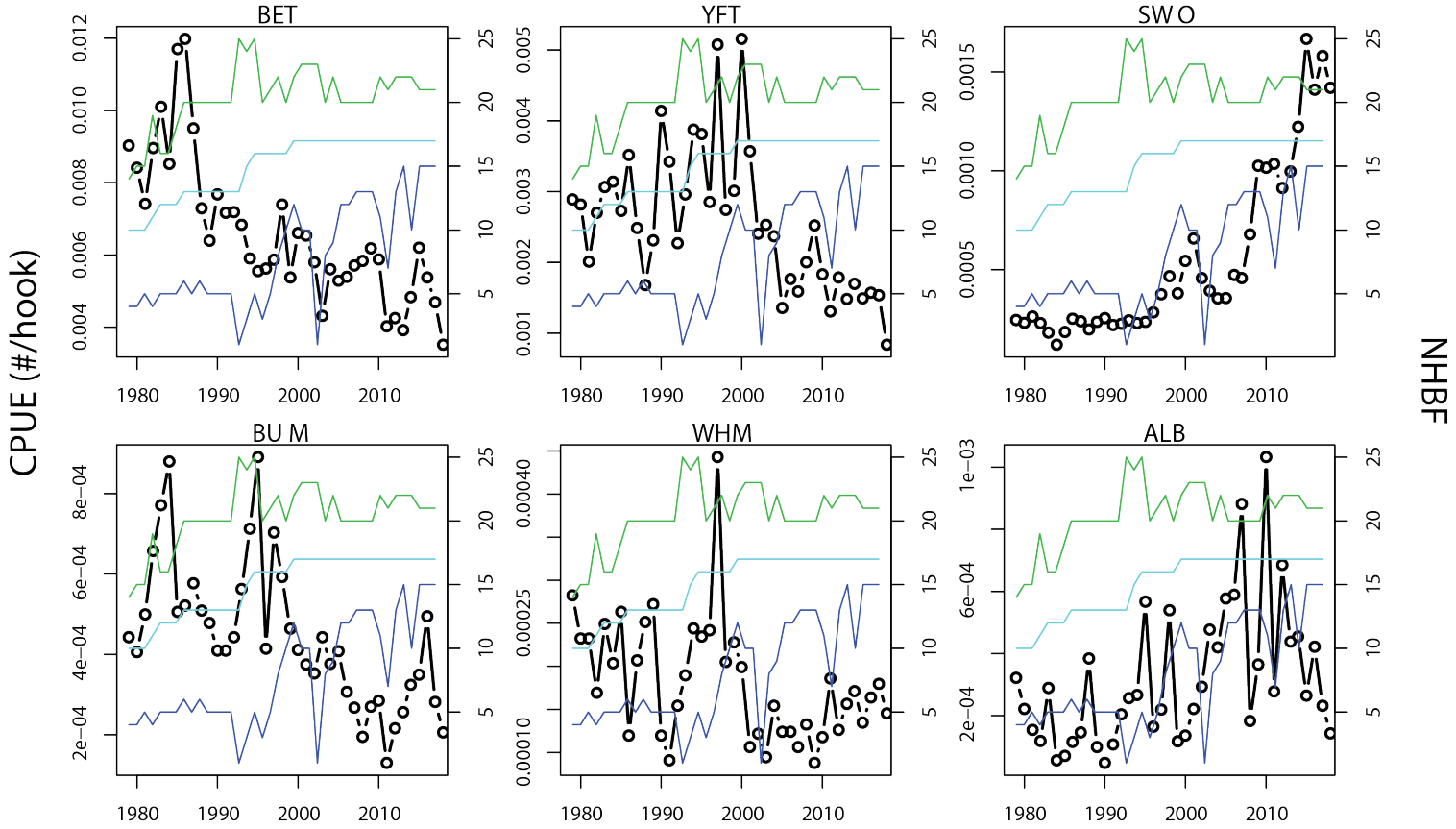
Estimation of Targeting Effects in the EPO Using Different Methods

Longline CPUE Indices Workshop, La Jolla, California, USA, 11-15 February 2019

Motivation

- Changes in targeting affect catch composition, and possibly relative abundance trends estimated from fisheries data.



