NPFC-2024-TWG CMSA08-WP06
January 22-25, 2024
@Niigata, Japan

Standardized abundance index for recruitment of chub mackerel from Northwest Pacific summer surveys up to 2023

Shota Nishijima, Yasuhiro Kamimura, Momoko Ichinokawa, Ryuji Yukami

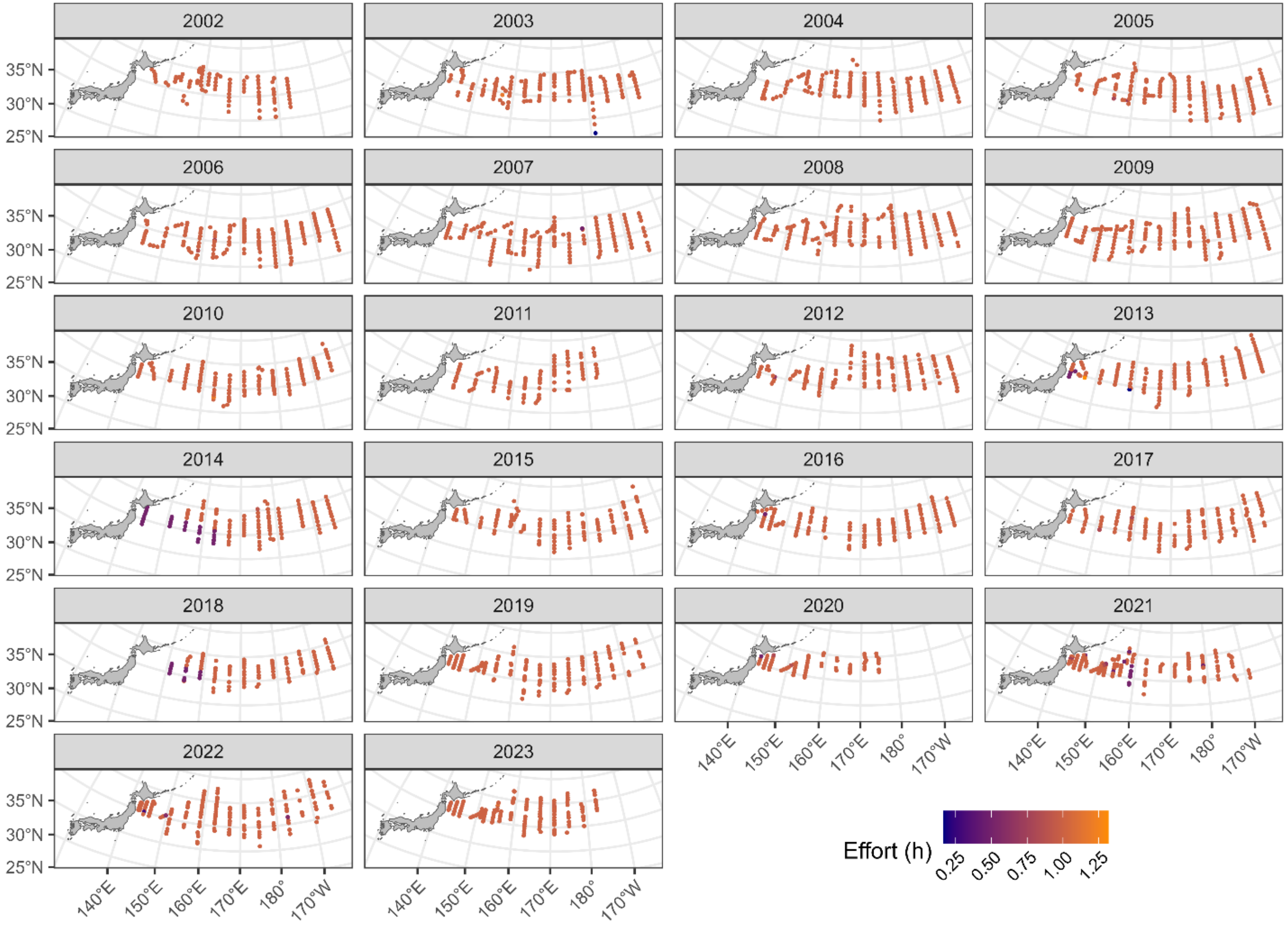
Fisheries Resources Institute,
Japan Fisheries Research and Education Agency (FRA)

Summary

- We conducted CPUE standardization of surface trawl surveys in summer for Pacific chub mackerel using the Vector Autoregressive Spatio-Temporal (VAST) model.
- We estimated local densities of young-of-the-year fish in the Northwest Pacific from 2002 to 2023 with consideration for environmental factors of sea surface temperature (SST) and 50m-depth temperature as well as spatial autocorrelation
- The analysis showed high levels of recruitment index have frequently occurred since 2013
- Model diagnostics found no serious problems in residual patterns.
- We propose this standardized recruitment index to be used as the abundance index of age 0 fish in the Technical Working Group for the Chub Mackerel Stock Assessment (TWG CMSA).

Summer surveys by Japan

Fig. 1A



- Japan (FRA) has conducted sea surface trawl surveys in the Northwest Pacific Ocean from June to July annually to collect biological and abundance information on small pelagic fish
- The standardized CPUE (catch number divided by sweeping time) of age 0 fish of CM has long been used as a recruitment index in the Japanese domestic stock assessment

Changes since the last document

	Previous WP	Current WP
Model	Delta-GLM-tree (Hashimoto et al. 2019)	VAST (Thorson et al. 2019)
Environmental covariate	SST, 50m-depth temperature (T50)	Principal components (PC1, PC2)
Years	2002-2021	2002-2023

- A previous paper showed that VAST demonstrated superior overall performance in CPUE standardization compared to generalized linear models or generalized additive models (Grüss et al. 2019)
- VAST was found to outperform the delta-GLM-tree in terms of Akaike Information Criterion (AIC) (Yukami et al. 2023)
- Used principal component analysis (PCA) to resolve a high correlation between SST and T50
- We extended the duration of years into 2023

Table 1

Year	Number of observations (stations)	Total sweeping time (h)	Total swept area (km^2)	Total catch (ind)	Number of observations with positive catch	Percentage of positive catch (%)
2001	58	59.00	12.02	113.5	9	15.52
2002	93	93.00	18.26	259.0	17	18.28
2003	157	155.37	30.55	4063.8	15	9.55
2004	179	178.50	36.35	21262.5	24	13.41
2005	164	162.95	31.12	2389.0	16	9.76
2006	163	162.63	30.19	39.0	3	1.84
2007	155	154.50	29.58	36441.0	24	15.48
2008	169	169.00	33.08	6024.0	16	9.47
2009	168	168.02	39.43	5568.0	25	14.88
2010	126	126.18	24.88	2504.0	18	14.29
2011	97	97.00	17.48	363.5	12	12.37
2012	135	134.85	25.12	4745.5	20	14.81
2013	125	122.48	26.27	183151.5	17	13.60
2014	122	108.95	20.29	884.8	5	4.10
2015	121	121.00	22.99	4358.6	19	15.70
2016	122	121.47	22.73	81005.6	32	26.23
2017	129	128.65	24.18	68441.9	18	13.95
2018	104	97.93	18.74	192845.9	23	22.12
2019	134	134.00	28.27	9998.5	26	19.40
2020	67	66.20	11.53	29231.4	28	41.79
2021	143	136.45	32.21	250694.6	60	41.96
2022	156	154.61	30.76	100144.9	55	35.26
2023	143	142.77	28.44	41228.2	53	35.33

- 100~300 individuals of 'mackerel' (chub + blue) were sampled per station, when more than 100 individuals were caught, for species identification and length measurement
- More than 100 stations except for 2001, 2002, 2011, 2020
- Sweeping time is almost one hour
- Percentages of positive catch were low than 30% until 2019, but became higher than 35% thereafter

Filtering rule

Filter Applied	Number of Records Remaining	Number Removed	Number of Records with Chub Mackerel Catch >0
Initial Data set	3,030	-	535
Remove data in 2001	2,972	58	526
Remove data with no SST	2,970	2	526
Remove data with no 50m-depth temperature	2,916	54	524

- Removed the samples of 2001 from analyzed data because the number of stations and covered range in the beginning year were small (N = 58)
- Removed samples with no SST (N = 2) or 50m-depth temperature (N = 54)
- The final sample size was 2,916

