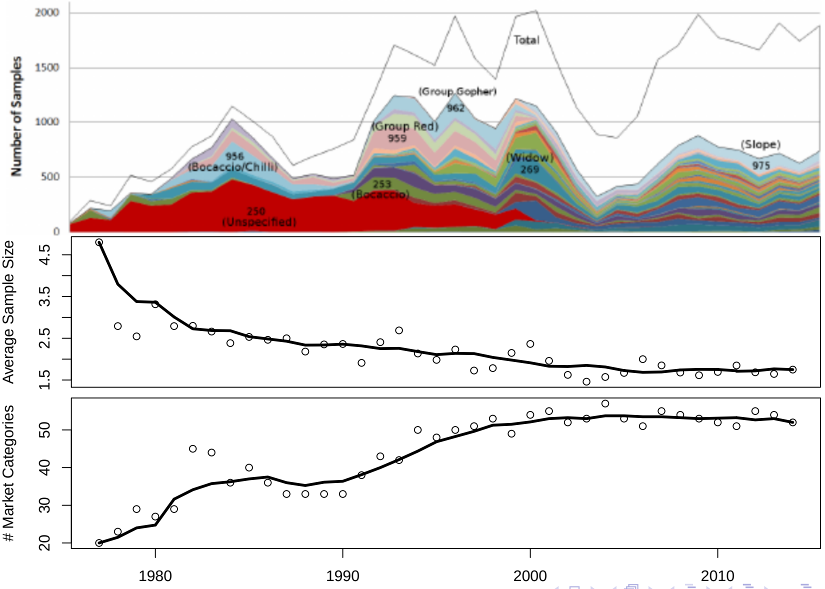
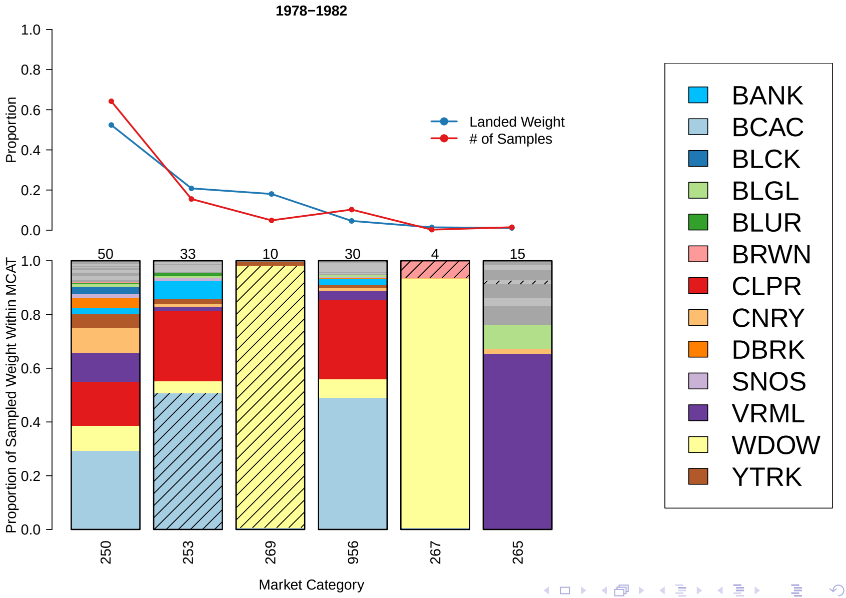
Improving Catch Estimation Methods in Sparsely Sampled, Mixed Stock Fisheries.