JPN, area A1
1979-2017

NHBF color legend
min: dark blue; median: light blue; max: green

Presentation overview

- Present four approaches implemented to estimate targeting *outside* of the CPUE standardization model:
    - Cluster analysis of proportion species catch (Hoyle *et al*. method);
    - Gaussian mixture analysis of relative BET CPUE residuals (Okamura *et al*. method);
    - Hybrid method (cluster analysis of relative CPUE residuals for multiple species).
    - Potential Target Species method (Satoh)
- Discuss preliminary results and future work.

Methods implemented

Hoyle *et al*. 2018 (SCR5 058) method

- Data gridded to call sign x month x year
- Compute proportion species catch, $p_{sj} = \frac{c_{sj}}{\sum_r c_{rj}}$, c = catch, s = species, j = grid cell
- Proportions centered and scaled: $\tilde{p}_{sj} = \frac{(p_{sj} - p_{s.})}{\sigma_s}$
- Agglomerative hierarchical clustering applied to $\{\tilde{p}_{sj}\}$
- Target variable: cluster id (from pruned dendrogram).

Okamura *et al*. 2018 (CJFAS) method

- Overall model: $\log(cpue_{si}) = \mathbf{X'}_{si}\boldsymbol{\alpha}_s + \boldsymbol{Z'}_{si}\boldsymbol{\beta}_s + \epsilon_{si}$

  X = covariates *unrelated to* targeting, Z = covariates related to targeting, i = set.
- Estimate target variable:
  - Fit: $\log(cpue_{si}) = \mathbf{X'}_{si}\boldsymbol{\alpha}_s + \delta_{si}$
  - Compute residuals: $v_{si} = \log(cpue_{si}) - \mathbf{X'}_{si}\boldsymbol{\alpha}_s$
  - Obtain relative residuals for species of interest (e.g., BET): $z_i = logit\left(\frac{\exp(v_{si})}{\sum_r \exp(v_{ri})}\right)$
  - Assume Gaussian mixture for z_i: $f(z_i) = \sum_k \pi_k N(\mu_{ki}, {\sigma_k}^2)$, for $\mu_{ki} = \boldsymbol{M'}_{ki}\boldsymbol{\omega}_k$, $\boldsymbol{\omega}_k$ targeting-related covariates.
  - Targeting variable: $t_i = argmax_h \frac{\hat{\pi}_h f(z_i, \mu_{hi} \dots)}{\sum_l \hat{\pi}_l f(z_i, \mu_{li} \dots)}$

Methods implemented

Hybrid method

- Compute relative residuals for all species, $z_{sj} = logit[median_{i \in cell j}\left(\frac{\exp(v_{si})}{\sum_r \exp(v_{ri})}\right)]$
- Agglomerative hierarchical clustering applied to $\{z_{sj}\}$
- Target variable: cluster id (from pruned dendrogram)

- **Satoh Potential Target Species (PTS) method**
- For each set in a 5° area x quarter x year, compute $PTS_{si} = \begin{cases} 0 \; if \; CPUE_{si} \leq p^{th} \; percentile \; of \; \{CPUE_{si}\} \\ 1 \; if \; CPUE_{si} > p^{th} \; percentile \; of \; \{CPUE_{si}\} \end{cases}$
- A set can have PTS_{si} = 1 for a single or multiple *s*, or PTS_{si} = 0 for all *s*.
- Retain those sets for which $\sum_s PTS_{si}$ = 1.
- Use CART to build a classification algorithm relating PTS_{si} to covariates.
- Predict *PTS* for all sets based on this classification algorithm.