Map of catch and CPUE of age-0 CM fish

Fig. 1B: Catch

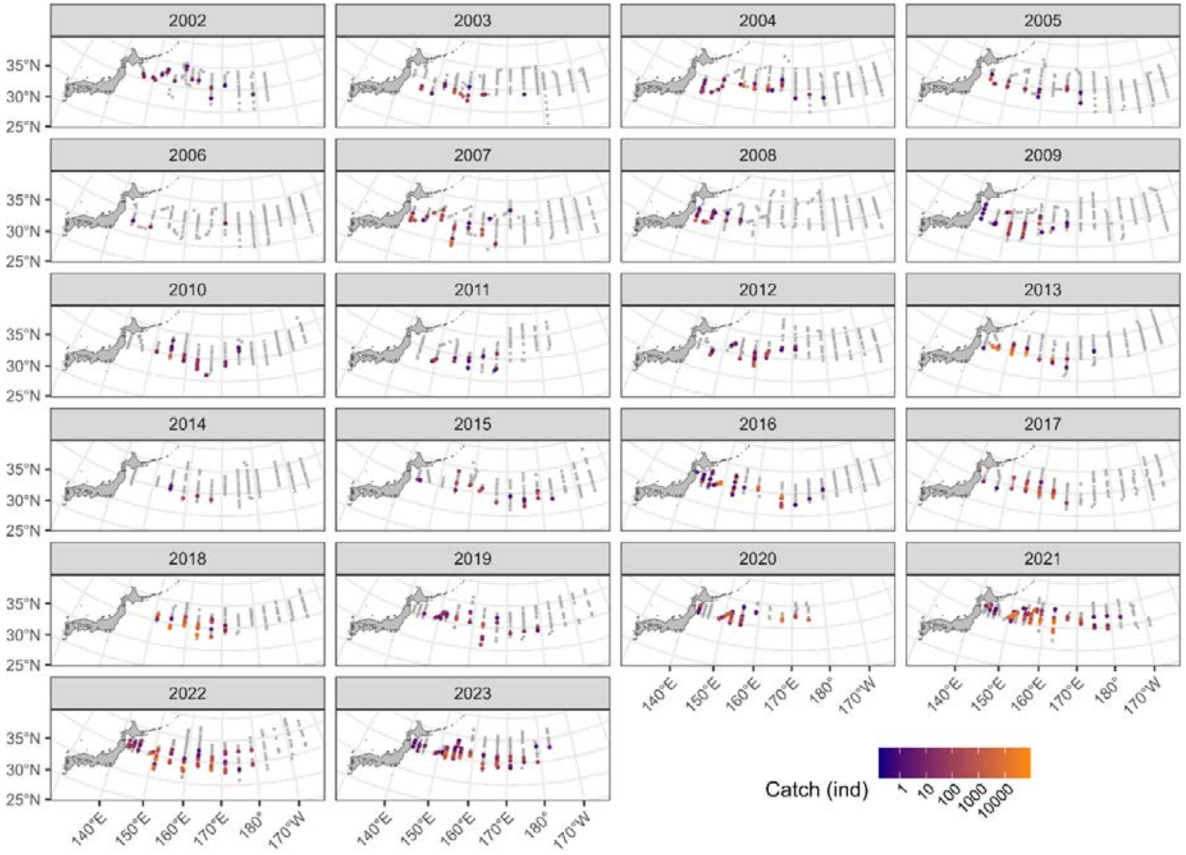
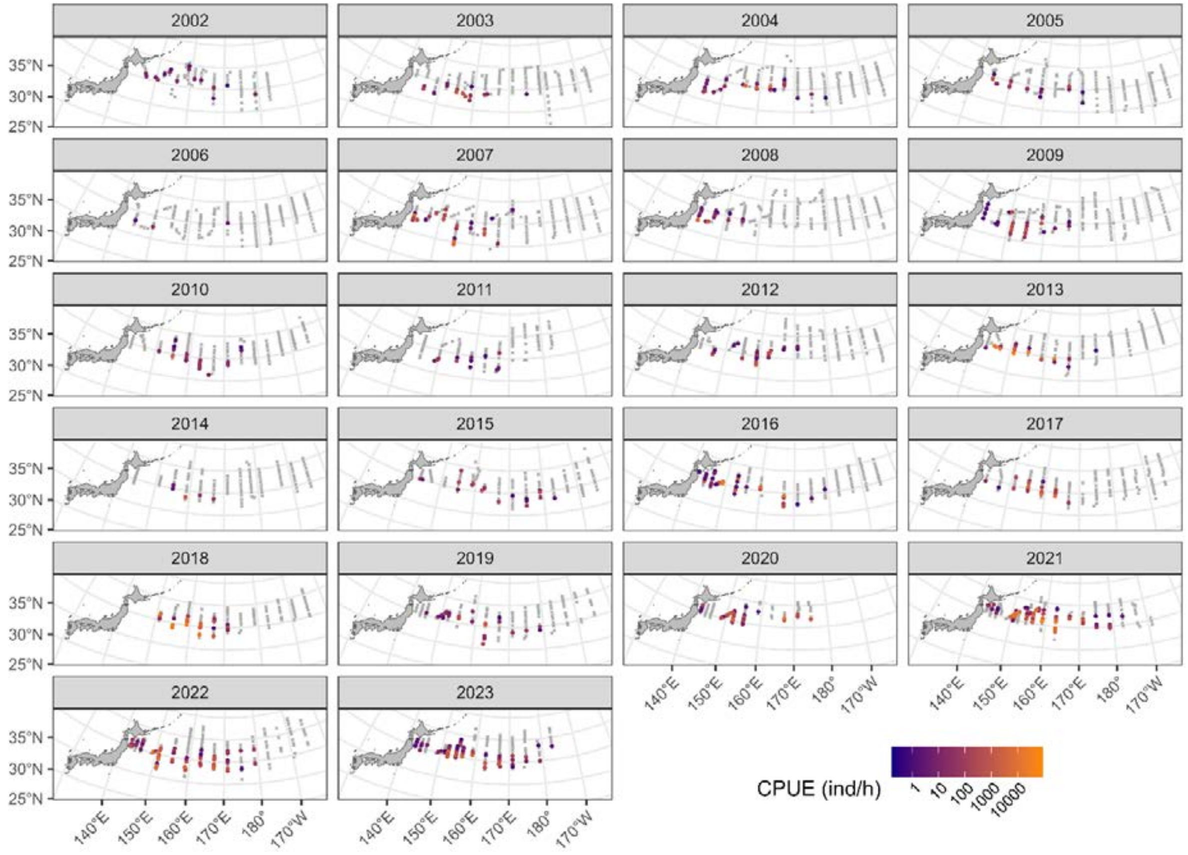