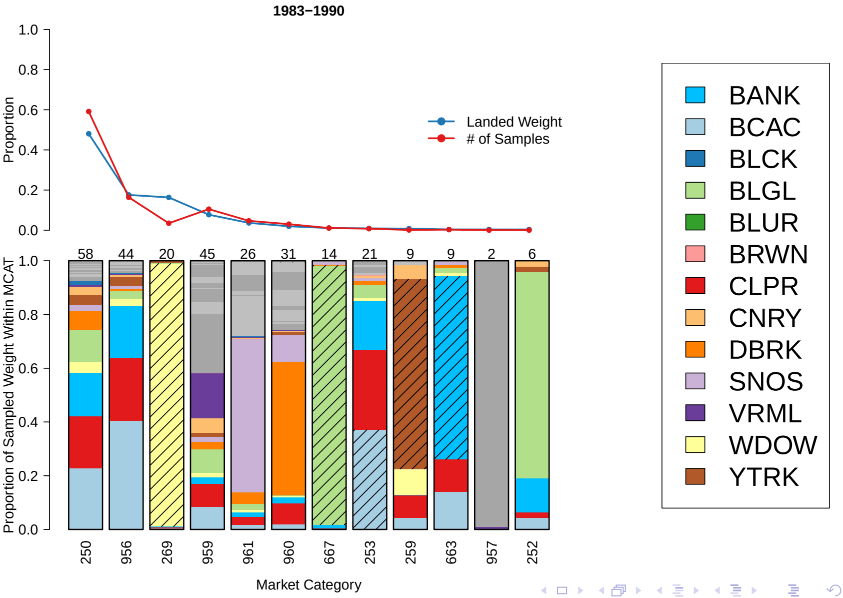
Nick Grunloh

UCSC :: CSTAR :: SWFSC :: NMFS

28 March 2018


Number of Samples
2000
1500
1000
500
0
Total
(Group Gopher)
962
(Group Red)
959
956
(Bocaccio/Chilli)
(Widow)
269
(Slope)
975
253
(Bocaccio)
250
(Unspecified)
Average Sample Size
4.5
3.5
2.5
1.5
Market Categories
50
40
30
20
1980
1990
2000
2010


1978−1982
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
Landed Weight
# of Samples
Proportion of Sampled Weight Within MCAT
50
33
10
30
4
15
250
253
269
956
267
265
Market Category
BANK
BCAC
BLCK
BLGL
BLUR
BRWN
CLPR
CNRY
DBRK
SNOS
VRML
WDOW
YTRK

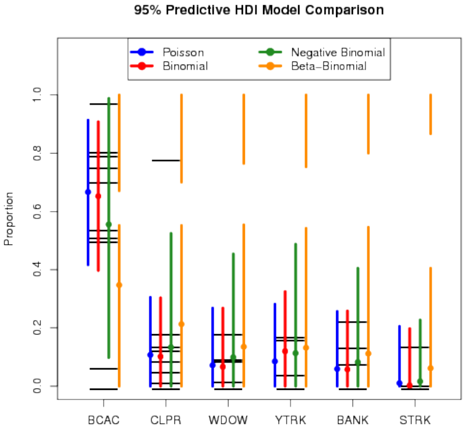


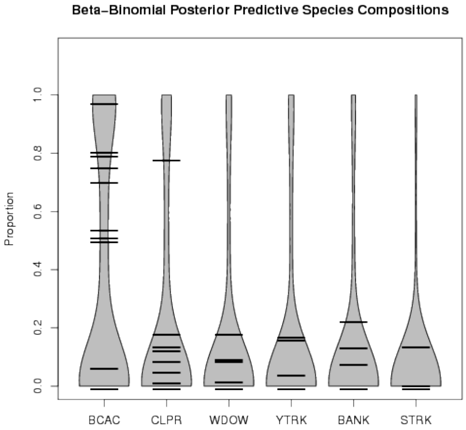
1983−1990
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
Landed Weight
# of Samples
58 44 20 45 26 31 14 21 9 9 2 6
Proportion of Sampled Weight Within MCAT
250 956 269 959 961 960 667 253 259 663 957 252
Market Category
BANK
BCAC
BLCK
BLGL
BLUR
BRWN
CLPR
CNRY
DBRK
SNOS
VRML
WDOW
YTRK

Likelihood

y_{ij}: i^{th} sample of the j^{th} species' integer weight from market category 250, in the Monterey port complex trawl fishery for the second quarter of 1982.

$$y_{ij} \sim \text{Pois}(\theta_j) \quad y_{ij} \sim \text{Bin}(\theta_j) \quad y_{ij} \sim \text{NB}(\theta_j, \phi) \quad y_{ij} \sim \text{BB}(\theta_j, \phi)$$


95% Predictive HDI Model Comparison
Poisson
Binomial
Negative Binomial
Beta−Binomial
Proportion
0.0
0.2
0.4
0.6
0.8
1.0
BCAC
CLPR
WDOW
YTRK
BANK
STRK


Beta−Binomial Posterior Predictive Species Compositions
Proportion
0.0
0.2
0.4
0.6
0.8
1.0
BCAC
CLPR
WDOW
YTRK
BANK
STRK

	Poisson	Binomial	NB	BB
MSE	0.06412	0.06264	0.05171	0.04479
Δ DIC	1001.41	1230.60	5.03	0
Δ WAIC	1079.95	1323.75	3.43	0
$pr(M\|y)$	≈ 0	≈ 0	$\approx 10^{-7}$	$\approx 1 - 10^{-7}$

Beta-Binomial Model

$$y_{ijkl m\eta} \sim \text{Beta-Binomial}\left(\mu_{jklm\eta},\ \sigma^2_{jklm\eta}\right)$$

$$\mu_{jklm\eta} = n\ \text{logit}^{-1}(\theta_{jklm\eta})$$

$$\sigma^2_{jklm\eta} = \mu_{jklm\eta}\left(1 - \frac{\mu_{jklm\eta}}{n}\right)\left(1 + (n-1)\ \rho\right)$$

$$\theta_{jklm\eta} = \beta_0 + \beta^{(s)}_j + \beta^{(p)}_k + \beta^{(g)}_l + \beta^{(t)}_{m\eta}$$

$y_{ijklm\eta}$: i^{th} sample of the j^{th} species' integer weight, in the k^{th} port, caught with the l^{th} gear, in the η^{th} quarter, of year m, for a particular market category.