Some method details

- All analyses:
  Japanese longline data, 1979 - 2017
  Six species: ALB, BET, BUM, WHM (MLS), SWO, YFT
  Area A1 (10°S-10°N, 110°W-150°W)
  Excluded data with reported targets of SWO or sharks
  Excluded data that did not catch any of the 6 species

- Okamura *et al*. method:
  Delta-lognormal GAM for CPUE
  each component: ~ year effect + te(lat, lon, k=small)
  Relative residuals: based on deviance residuals
  Covariates for Gaussian mixture: month, NHBF

- Satoh method:
  Targeting threshold: 85th percentile of CPUE
  Covariates: quarter (month), 5° latitude, 5° longitude, NHBF

- All methods implemented in R; packages *fastcluster, mgcv, flexmix, rpart* and *randomForest.*

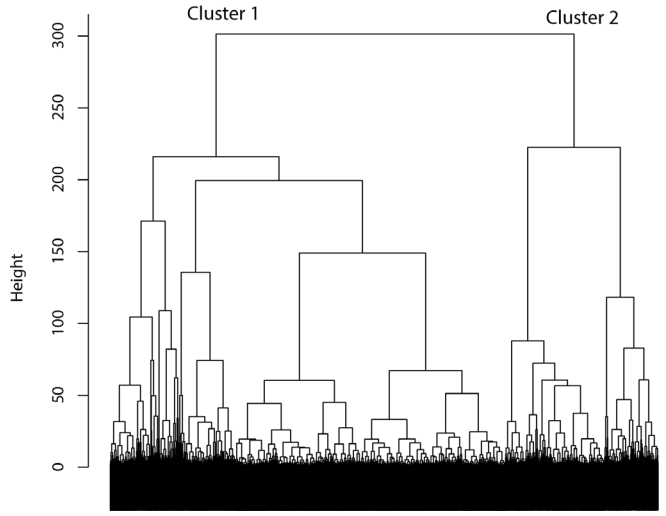
Preliminary results: PTS method

- After removing sets with no PTS or multiple PTS, 35% of sets retained.
- The CART tree based on the 1-se rule had one terminal node.
- A random forest (RF) algorithm was built to predict PTS.
- Little PTS predictive ability was obtained using covariates: quarter (month), 5° latitude and longitude, and NHBF.
- The low misclassification error for ALB is likely due to a strong relationship with latitude.

PTS	RF misclassification error
ALB	0.13
BET	0.80
BUM	0.82
WHM	0.89
SWO	0.90
YFT	0.84

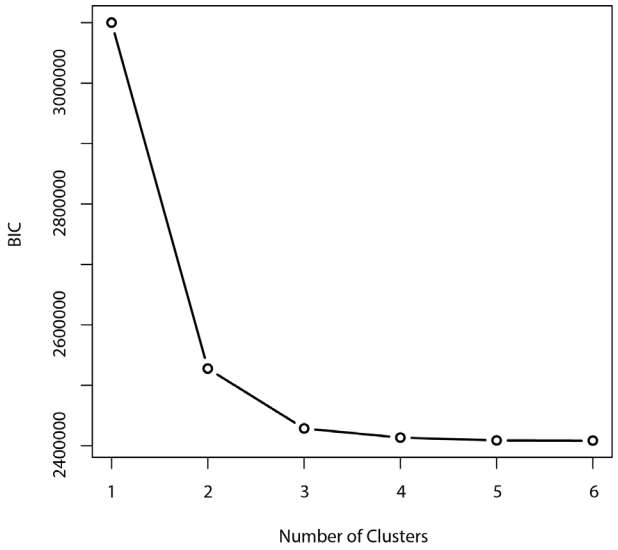
Variable importance	ALB	BET	BUM	WHM	SWO	YFT
month.fac	0.083	0.022	0.022	0.005	0.017	0.017
latc5	0.316	0.017	0.020	0.001	0.013	0.011
lonc5	0.073	0.007	0.008	-0.001	0.003	0.011
nhbf	0.056	0.010	0.006	0.005	0.005	0.009