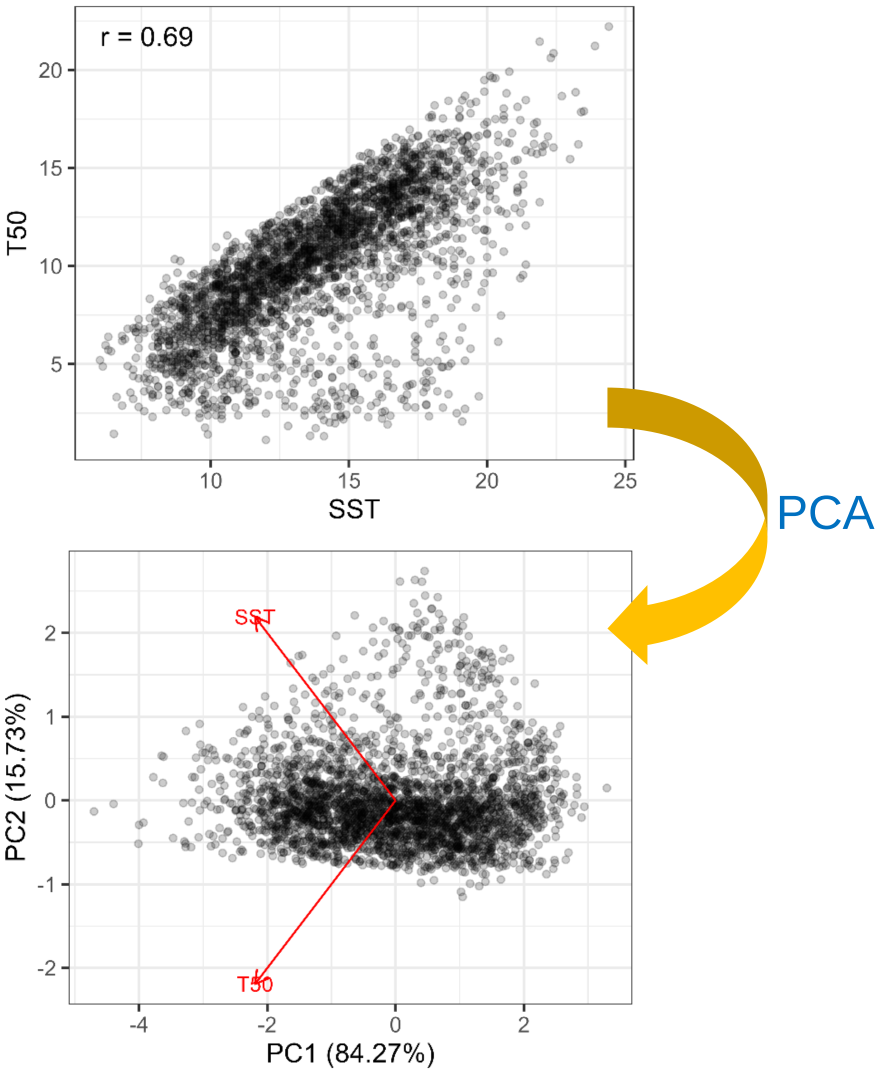
Fig. 1C: CPUE



- Catch and CPUE patterns are quite similar because of effort is almost 1 (hour)
- Age 0 fish of CM were likely to be caught in southern areas

Principal component analysis (PCA)

Fig. 2



- *In situ* SST and T50 were highly correlated with r = 0.69 of Pearson's correlation coefficient
- Such collinearity in multiple regression models could destabilize parameter estimates and prediction to new data, suggesting that it might be problematic in the interpretation of results and model predictions in CPUE standardization
- Conducted the PCA and used PC1 and PC2 calculated from the analysis as orthogonal covariates
- PC1 was negatively correlated with SST and T50, indicating a common component of SST and T50.
- PC2 was positively correlated with SST but negatively with T50, reflecting a difference between SST and T50.
- The proportion of variance of PC1 and PC2 were 84.3% and 15.7%, respectively