$j \in \{1, ..., J\}$ Species
$k \in \{1, ..., K\}$ Ports
$l \in \{1, ..., L\}$ Gears
$m \in \{1, ..., M\}$ Years
$\eta \in \{1, ..., H\}$ Quarters

Time Model

(M1)

$$\beta_{m\eta}^{(t)} = \beta_{m}^{(y)} + \beta_{\eta}^{(q)}$$
$$\beta_{m}^{(y)} \sim N(0, 32^2)$$
$$\beta_{\eta}^{(q)} \sim N(0, 32^2)$$

(M2)

$$\beta_{m\eta}^{(t)} = \beta_{m}^{(y)} + \beta_{\eta}^{(q)}$$
$$\beta_{m}^{(y)} \sim N(0, v^{(y)})$$
$$\beta_{\eta}^{(q)} \sim N(0, v^{(q)})$$

(M3)

$$\beta_{m\eta}^{(t)} = \beta_{m}^{(y)} + \beta_{\eta}^{(q)} + \beta_{m\eta}^{(y:q)}$$
$$\beta_{m}^{(y)} \sim N(0, v^{(y)})$$
$$\beta_{\eta}^{(q)} \sim N(0, v^{(q)})$$
$$\beta_{m\eta}^{(y:q)} \sim N(0, v)$$

(M4)

$$\beta_{m\eta}^{(t)} = \beta_{m\eta}^{(y:q)}$$
$$\beta_{m\eta}^{(y:q)} \sim N(0, v)$$

(M5)

$$\beta_{m\eta}^{(t)} = \beta_{m\eta}^{(y:q)}$$
$$\beta_{m\eta}^{(y:q)} \sim N(0, v_{\eta})$$

(M6)

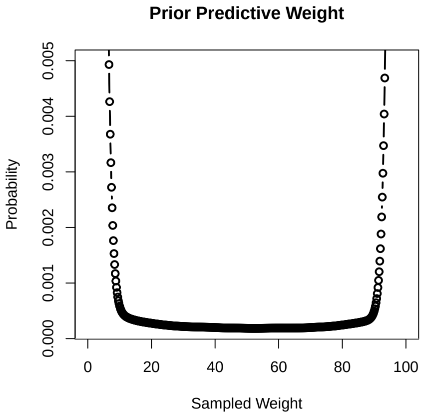
$$\beta_{m\eta}^{(t)} = \beta_{m\eta}^{(y:q)}$$
$$\beta_{m\eta}^{(y:q)} \sim N(0, v_{m})$$

Priors

$$\beta_0 \propto 1$$
$$\beta_j^{(s)} \sim N(0, 32^2)$$
$$\beta_k^{(p)} \sim N(0, 32^2)$$
$$\beta_l^{(g)} \sim N(0, 32^2)$$

$$\text{logit}(\rho) \sim N(0, 2^2)$$

$$v \sim IG(1,\ 2x10^3) \quad \forall \quad v$$



1978-1982

	M1	M2	M3	M4	M5	M6
MSE	0.12725	0.12704	0.12680	0.12237	0.12724	0.12657
Δ DIC	2558.56	2259.94	2013.21	0	2175.32	2174.71
Δ WAIC	2562.65	2263.58	2009.32	0	2171.18	2170.56
$pr(M\|y)$	≈ 0	≈ 0	≈ 0	≈ 1	≈ 0	≈ 0

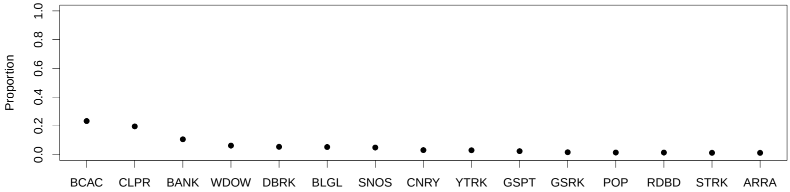
1983-1990

	M1	M2	M3	M4	M5	M6
MSE	0.12968	0.12820	0.12604	0.12604	0.12795	0.12724
Δ DIC	2865.17	2851.10	0	217.21	152.67	352.29
Δ WAIC	2847.76	2836.00	0	625.09	512.61	2170.39
$pr(M\|y)$	≈ 0	≈ 0	≈ 0.289	≈ 0.319	≈ 0.322	≈ 0.070

Posterior Predictive Weight

$$p(y^*_{jklm\eta}|\mathbf{y}) = \iint \mathrm{BB}\left(y^*_{jklm\eta}|\mu_{jklm\eta}, \sigma^2_{jklm\eta}\right) P\left(\mu_{jklm\eta}, \sigma^2_{jklm\eta}|\mathbf{y}\right) d\mu_{jklm\eta} d\sigma^2_{jklm\eta}$$