Preliminary results: other methods



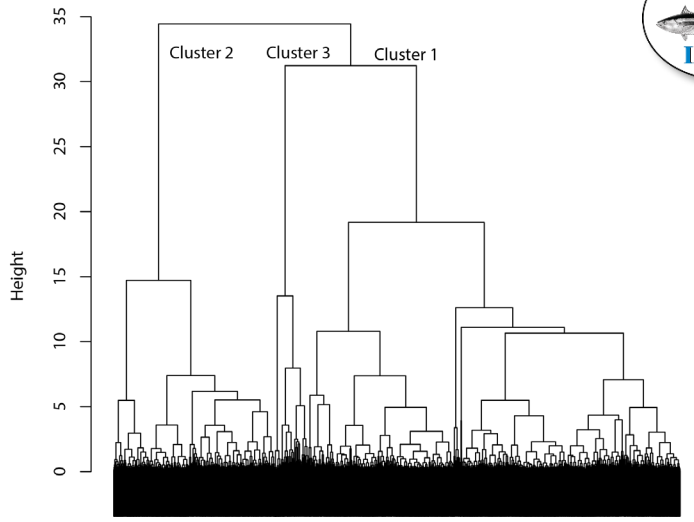
Hoyle et al. method

Cluster	# Sets	# Unique call signs
1	506,003 (73%)	1,122
2	187,505 (27%)	991
Total	693,508	1,158



Okamura et al. method

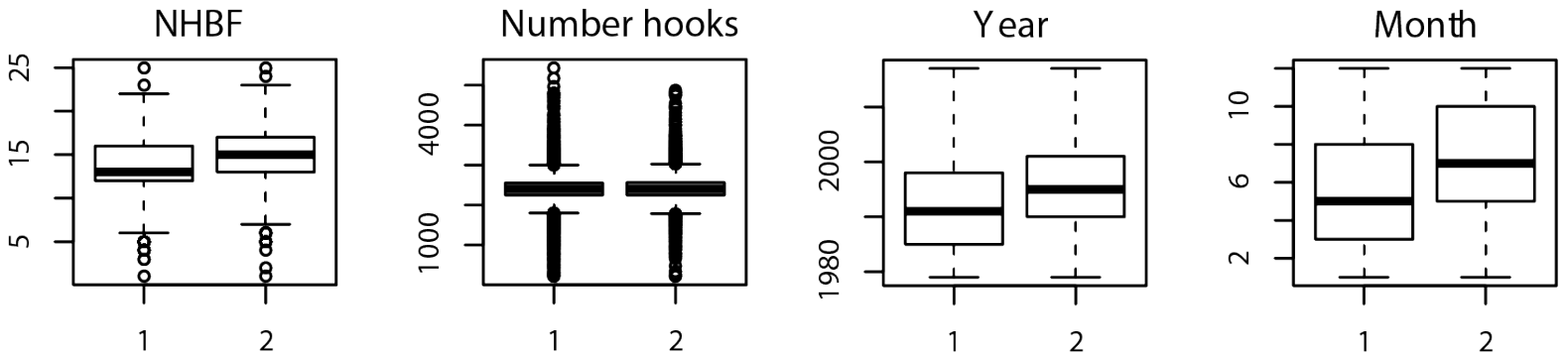
Cluster	# Sets	# Unique call signs
1	466,030 (67%)	1,152
2	181,507 (26%)	1,145
3	45,971 (7%)	1,105
Total	693,508	1,158



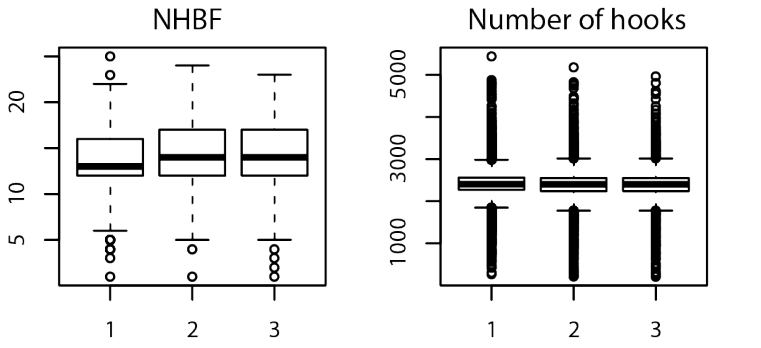
Hybrid method

Cluster	# Sets	# Unique call signs
1	487,088 (70%)	1,120
2	176,062 (25%)	1,026
3	30,358 (4%)	711
Total	693,508	1,158