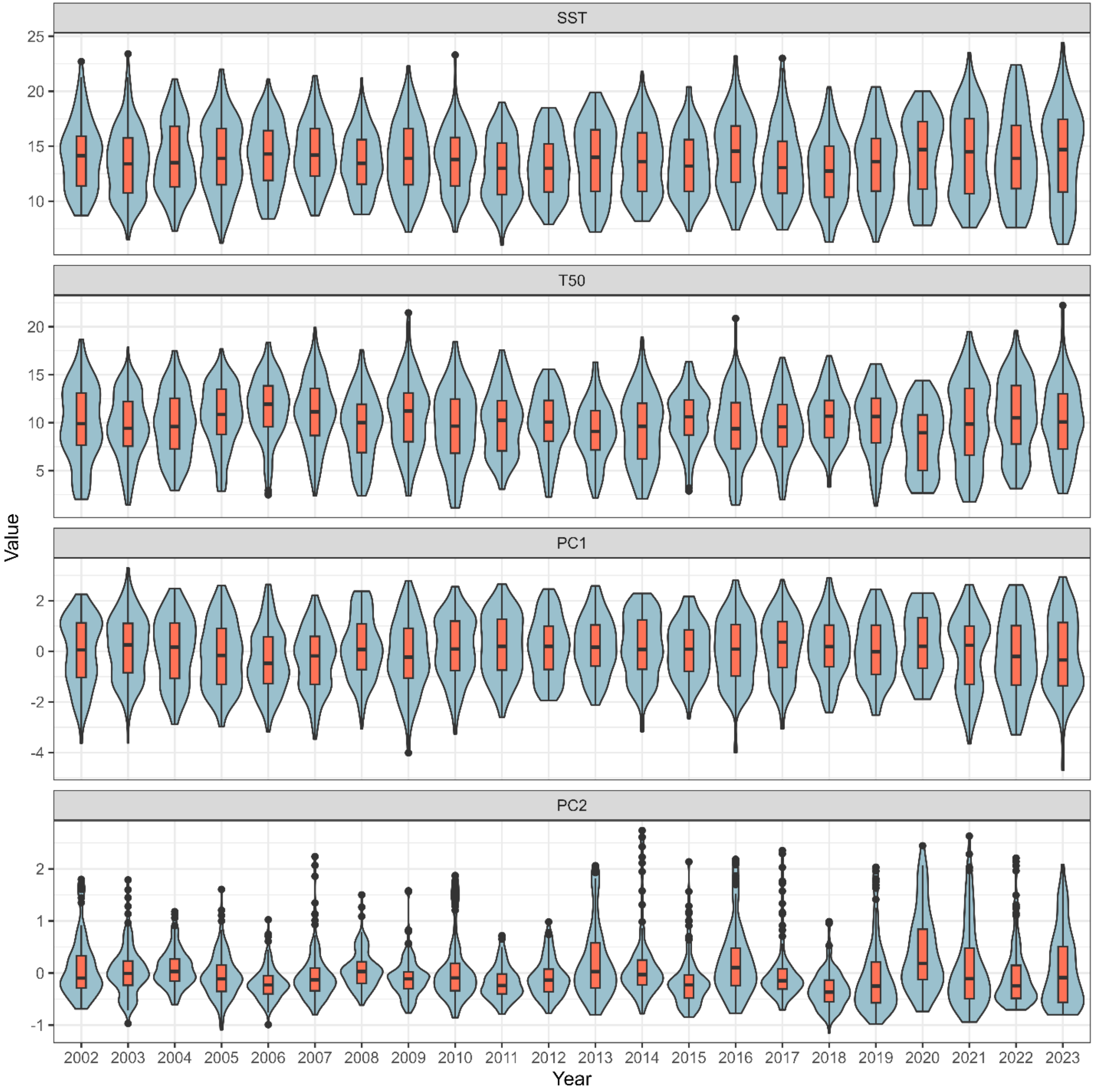


Fig. 3

SST, T50, PC1, and PC2 did not show any systematic patterns over the years

Spatial patterns of SST and T50 in each year

Fig. 4A

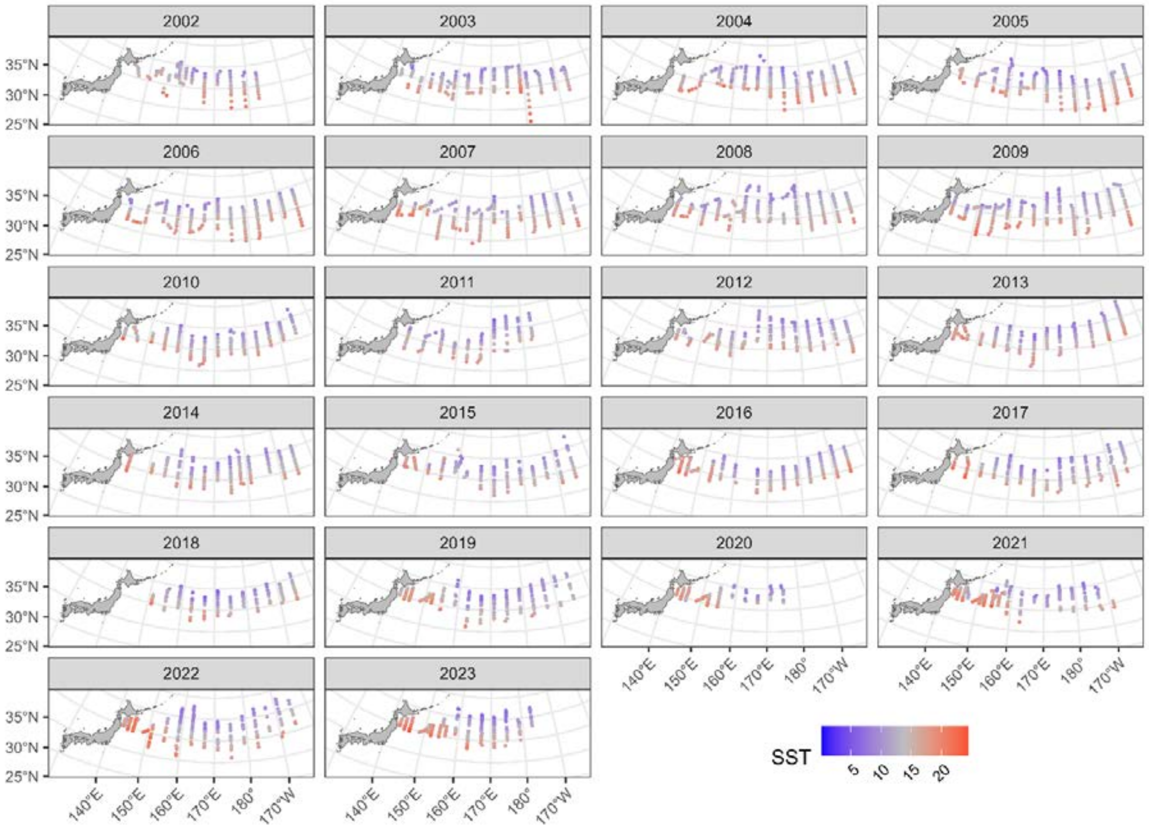
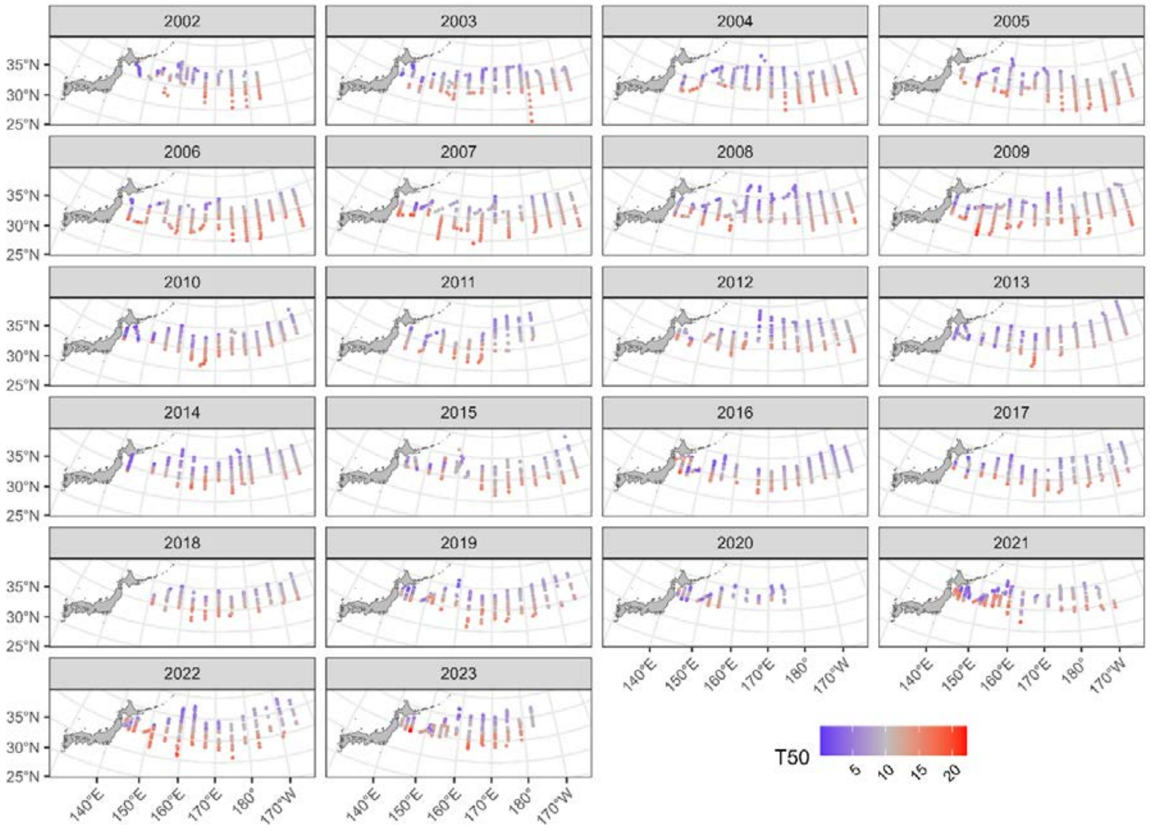


Fig. 4B



- SST and T50 tended to be higher in the south than in the north

Spatial patterns of PC1 and PC2 in each year

Fig. 4C

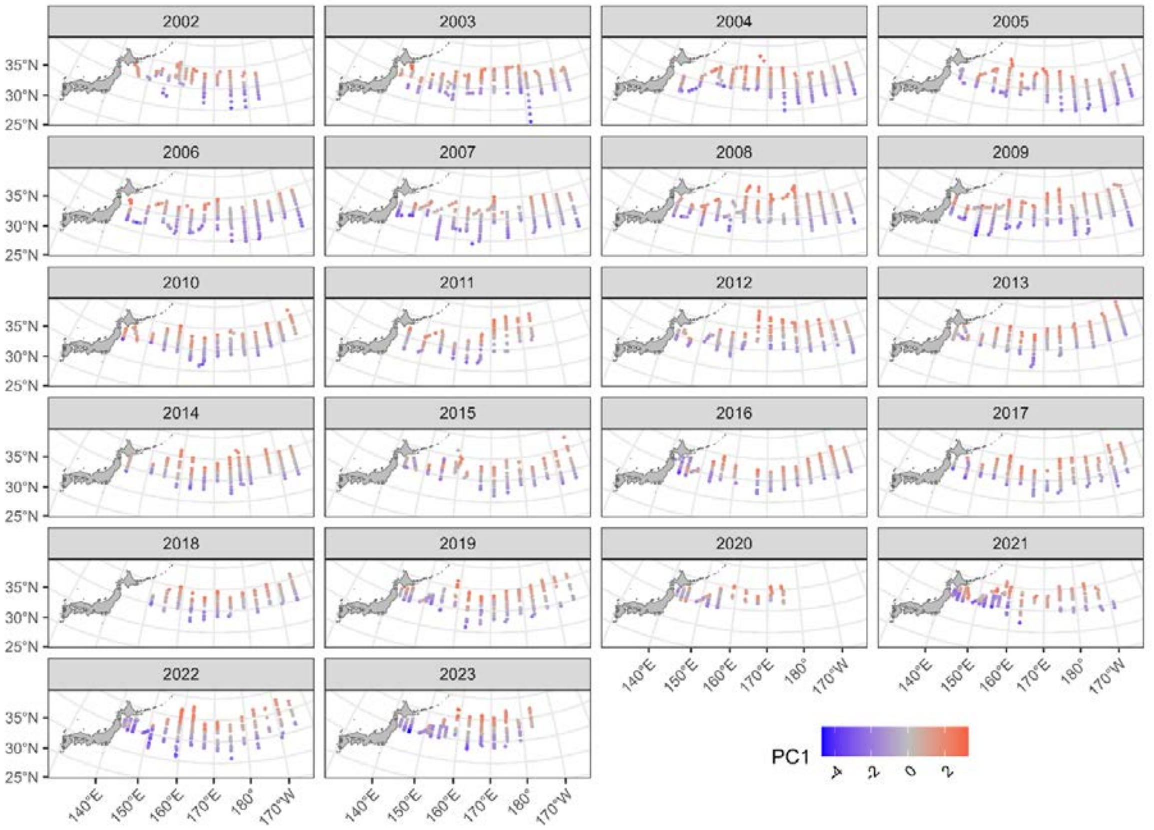
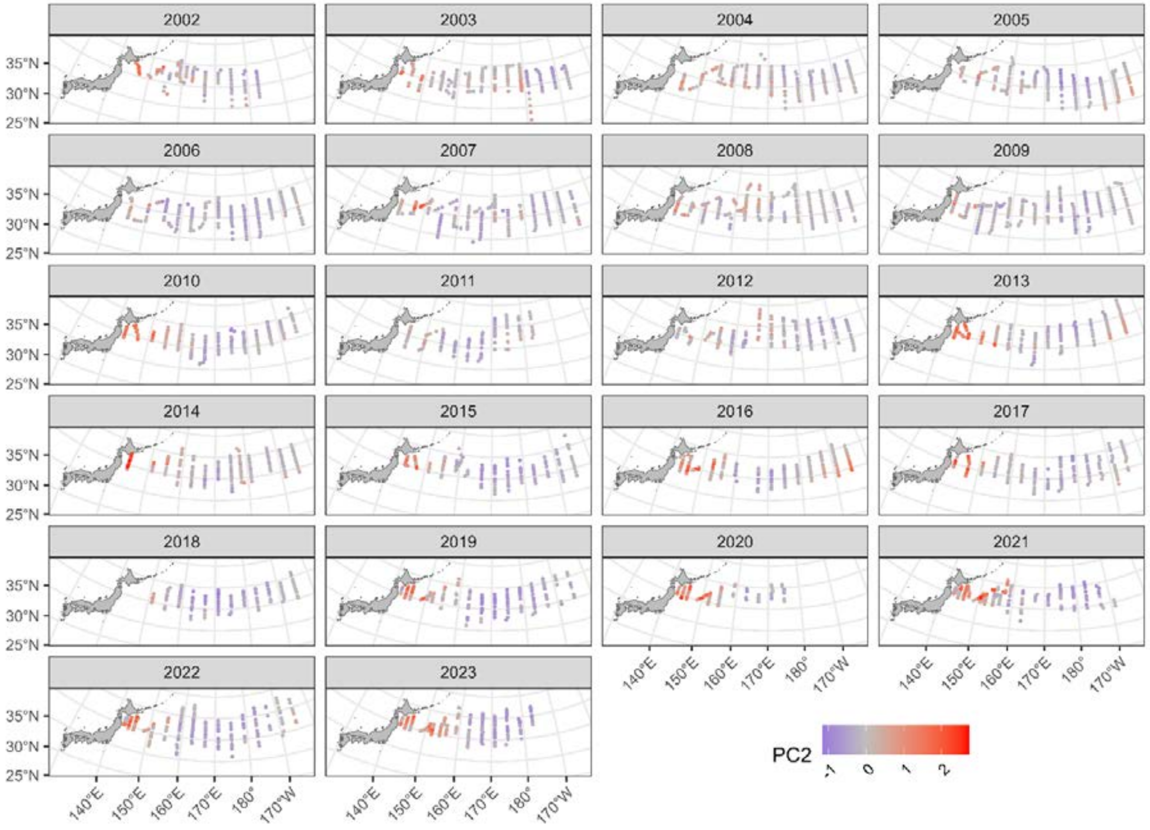