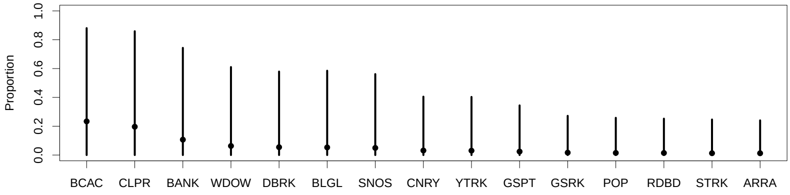
$$\pi^*_{jklm\eta} = \frac{y^*_{jklm\eta}}{\sum_j y^*_{jklm\eta}} \quad \mathbf{y}^*_{klm\eta} \neq \mathbf{0}$$



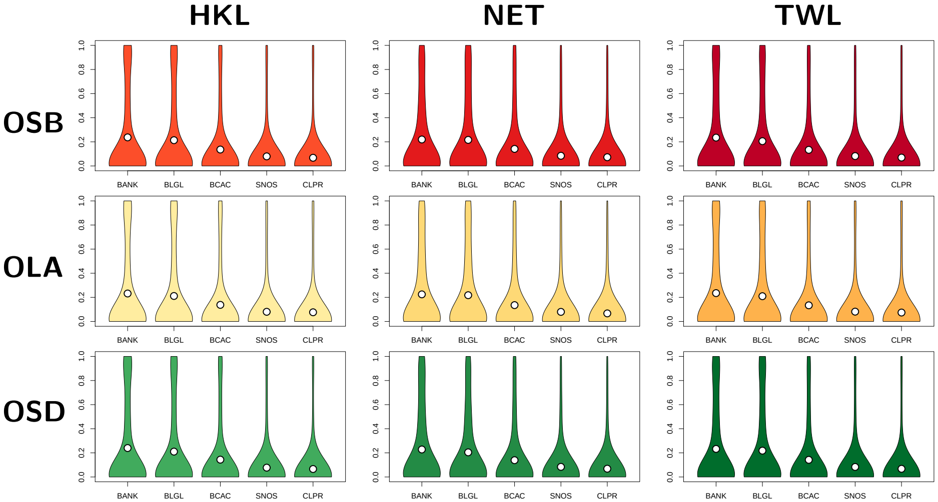
Posterior Predictive Weight

$$p(y^*_{jklm\eta}|\mathbf{y}) = \iint \mathrm{BB}\left(y^*_{jklm\eta}|\mu_{jklm\eta}, \sigma^2_{jklm\eta}\right) P\left(\mu_{jklm\eta}, \sigma^2_{jklm\eta}|\mathbf{y}\right) d\mu_{jklm\eta} d\sigma^2_{jklm\eta}$$

$$\pi^*_{jklm\eta} = \frac{y^*_{jklm\eta}}{\sum_j y^*_{jklm\eta}} \quad \mathbf{y}^*_{klm\eta} \neq \mathbf{0}$$



Single Quarter Hindcast (MCAT 250)



Predictive Accuracy: 1978-1982

Market Category	Tons Landed	68%	95%	99%
250: Unspecified	36539.3	67.1%	96.1%	98.7%
253: Bocaccio	14512.9	67.3%	96.3%	98.9%
269: Widow	12575.6	68.2%	88.8%	90.2%
262: Thornyheads	8512.2	67.4%	93.8%	95.3%
956: G. Boc./Chili.	3213.7	68.3%	96.7%	99.2%
265: Yelloweye	774.8	69.6%	96.0%	97.8%
270: Splitnose	458.7	68.6%	93.6%	96.7%
959: G. Red	225.1	68.5%	96.3%	98.1%
961: G. Rosefish	162.1	69.3%	93.2%	95.3%
Average*		68.3%	94.5%	96.7%

Predictive Accuracy: 1983-1990

Market Category	Tons Landed	68%	95%	99%
250: Unspecified	55332	68.1%	96.0%	99.0%
262: Thornyheads	27929	68.5%	95.1%	95.9%
956: G. Boc./Chili.	20227	67.5%	96.2%	99.0%
269: Widow	18802	68.6%	94.2%	94.7%
959: G. Red	8883	67.4%	96.4%	99.0%
961: G. Rosefish	4179	68.6%	94.6%	97.8%
960: G. Small	2223	68.0%	96.1%	98.6%
667: Blackgill	1213	69.4%	92.5%	93.5%
253: Bocaccio	1029	69.3%	97.1%	98.9%
259: Yellowtail	868	83.8%	91.9%	92.9%
663: Bank	432	68.1%	94.1%	96.3%
245: Cowcod	273	60.8%	94.9%	97.7%
270: Splitnose	3	67.9%	94.2%	96.7%
Average*		68.9%	94.9%	96.9%

Speciated Landings

If $\lambda_{\cdot klm\eta}$ is the observed landings of **all species** in the k^{th} port, caught with the l^{th} gear, in the η^{th} quarter, of year m, in particular market category. Then,