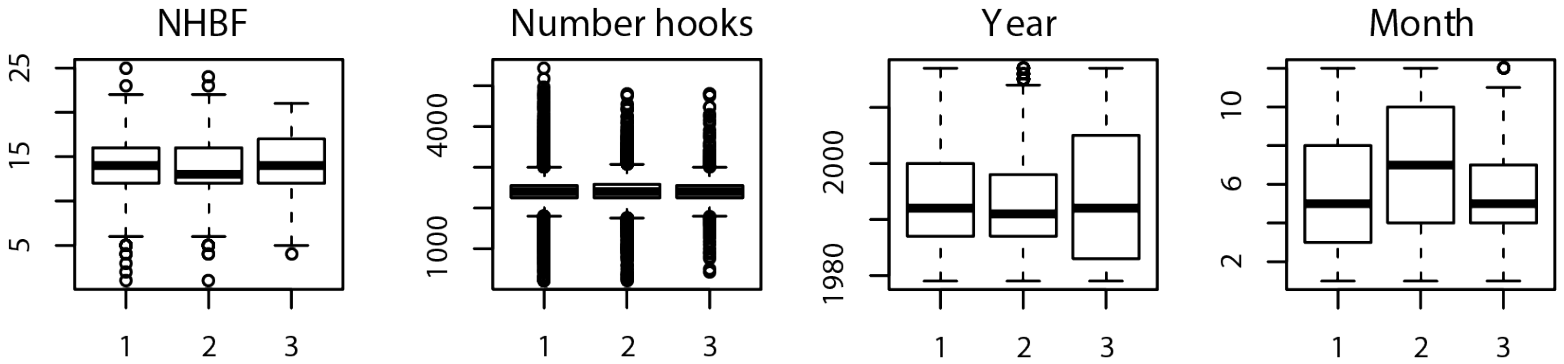
Preliminary results: cluster characteristics