Fig. 4D



- PC1, which was negatively correlated with SST and T50, was thus higher in the north
- PC2 tended to be higher off the Pacific coast of Japan

Model description of the VAST

1st predictor for encounter probability

$$p_1(i) = \beta_1(t_i) + \omega_1(s_i) + \varepsilon_1(s_i, t_i) + \sum_{k_1}^{n_{k1}} \lambda_1(k_1) Q_i(i, k_1)$$

2nd predictor for positive catch rate when encountered

$$p_2(i) = \underbrace{\beta_2(t_i)}_{\text{temporal}} + \underbrace{\omega_2(s_i)}_{\text{spatial}} + \underbrace{\varepsilon_2(s_i, t_i)}_{\text{spatio-temporal}} + \underbrace{\sum_{k_2}^{n_{k2}} \lambda_2(k_2) Q_i(i, k_2)}_{\text{catchability covariate}}$$

The encounter probability transformed the inverse function of logit link

$$r_1(i) = \text{logit}^{-1} p_1(i)\ ,$$

The positive catch rate transformed the inverse function of log (i.e., exp)

$$r_2(i) = a_i \times \log^{-1} p_2(i)\ .$$

($a_i = 1$ in this study)

Binomial model

The probability density function

$$\Pr(b_i = B) = \begin{cases} 1 - r_1(i) & \text{if } B = 0 \\ r_1(i) \times g\{B \mid r_2(i), \sigma_m^2\} & \text{if } B > 0 \end{cases}$$

Function for Gamma distribution

Used covariates and other settings

Table 3

Variable	Symbol[1]	Number of categories	Detail	Note
Year	$\beta(t)$	22	2002-2023	Categorical variable with fixed effect
Spatial	$\omega(s)$	-	Average over years	Estimated as random effects by SPDE approximation
Spatio-temporal	$\varepsilon(s,t)$	-	Assume independence of each year	Estimated as random effects by SPDE approximation
PC1	$\lambda(k)Q_i(i,k))$	-	Negative correlation for SST and T50	Continuous variable as a catchability covariate
PC1 squared	$\lambda(k)Q_i(i,k))$	-	Squared PC1	Continuous variable as a catchability covariate
PC2	$\lambda(k)Q_i(i,k))$	-	Positive correlation for SST and negative correlation for T50	Continuous variable as a catchability covariate
PC2 squared	$\lambda(k)Q_i(i,k))$	-	Squared PC1	Continuous variable as a catchability covariate
PC1 X PC2	$\lambda(k)Q_i(i,k))$	-	Interaction between the two PC axes	Continuous variable as a catchability covariate

- The number of knots was set as 100
- The effect of year was estimated as a categorical variable by fixed effects
- PC1, PC2, their squared terms, and their 1st order interaction were treated as catchability covariates because it was assumed that they reflected local conditions at observation affecting catchability rather than abundance of the year

Model selection

Table 4

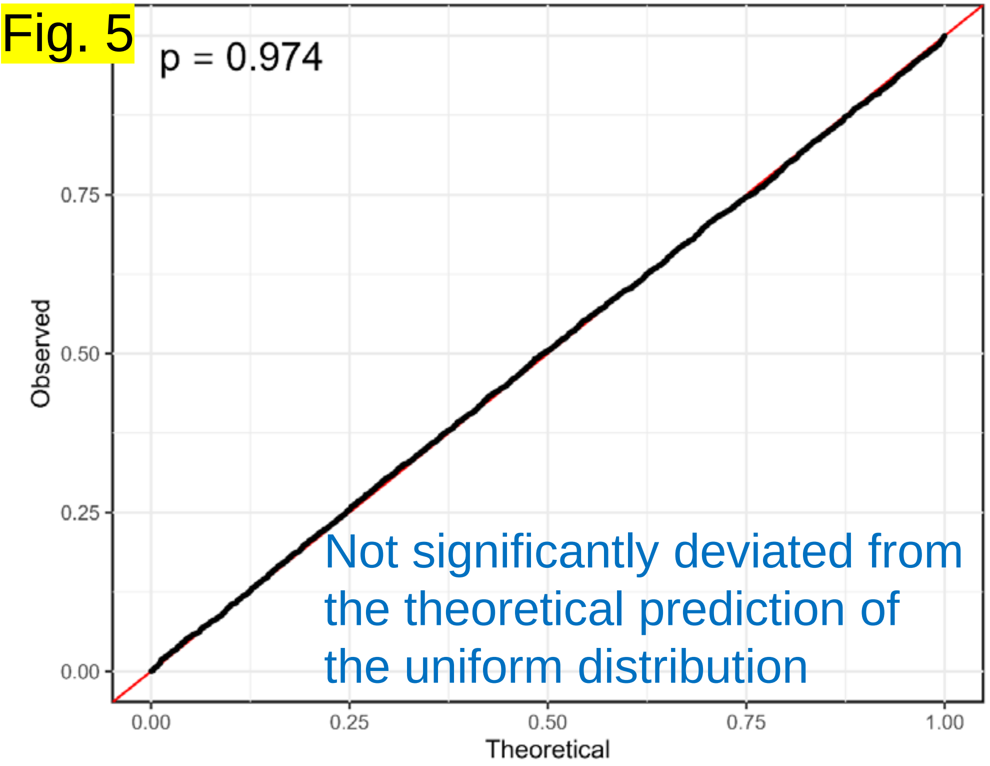
Rank	PC1	PC1 squared	PC2	PC2 squared	PC1×PC2	Df	logLik	AICc	ΔAICc
1	B,G	B	B,G	B	B	58	-4245.74	8609.88	0.00
2	B,G	B	B,G	B,G	B	59	-4244.84	8610.17	0.28
3	B,G	B,G	B,G	B	B	59	-4245.13	8610.73	0.85
4	B,G	B,G	B,G	B,G	B	60	-4244.28	8611.13	1.24
5	B,G	B	B,G		B	57	-4247.43	8611.17	1.29
6	B,G	B	B,G	G	B	58	-4246.53	8611.45	1.57
7	B,G	B	B,G	B,G	B,G	60	-4244.62	8611.80	1.92
8	B,G	B	B,G	B	B,G	59	-4245.74	8611.96	2.08
9	B,G	B,G	B,G		B	58	-4246.81	8612.02	2.13
10	B,G	B,G	B,G	G	B	59	-4245.97	8612.41	2.53
11	B,G	B,G	B,G	B	B,G	60	-4244.96	8612.49	2.61
12	B,G	B	B,G	G	B,G	59	-4246.30	8613.09	3.20
13	B,G	B,G	B,G	B,G	B,G	61	-4244.27	8613.19	3.30
14	B,G	B	B,G		B,G	58	-4247.43	8613.25	3.36
15	B,G	B,G	B,G		B,G	59	-4246.65	8613.77	3.89
16	B,G	B,G	B,G	G	B,G	60	-4245.95	8614.47	4.59
17	B,G	B	B	B	B	57	-4251.96	8620.24	10.36
18	B,G	B,G	B	B	B	58	-4251.46	8621.31	11.43
19	B,G	B	B		B	56	-4253.65	8621.53	11.64
20	B	B	B,G	B	B	57	-4252.66	8621.64	11.76