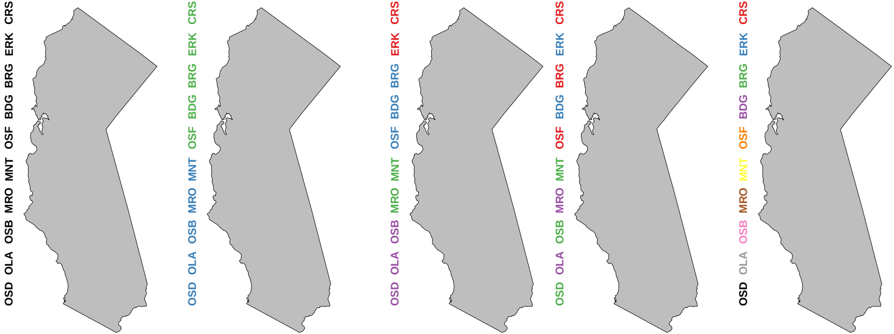
$$\lambda^*_{jklm\eta} = \lambda_{\cdot klm\eta}\pi^*_{jklm\eta}$$

$$\lambda^*_{jklm\cdot} = \sum_{\eta} \lambda^*_{jklm\eta}$$

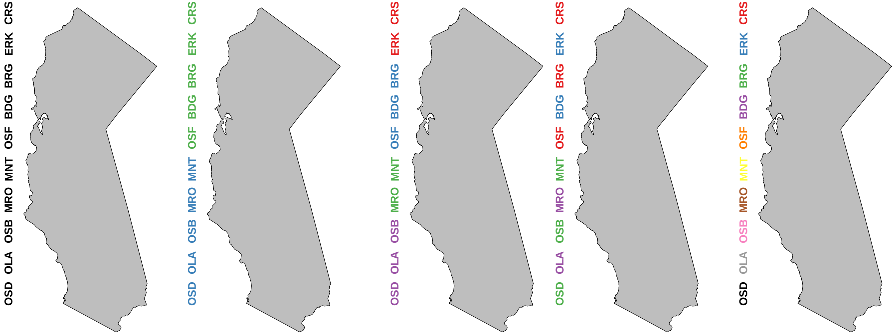
$$\lambda^*_{j\cdot lm\cdot} = \sum_{k}\sum_{\eta} \lambda^*_{jklm\eta}$$

$$\lambda^*_{j\cdot\cdot m\cdot} = \sum_{l}\sum_{k}\sum_{\eta} \lambda^*_{jklm\eta}$$

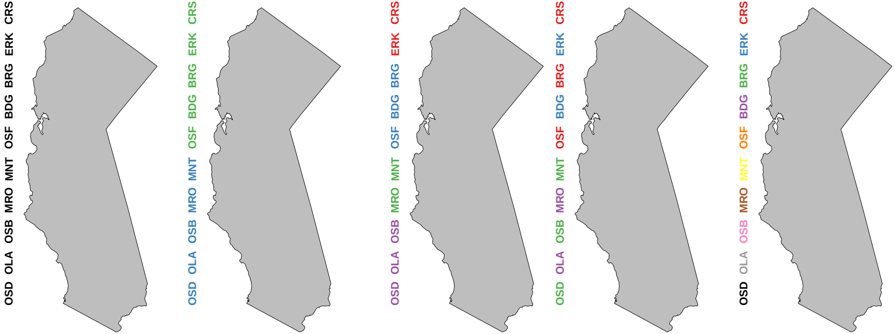
$$\mathrm{MSE}(\hat\theta) = \mathbb{E}\left[\, (\hat\theta-\theta)^2 \,\right] = \overbrace{\mathbb{E}\left[\, \left(\hat\theta - \mathbb{E}(\hat\theta)\right)^2 \,\right]}^{\mathrm{Var}(\hat\theta)} + \overbrace{\left(\, \mathbb{E}(\hat\theta) - \theta \,\right)^2}^{\mathrm{Bias}(\hat\theta,\theta)^2}$$



$$B_K = \sum_{\kappa=0}^{K} \frac{1}{\kappa!} \left(\sum_{j=0}^{\kappa} (-1)^{\kappa-j} \binom{\kappa}{j} j^K \right)$$



$$B_{10} = 115975$$


OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS

$$\bar{B}_{10} = 61136$$

$$\hat{B}_{10} = 512$$


OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS
OSD OLA OSB MRO MNT OSF BDG BRG ERK CRS

$$\hat{\hat{B}}_{10} = 274$$

Bayesian Model Averaging (BMA)

Consider a set of Models (M) indexed by ι:

$$\omega_\iota = Pr(M_\iota|y) = \frac{p(y|M_\iota)p(M_\iota)}{\sum_\iota p(y|M_\iota)p(M_\iota)}$$

$$\bar{p}(\boldsymbol{\theta}|\mathbf{y}) = \sum_\iota \omega_\iota p(\boldsymbol{\theta}|\mathbf{y}, M_\iota)$$

if f only depends on M through $\boldsymbol{\theta}$, then

$$\bar{p}(y^*|\mathbf{y}) = \int f(y^*|\boldsymbol{\theta})\bar{p}(\boldsymbol{\theta}|\mathbf{y})\mathbf{d}\boldsymbol{\theta}$$

* Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical science*, 382-401.