Hoyle et al. method

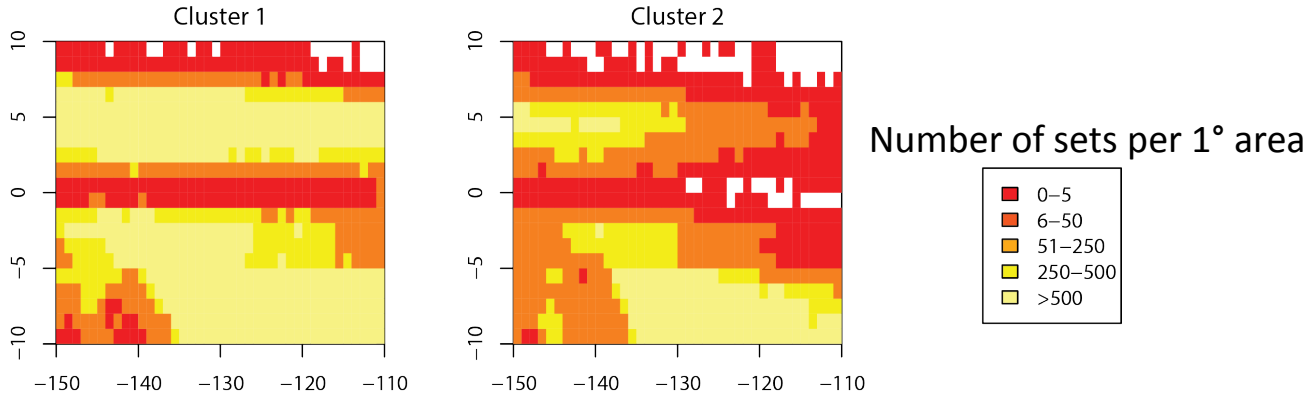


Okamura et al. method

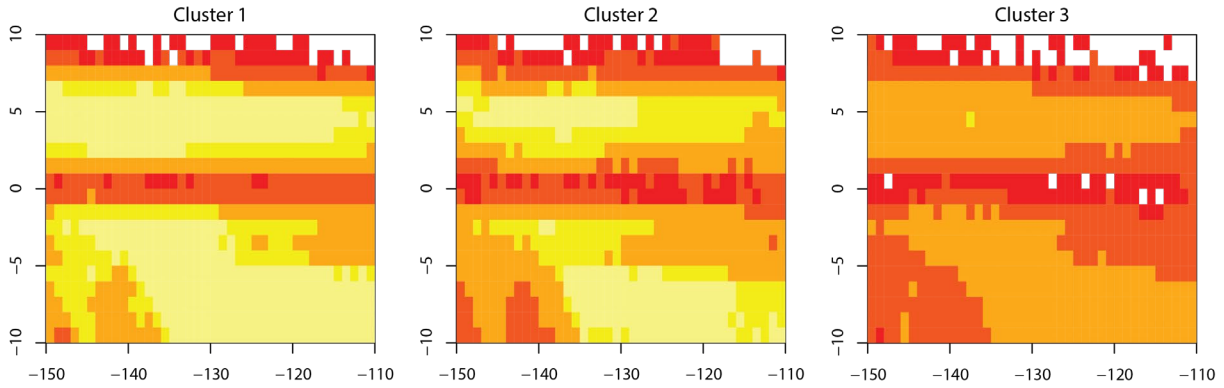


Hybrid method

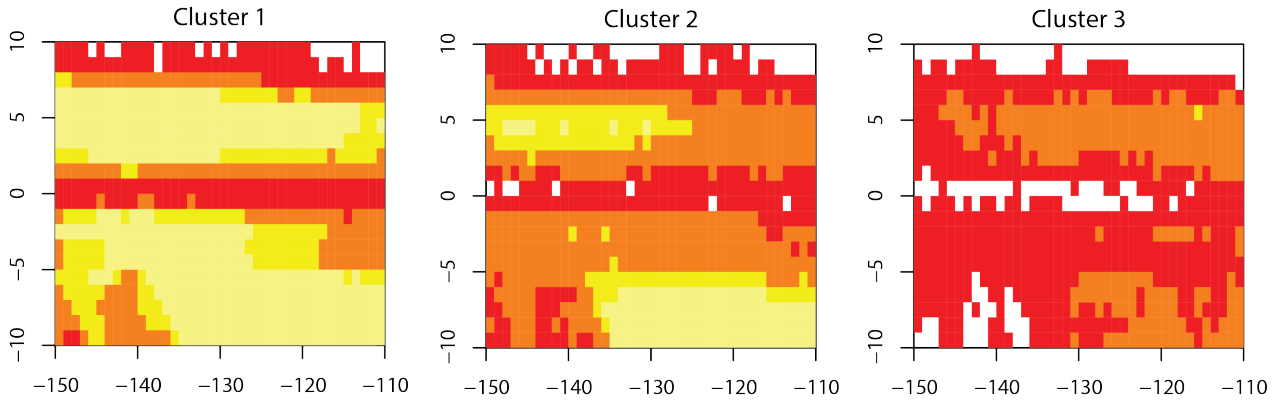
Preliminary results: spatial distribution of clusters