- Model selection was conducted using exhaustive search based on Akaike Information Criterion with correction (AICc).
- All the covariates were selected for encounter probability (B) in the top four models
- The linear effects of PC1 and PC2 were only selected for positive catch rate when encountered (G) in the best model
- The percent deviance explained was 57.5%.

Model diagnostics for scaled residuals

- Generated scaled residuals using the R package 'DHARMa' (Hartig 2022) for model diagnostics
- This package enables to simulate the scaled residuals which should theoretically follow the uniform distribution from zero to one

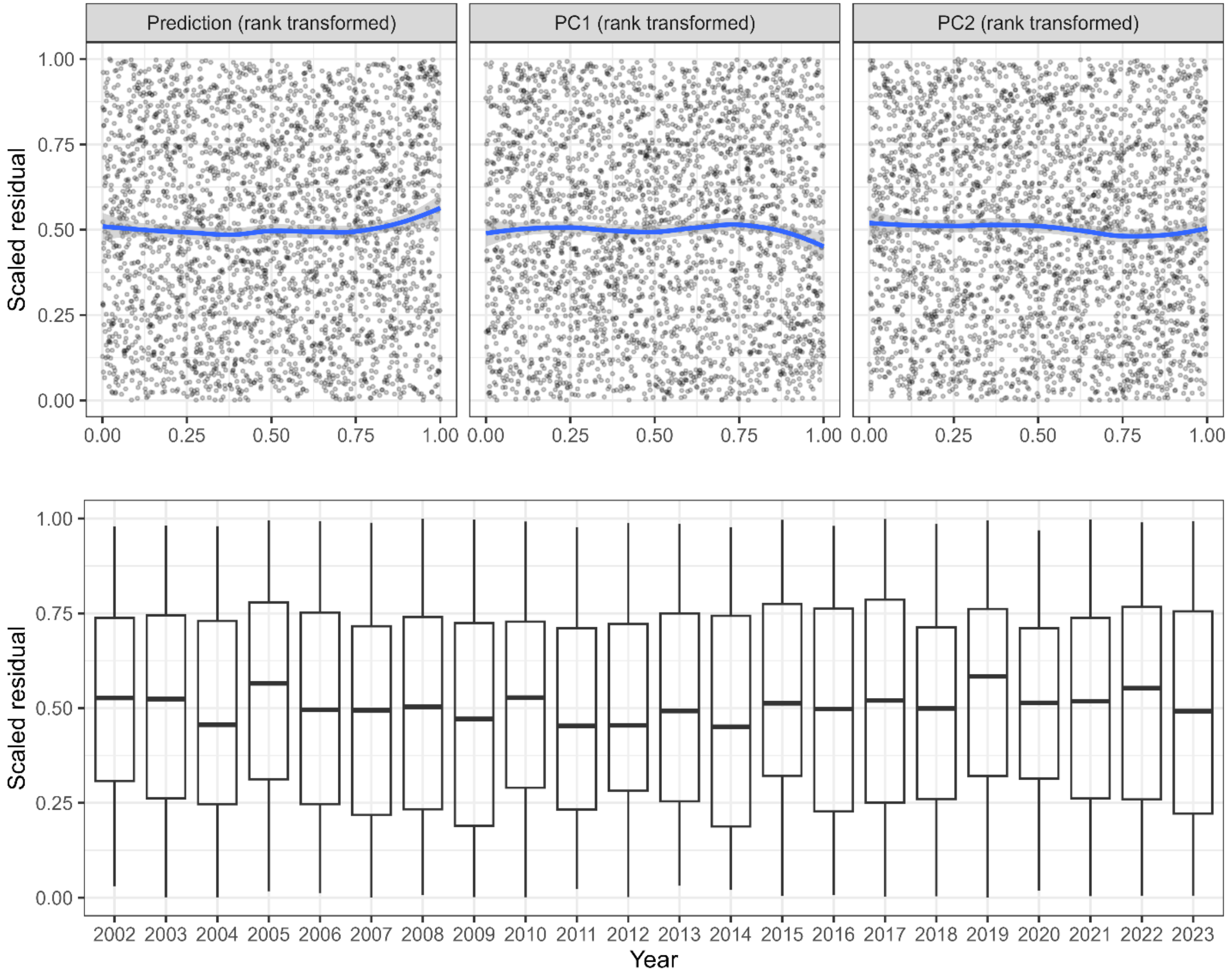
Fig. 5



Not significantly deviated from the theoretical prediction of the uniform distribution