1978-1982

MCAT 250										
ω	0.32	0.14	0.13	0.12	0.02	0.02	0.02	0.02	0.02	0.02
CRS										
ERK										
BRG										
BDG										
OSF										
MNT										
MRO										
~~OSB~~										
~~OLA~~										
~~OSD~~										

1978-1982

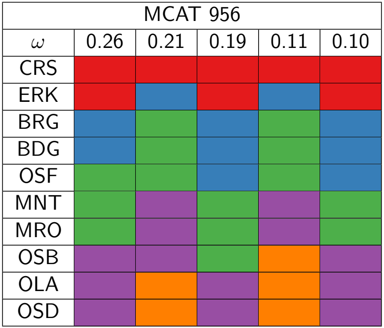
MCAT 253					
ω	0.14	0.14	0.14	0.10	0.06
CRS					
ERK					
BRG					
BDG					
OSF					
MNT					
MRO					
~~OSB~~					
~~OLA~~					
~~OSD~~					

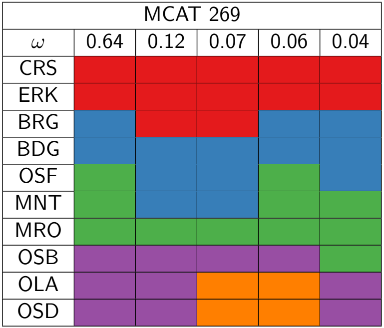
MCAT 269					
ω	0.19	0.14	0.14	0.13	0.07
CRS					
ERK					
BRG					
BDG					
OSF					
MNT					
MRO					
~~OSB~~					
~~OLA~~					
~~OSD~~					

1983-1990

MCAT 250							
ω	0.73	0.25	0.00	0.00	0.00	0.00	0.00
CRS							
ERK							
BRG							
BDG							
OSF							
MNT							
MRO							
OSB							
OLA							
OSD							

1983-1990


MCAT 956
ω
0.26
0.21
0.19
0.11
0.10
CRS
ERK
BRG
BDG
OSF
MNT
MRO
OSB
OLA
OSD


MCAT 269
ω
0.64
0.12
0.07
0.06
0.04
CRS
ERK
BRG
BDG
OSF
MNT
MRO
OSB
OLA
OSD

Conclusion

Bayesian Model Based Statistics:

- Model Overdispersion
- Uncertainty Estimation
- Mechanisms for Pooling
- Out-of-Sample Predictions

