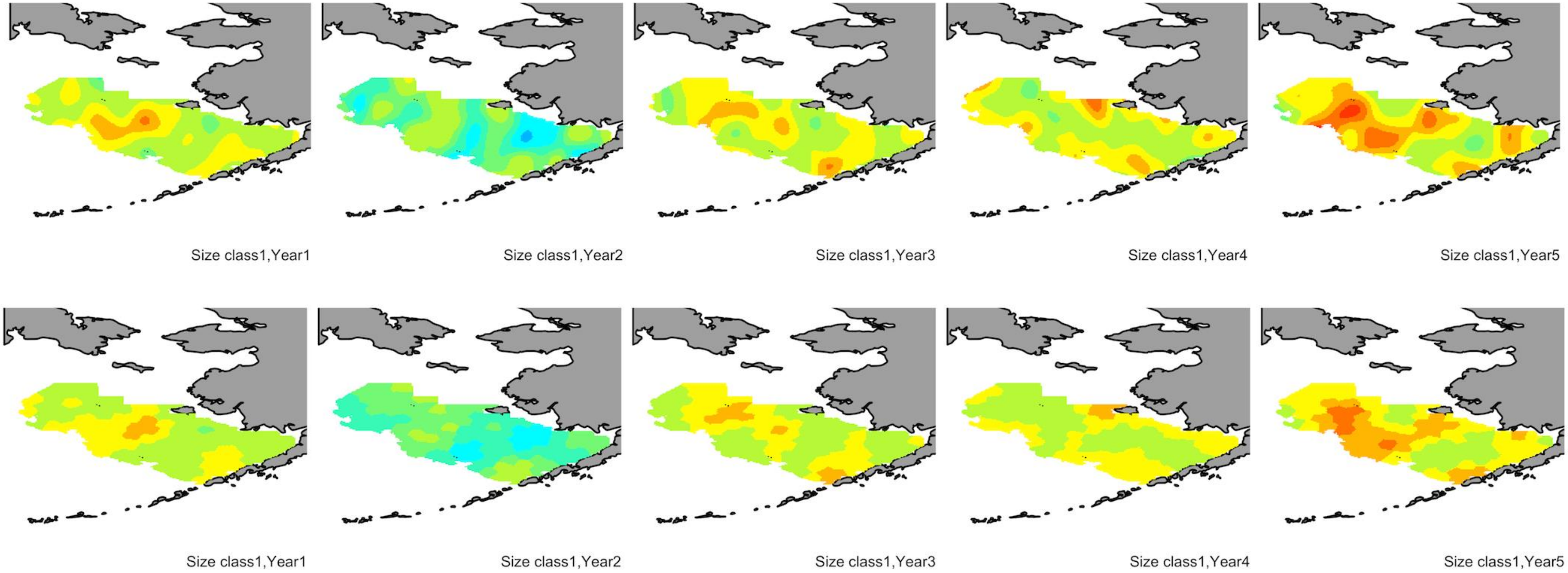
— 1 replicate

Survey: $\sigma = 0.15$; **Catch:** $\sigma = 0.20$



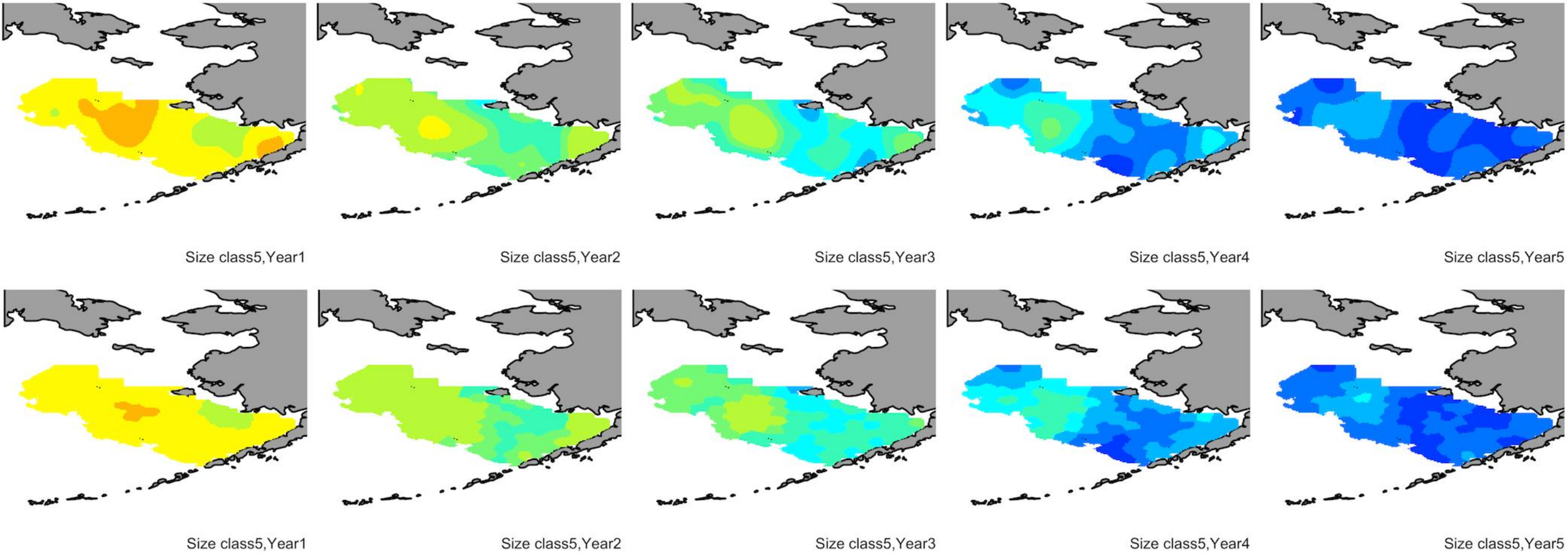
Simulation (upper panel) **vs. Estimation** (lower panel)