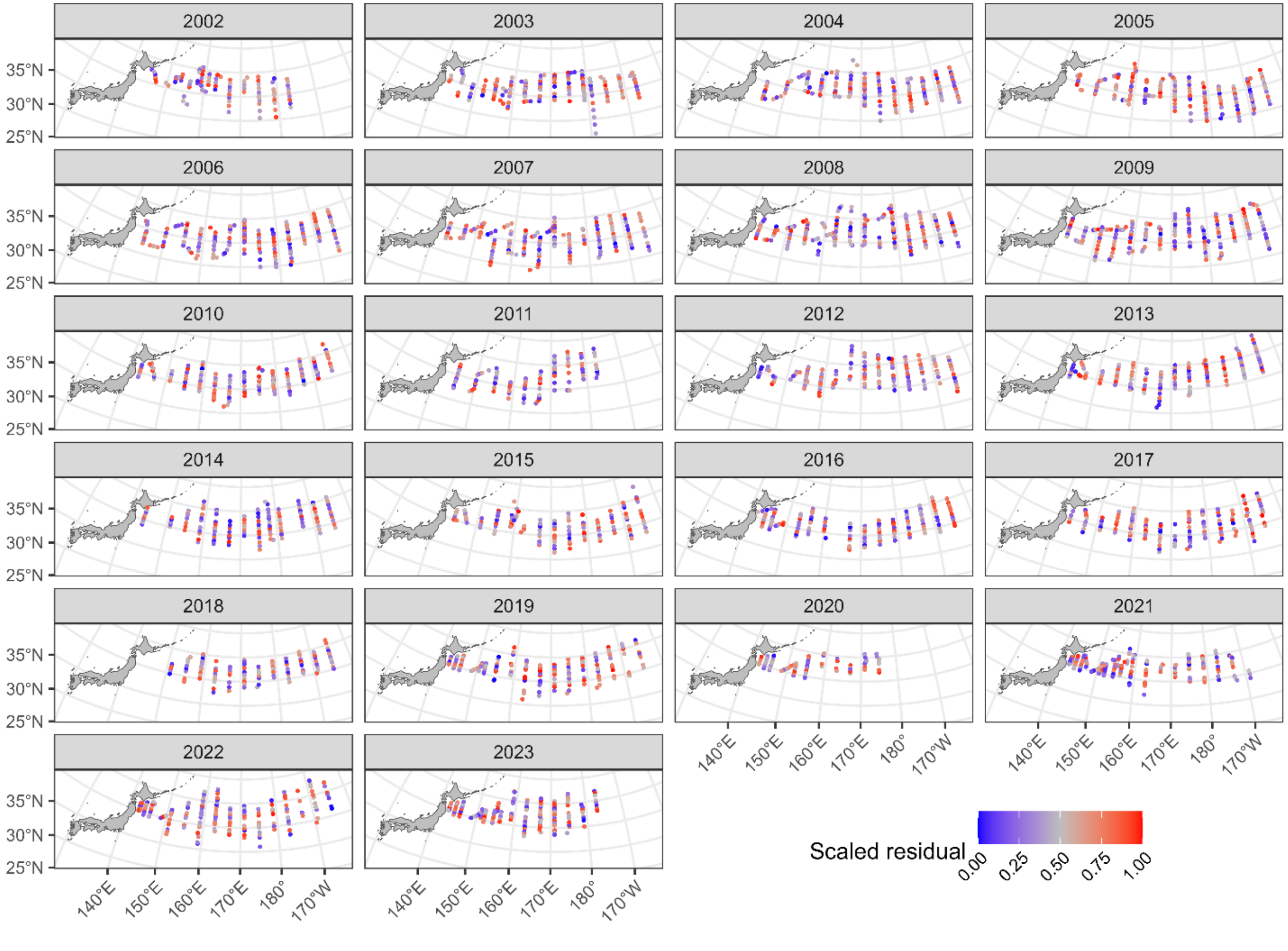
Fig. 6

The averages were not deviated from the theoretical average (0.5) in response to predicted values and covariates



Map of scaled residuals in each year

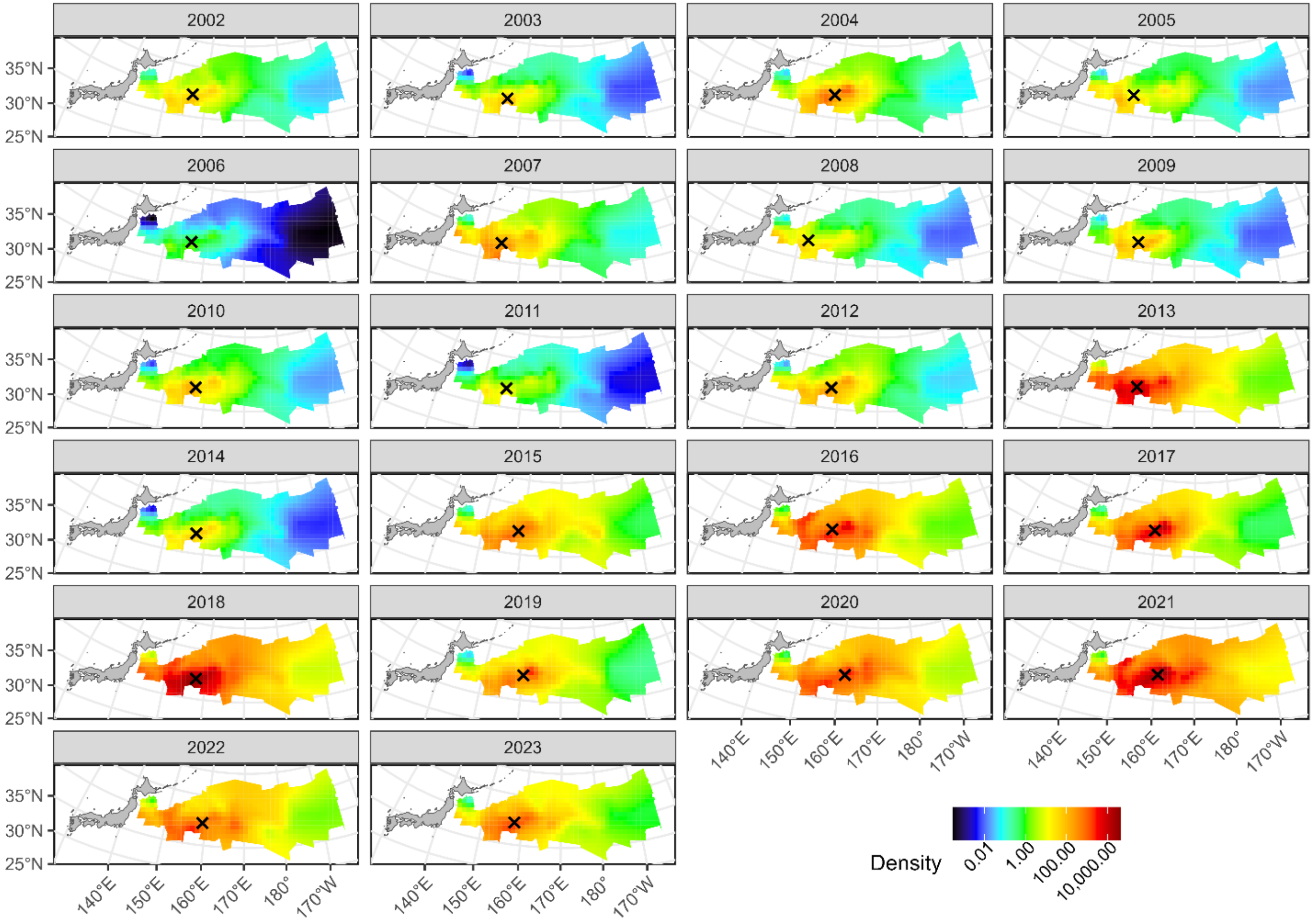
Fig. 7



No systematic spatial patterns in scaled residuals

Map of estimated densities

Fig. 8



- Local densities were estimated from the product of encounter probability and positive catch rate when encountered

$$d(s,t) = r_1^*(s,t) \times r_2^*(s,t)$$

- The terms of catchability covariates were dropped off (assuming λ = 0)

- Estimated densities of YOY fish were low until 2012, but increased thereafter

- The centroid of fish distributions was relatively constant over the years, averaging 157.4º E and 39.2º N

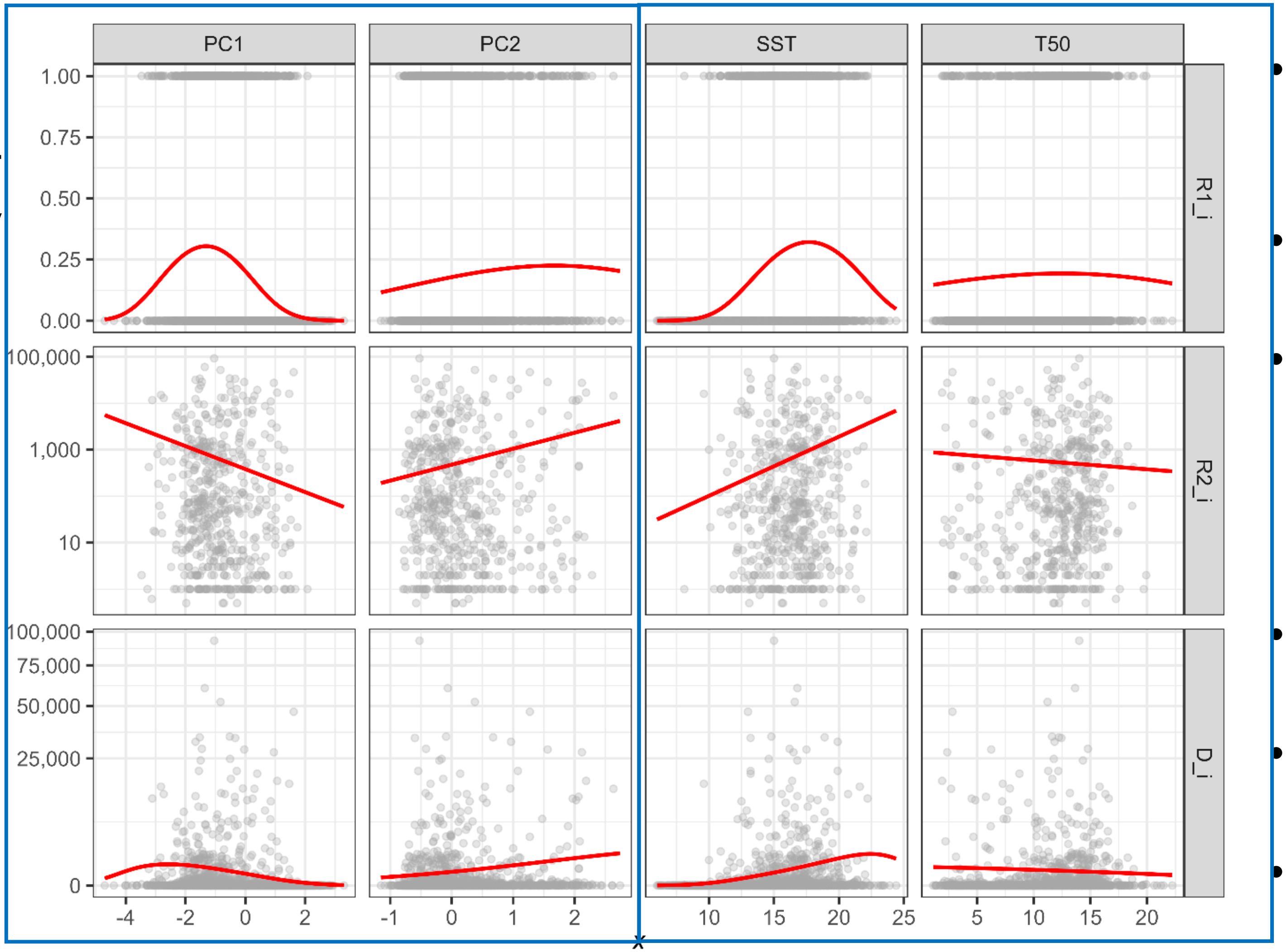
Relationships between covariates and CPUE