Hoyle et al. method

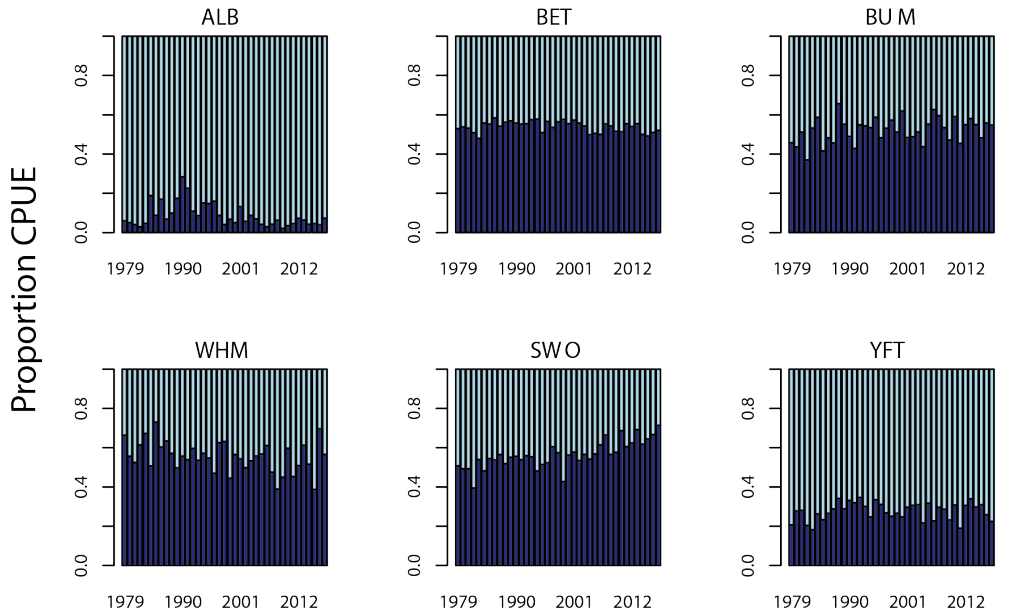


Okamura et al. method

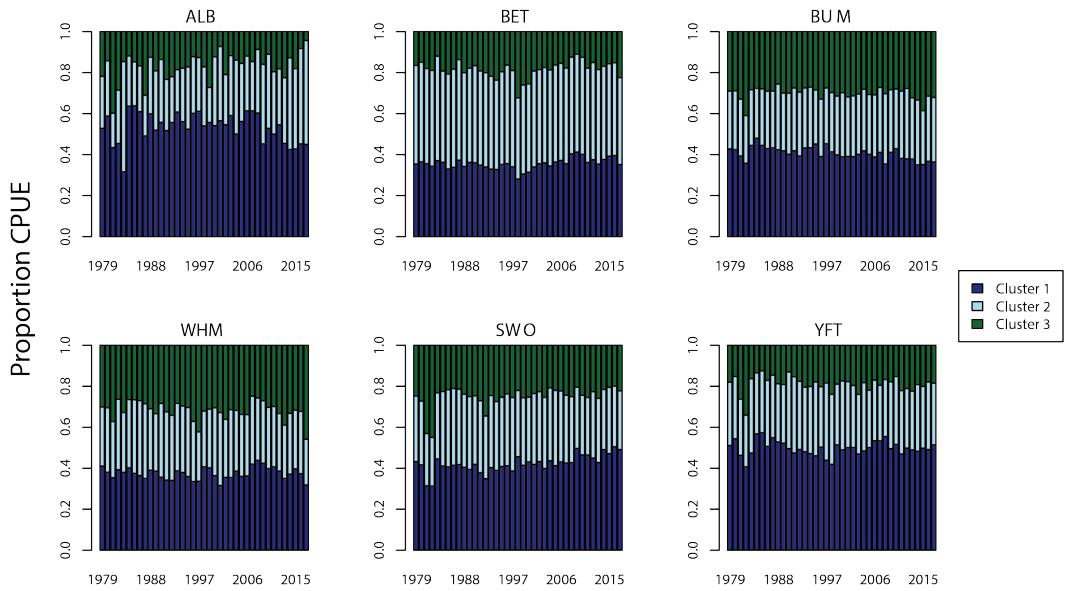


Hybrid method

Preliminary results: proportion species annual CPUE by cluster

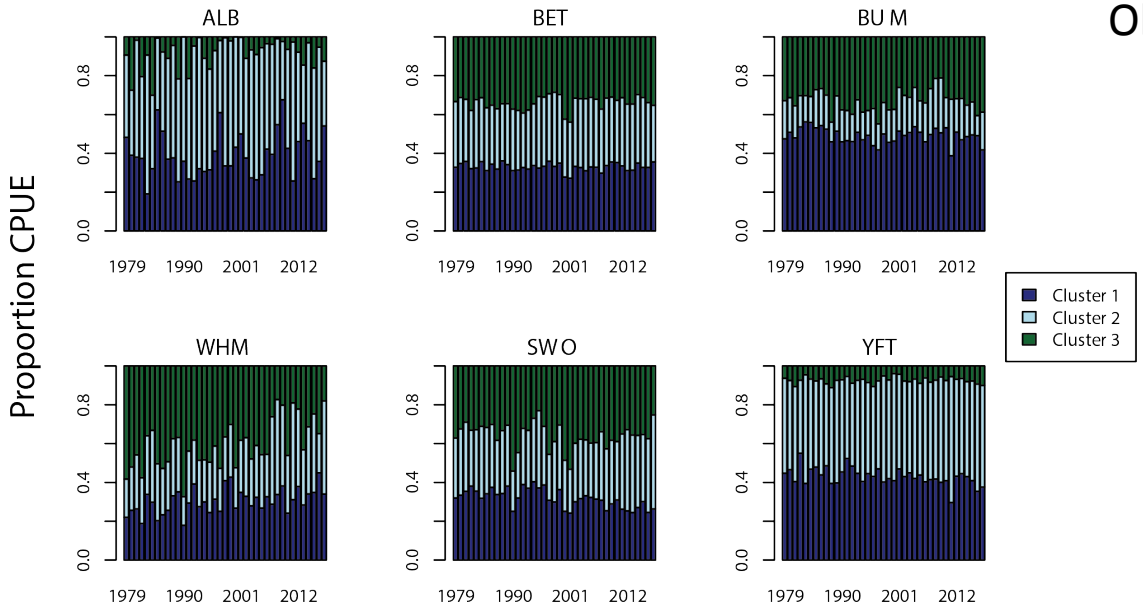


Hoyle et al. method



Okamura et al. method

Hybrid method



Proportion CPUE for species s

$$= \frac{CPUE_s}{\sum_r CPUE_r}$$

Preliminary results: comparison of cluster assignments

- There appears to be little correspondence between cluster assignments for these methods.

	Hybrid		
Okamura et al.	1	2	3
1	0.70	0.27	0.03
2	0.69	0.24	0.07
3	0.74	0.20	0.06
Hybrid proportions	0.70	0.25	0.04

	Hoyle et al.	
Hybrid	1	2
1	0.73	0.27
2	0.68	0.32
3	0.97	0.03
Hoyle et al. proportions	0.73	0.27

	Hoyle et al.	
Okamura et al.	1	2
1	0.71	0.29
2	0.78	0.22
3	0.73	0.27
Hoyle et al. proportions	0.73	0.27

Summary of preliminary results

- It appears BET may have always been a target in area A1 during 1979-2017.

- Given the temporal changes in CPUE, this could indicate that changes in fishing strategies to catch other species (e.g., secondary targets) do not strongly impact the ability to catch BET.

- The greatest contrast in proportion CPUE among clusters was seen for the Hoyle *et al*. method and least for Okamura *et al*. method.

What's next

- Apply methods to other longline fleets and other assessment areas.
- Run sensitivity analyses with respect to configuration of the two components of the Okamura *et al*. method (e.g., covariates and smoothing in GAM; covariates used in Gaussian mixture, etc).
- Simulations
- Investigate possible improvements to the hybrid method, such as:
    - Fitting a multivariate Gaussian mixture to multiple species residuals;
    - Developing an iterative fitting procedure to better separate targeting effects from density effects.

Thank you! Questions?