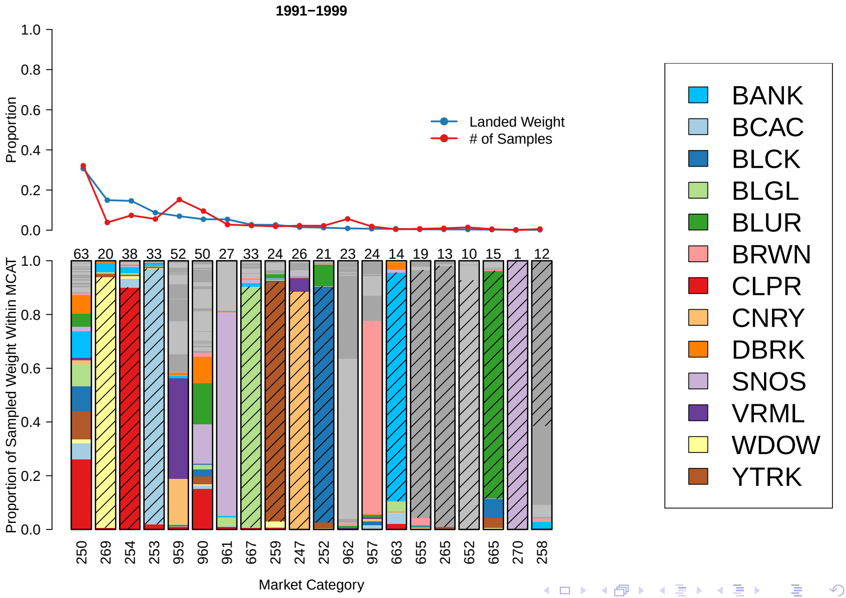
Conclusion

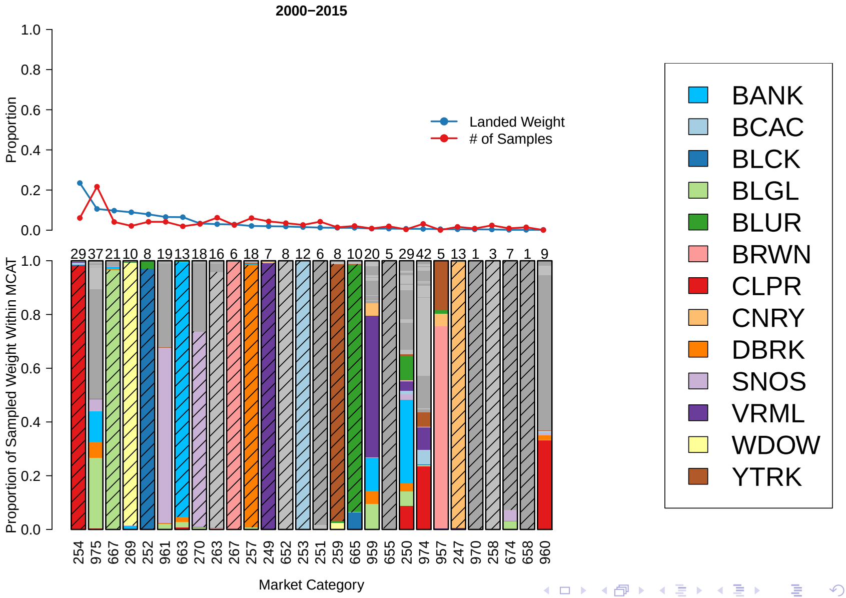
Bayesian Model Based Statistics:

- Model Overdispersion
- Uncertainty Estimation
- Mechanisms for Pooling
- Out-of-Sample Predictions

Future Directions:

- Additional Predictors
- Multivariate Models
- Time Series Hindcasting
- Dirichlet Process Models


1991–1999
Proportion
Landed Weight
# of Samples
63 20 38 33 52 50 27 33 24 26 21 23 24 14 19 13 10 15 1 12
Proportion of Sampled Weight Within MCAT
250 269 254 253 959 960 961 667 259 247 252 962 957 663 655 265 652 665 270 258
Market Category
BANK
BCAC
BLCK
BLGL
BLUR
BRWN
CLPR
CNRY
DBRK
SNOS
VRML
WDOW
YTRK


2000–2015
Proportion
Landed Weight
# of Samples
Proportion of Sampled Weight Within MCAT
29 37 21 10 8 19 13 18 16 6 18 7 8 12 6 8 10 20 5 29 42 5 13 1 3 7 1 9
254 975 667 269 252 961 663 270 263 267 257 249 652 253 251 259 665 959 655 250 974 957 247 970 258 674 658 960
Market Category
BANK
BCAC
BLCK
BLGL
BLUR
BRWN
CLPR
CNRY
DBRK
SNOS
VRML
WDOW
YTRK

ρ Posterior

MCAT	Mean	Median	SD
250	0.55	0.55	0.004
253	0.39	0.39	0.001
262	0.35	0.35	0.008
265	0.64	0.64	0.002
269	0.52	0.52	0.019
270	0.53	0.54	0.020
956	0.35	0.35	0.007
959	0.47	0.47	0.070
961	0.55	0.55	0.004

1978-1982

MCAT	Mean	Median	SD
245	0.65	0.65	0.014
250	0.51	0.51	0.002
253	0.47	0.47	0.010
259	0.75	0.75	0.009
262	0.41	0.41	0.001
269	0.57	0.57	0.046
270	0.74	0.75	0.027
663	0.51	0.51	0.001
667	0.49	0.49	0.022
956	0.43	0.43	0.003
959	0.55	0.55	0.004
960	0.45	0.45	0.004
961	0.59	0.59	0.001

1983-1990

v Posterior

MCAT	Mean	Median	SD
250	12915.85	18523.12	8699.87
253	22747.87	23063.76	1535.53
262	20254.41	20506.36	2581.87
265	15846.22	16694.98	7601.15
269	20135.05	19975.15	4667.11
270	19931.96	19955.13	6033.35
956	19659.11	19795.60	1227.99
959	19159.69	13375.80	19256.94
961	18631.44	19498.31	7970.44

1978-1982

MCAT	Mean	Median	SD
245	20211.82	20204.95	1276.83
250	236.03	192.53	134.67
253	20455.18	20140.50	1521.72
259	20246.14	20186.61	898.99
262	20445.49	20348.56	343.70
269	34386.49	25951.03	24030.32
270	20253.34	19908.07	9269.02
663	19563.87	19624.09	331.04
667	20089.55	20078.27	2723.34
956	20581.67	20664.71	913.92
959	19242.41	18707.09	5076.03
960	20059.66	20012.80	1703.89
961	20127.69	20141.04	580.80

1983-1990

Proof: Species Comps Sum to One... as do Their Means.