- density; stochastic data (**Survey:** $\sigma = 0.15$; **Catch:** $\sigma = 0.2$)



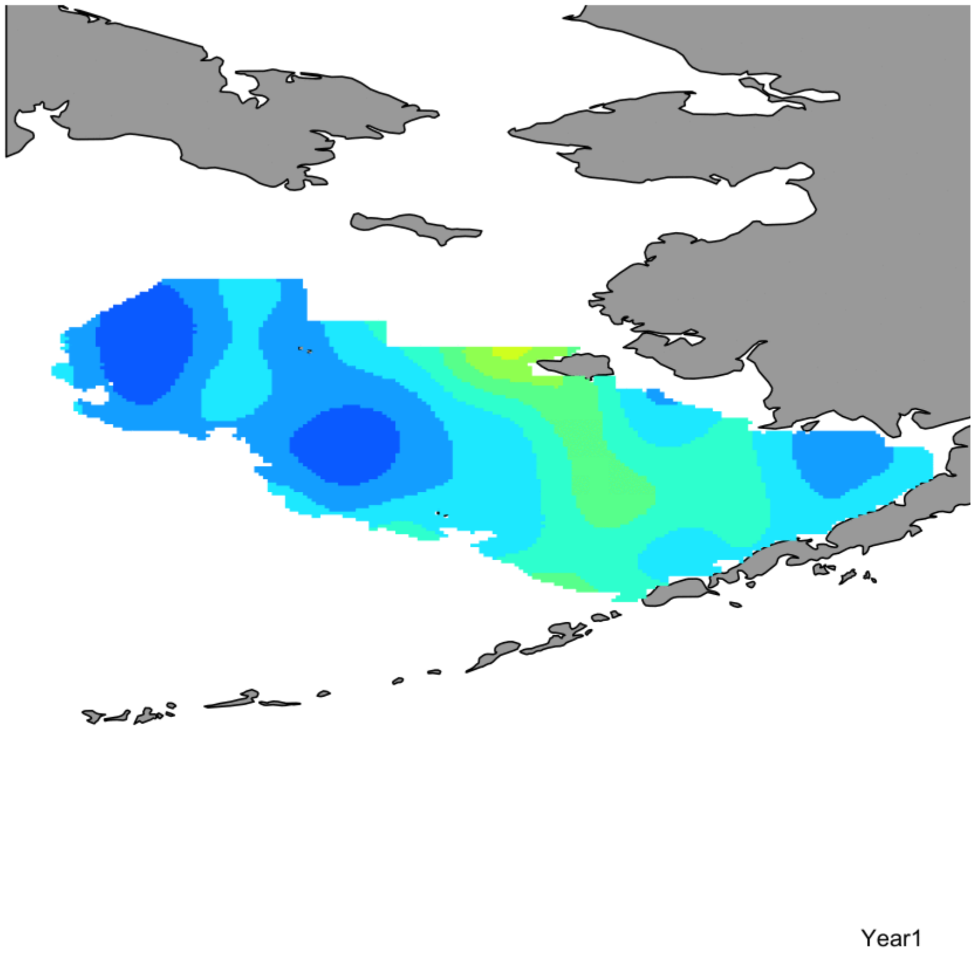
Simulation (upper panel) **vs. Estimation** (lower panel)