Fig. 9: Partial dependence plots

Encounter probability

Positive CPUE

Overall CPUE



- Concave-down responses of encounter probability to PC1 and PC2
- Linear relationship of positive CPUE to PC1 and PC2
- Assuming that the original variables SST and T50 change "independently," the responses to changes in each variable were examined
- SST had a greater influence than T50.
- The probability of positive catch peaked around 17.5° C for SST,
- The overall CPUE is highest at temperatures exceeding 20° C.

Yearly trends of nominal and standardized CPUE

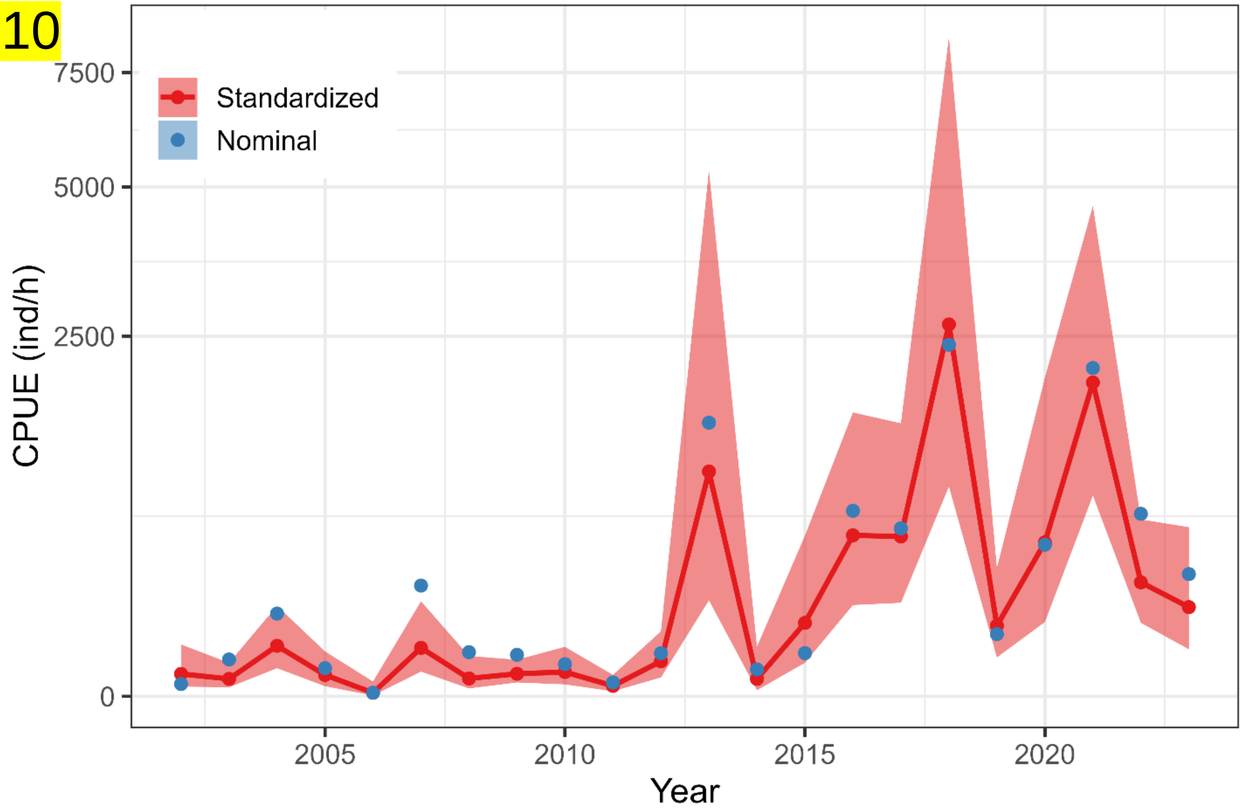
Average density (CPUE)

$$I(t) = \frac{\sum_{s=1}^{n_s}\left(a(s) \times d(s,t)\right)}{\sum_{s=1}^{n_s} a(s)}$$

area density

$$\frac{\text{Abundance}}{\text{Total area}}$$

Fig. 10



- Standardized CPUE remained low until 2012, but high values were frequently observed since 2013
- Especially in 2013, 2018, and 2021, the values were the highest, but compared to those, the values for the past two years (2022-2023) are not as elevated
- The yearly trend of the standardized CPUE was not greatly different from that of the nominal CPUE

Values and uncertainties of the nominal and standardized CPUE

Table 6

Year	Nominal (ind/h)	Standardized (ind/h)	CV	Lower 95%CI	Upper 95%CI
2002	2.94	9.75	0.39	1.83	51.94
2003	26.22	5.88	0.35	1.59	21.72
2004	132.07	49.22	0.33	15.34	157.98
2005	15.31	8.85	0.37	2.01	39.00
2006	0.24	0.25	0.62	0.01	4.26
2007	236.63	45.35	0.34	11.81	174.17
2008	37.65	6.20	0.40	1.22	31.43
2009	33.33	9.85	0.28	3.75	25.90
2010	19.97	11.46	0.36	2.78	47.30
2011	3.75	2.12	0.36	0.50	8.90
2012	35.95	23.76	0.32	6.93	81.49
2013	1443.45	974.09	0.43	177.92	5333.00
2014	14.03	5.89	0.49	0.73	47.43
2015	36.02	104.11	0.38	21.78	497.80
2016	663.42	499.73	0.30	160.72	1553.78
2017	543.68	492.72	0.31	168.96	1436.87
2018	2382.26	2665.93	0.32	848.68	8374.38
2019	74.62	96.33	0.32	29.20	317.73
2020	443.27	456.79	0.36	106.64	1956.65
2021	2077.32	1898.33	0.25	777.15	4637.02
2022	642.11	250.71	0.24	104.26	602.90
2023	288.17	153.54	0.35	42.70	552.14

- The coefficient of variation (CV) of the standardized CPUE was in the range of 0.24−0.49 for almost all years
- In 2006, when the standardized CPUE was the lowest, CV was highest (0.62)

Recommendation

- The standardized index obtained from this analysis cover a long time series from periods of poor chub mackerel recruitment in the Pacific to times of high recruitment
- The surveys covered a broad area in the Northwestern Pacific Ocean
- The cutting-edge VAST model was used for CPUE standardization
- Model diagnostics showed favorable results
- Propose utilizing the standardized index from the summer survey as an abundance index of recruitment (the numbers of age 0 fish) for the chub mackerel stock assessment in TWG CMSA