If y_{jk} is the k^{th} draw, $k \in \{1, ..., K\}$, of the posterior predictive weight of species j in a particular stratum. Then,

$$\pi_{jk} = \frac{y_{jk}}{\sum_j y_{jk}} \quad \mathbf{y}_k \neq \mathbf{0}. \tag{1}$$

The predictive mean for species j is,

$$\hat{\pi}_j = \frac{\sum_k^K \pi_{jk}}{K}. \tag{2}$$

Summing $\hat{\pi}_j$ across species, it follows from (1) and (2) that,

$$\sum_j \hat{\pi}_j \stackrel{(2)}{=} \sum_j \frac{\sum_k^K \pi_{jk}}{K} = \frac{\sum_k^K \sum_j \pi_{jk}}{K} \stackrel{(1)}{=} \frac{\sum_k^K \sum_j \frac{y_{jk}}{\sum_j y_{jk}}}{K} = \frac{\sum_k^K 1}{K} = \frac{K}{K} = 1. \quad \blacksquare$$

Landings Weighted

$$y_{ijklm\eta} \sim \text{Beta-Binomial}\Big(\mu(\theta_{jklm\eta}),\ \sigma^2(\theta_{jklm\eta}, \rho)\Big)$$

$$\theta_{jklm\eta} = \beta_0 + \beta_j^{(s)} + \beta_k^{(p)} + \beta_l^{(g)} + \beta_{m\eta}^{(y:q)} + \boldsymbol{\beta^{(\ell)}\ell}$$

$$\beta_0 \propto 1$$
$$\beta_j^{(s)} \sim N(0, 32^2)$$
$$\beta_k^{(p)} \sim N(0, 32^2)$$
$$\beta_l^{(g)} \sim N(0, 32^2)$$
$$\boldsymbol{\beta^{(\ell)} \sim N(0, 32^2)}$$

$$\text{logit}(\rho) \sim N(0, 2^2)$$

$$\beta_{m\eta}^{(y:q)} \sim N(0, v)$$
$$v \sim IG(1, 2x10^3)$$
$$\forall \quad v$$

$$j \in \{1, ..., J\} \text{ Species}$$
$$k \in \{1, ..., K\} \text{ Ports}$$
$$l \in \{1, ..., L\} \text{ Gears}$$
$$m \in \{1, ..., M\} \text{ Years}$$
$$\eta \in \{1, ..., H\} \text{ Quarters}$$

Vessel Effects

$$y_{ijklm\eta\nu} \sim \text{Beta-Binomial}\Big(\mu(\theta_{jklm\eta\nu}),\ \sigma^2(\theta_{jklm\eta\nu}, \rho)\Big)$$

$$\theta_{jklm\eta} = \beta_0 + \beta_j^{(s)} + \beta_k^{(p)} + \beta_l^{(g)} + \beta_{m\eta}^{(y:q)} + \boldsymbol{\beta_\nu^{(v)}}$$

$$\beta_0 \propto 1$$
$$\beta_j^{(s)} \sim N(0, 32^2)$$
$$\beta_k^{(p)} \sim N(0, 32^2)$$
$$\beta_l^{(g)} \sim N(0, 32^2)$$

$$\text{logit}(\rho) \sim N(0, 2^2)$$
$$\beta_{m\eta}^{(y:q)} \sim N(0, v)$$
$$\boldsymbol{\beta_\nu^{(v)} \sim N(0, v)}$$
$$v \sim IG(1, 2x10^3)$$
$$\forall \quad v$$

$$j \in \{1, ..., J\} \text{ Species}$$
$$k \in \{1, ..., K\} \text{ Ports}$$
$$l \in \{1, ..., L\} \text{ Gears}$$
$$m \in \{1, ..., M\} \text{ Years}$$
$$\eta \in \{1, ..., H\} \text{ Quarters}$$
$$\boldsymbol{\nu \in (1, ..., N) \text{ Vessels}}$$

Species:Gear Interactions

$$y_{ijklm\eta} \sim \text{Beta-Binomial}\Big(\mu(\theta_{jklm\eta}),\ \sigma^2(\theta_{jklm\eta}, \rho)\Big)$$

$$\theta_{jklm\eta} = \beta_0 + \beta_j^{(s)} + \beta_k^{(p)} + \beta_l^{(g)} + \beta_{m\eta}^{(y:q)} + \boldsymbol{\beta_{jl}^{(s:g)}}$$

$$\text{logit}(\rho) \sim N(0, 2^2)$$

$$\beta_0 \propto 1$$

$$\beta_j^{(s)} \sim N(0, 32^2)$$

$$\beta_k^{(p)} \sim N(0, 32^2)$$

$$\beta_l^{(g)} \sim N(0, 32^2)$$

$$\beta_{m\eta}^{(y:q)} \sim N(0, v)$$

$$\boldsymbol{\beta_{jl}^{(s:g)} \sim N(0, v)}$$

$$v \sim IG(1, 2x10^3)$$

$$\forall \quad v$$

$j \in \{1, ..., J\}$ Species

$k \in \{1, ..., K\}$ Ports

$l \in \{1, ..., L\}$ Gears

$m \in \{1, ..., M\}$ Years

$\eta \in \{1, ..., H\}$ Quarters