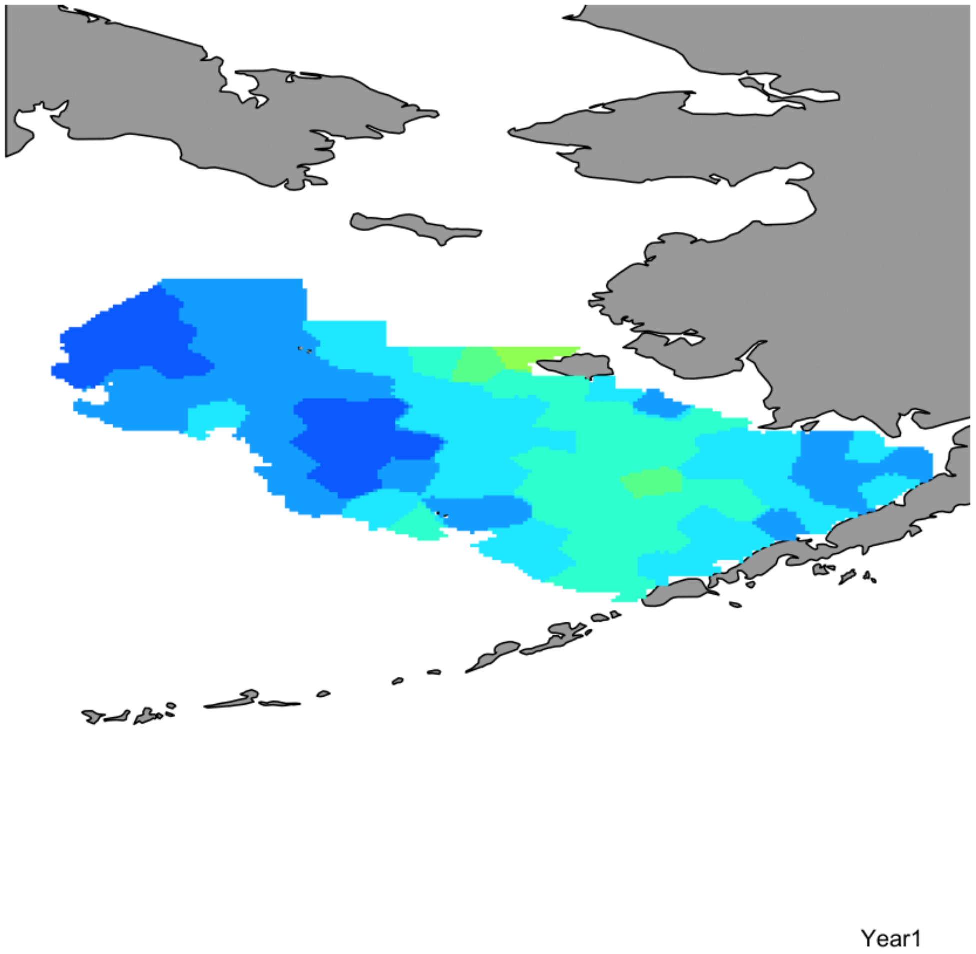
- density; stochastic data (**Survey:** $\sigma = 0.15$; **Catch:** $\sigma = 0.2$)



Simulation vs. Estimation

(fishing mortality; stochastic data)





Next steps

- Further testing
  - increasing dimension, e.g., # of size classes and years
  - model misspecifications, e.g., movement
- Comparing with non-spatial assessment models
- Application (snow crab)

Thank you !