NPFC-2024-TWG CMSA08-WP04 (Rev. 1)

**CPUE Standardization Protocol for Chub Mackerel**

(revised in September 2023 at TWG CMSA07, document template added in November 2023)

CPUE is catch per unit effort obtained either from fishery independent or fishery dependent data.

The use of CPUE in a stock assessment implicitly assumes that CPUE is proportional to stock abundance/biomass. However, many factors other than stock abundance/biomass may influence CPUE. Thus, any other factors, other than stock abundance/biomass, that may influence CPUE should be removed from the CPUE index. The process of reducing/removing the impacts of these factors on CPUE is referred to as CPUE standardization.

The following protocol is developed for the CPUE standardization:

1. Provide a description of the type of data (logbook, observer, survey, etc. ), and the "resolution" of the data (aggregated, set-by-set etc..). This description should also include the representativeness of the data in two tables: (1st table) Number of observations, % Coverage of CPUE fleet (catch), % Coverage of CPUE fleet (effort), Total Catch CPUE fleet (mt), Total Effort CPUE fleet, Percentage of overall catch by member (across all fleets/gears); and (2nd table) Number of records remaining, Number removed, Number of records with chub mackerel catch >0.
2. Conduct a thorough literature review to identify potential explanatory variables (i.e., spatial, temporal, environmental, and fisheries variables) that may influence CPUE values.
3. Plot annual/monthly spatial catch, effort and nominal CPUE distributions and determine temporal and spatial resolution for CPUE standardization.
4. Make scatter plots (for continuous variables) and/or box plots (for categorical variables) and present correlation matrix if possible to evaluate correlations between each pair of those variables.
5. Describe selected explanatory variables based on (2)-(4) to develop full model for the CPUE standardization.
6. Specify model type and software (packages) and fit the data to the assumed statistical models (i.e., GLM, GAM, Delta-lognormal GLM, Neural Networks, Regression Trees, Habitat based models, and Statistical habitat based models).
7. Evaluate and select the best model(s) using methods such as likelihood ratio test, information criterions, cross validation etc.
8. Provide diagnostic plots to support the chosen model is appropriate and assumption are met (QQ plot and residual plots along with predicted values and important explanatory variables, etc.).
9. Present estimated values of parameters and uncertainty in the parameters in table.
10. Present relationship between dependent variable and independent variables. Check whether it is interpretable.
11. Extract yearly standardized CPUE and standard error by a method that is able to account for spatial heterogeneity of effort, such as least squares mean or expanded grid. If the model includes area and the size of spatial strata differs or the model includes interactions between time and area, then standardized CPUE should be calculated with area weighting for each time step. Model with interactions between area and season or month requires careful consideration on a case by case basis. Provide details on how the CPUE index was extracted.
12. Calculate uncertainty (SD, CV, CI) for standardized CPUE for each year. Provide detailed explanation on how the uncertainty was calculated.
13. Provide a table and a plot of nominal and standardized CPUEs over time. When the trends between nominal and standardized CPUE are largely different, explain the reasons (e.g. spatial shift of fishing efforts), whenever possible.

DOCUMENT TEMPLATE FOR PRESENTING STANDARDIZED CPUE OF CHUB MACKEREL

**Title: Standardized CPUE of chub mackerel (*Scomber japonicus*) caught by the MEMBER’s fishery/survey up to 20XX**

Author’s Name(s)

Affiliation(s)

**1. Background of the chub mackerel fishery**

* Description of the chub mackerel fishery of corresponding member.
* Identify potential explanatory variables that may influence CPUE values by conducting a thorough literature review

**2. METHOD**

*2.1 The data*

* Description of the type of data and the "resolution" of the data
* Catch and effort information, including the proportion of chub mackerel within the observed catch, with the representativeness of the data by CPUE FLEET (Table 1).
* Filter "rules" used on data for CPUE standardization and the effect on the overall sample size (Table 2)
* Annual/monthly spatial distribution patterns of catch, effort and nominal CPUE (Fig. 1)
* Provide a summary of covariates considered in the CPUE standardization (Table 3). Include details on the data source, resolution and coverage of the data over time and space. Provide justification supporting the relationship between the covariate and catchability that is assumed by including the covariate in the CPUE standardization.
* Scatter plots (for continuous variables) and/or box plots (for categorical variables) (Fig. 2) and, if possible, a correlation matrix to evaluate correlations between each pair of those variables (Fig. 3)

*2.2 Full model description and model selection*

* Model type and assumptions of the statistical model, assumed error distribution, model selection method, and software (packages) used
* Formulation of full model

*2.3 Yearly trend extraction*

* The method for extracting yearly standardized CPUE and its uncertainty such as standard error. The method can account for spatial heterogeneity of effort, such as least squares mean or expanded grid. If the model includes area and the size of the spatial strata are different, or if the model includes interactions between time and area, then standardized CPUE with area weighting should be calculated for each time step. Models with interactions between area and season or month require careful consideration on a case-by-case basis.

**3. RESULT and 4. DISCUSSION**

* Result of the model selection, at least for the full, null and best models (Table 4[[1]](#footnote-2)) with interpretation of the selected model
* Model diagnosis: Analysis of deviance table (Table 5), tendencies of the residuals and Q-Q plot etc. (Fig. 4) and percentage of the deviance explained if available.
* Estimated relationship between response and explanatory variables (Fig. 5\*). Check whether it is interpretable.
* Estimated values of parameters and uncertainty in the parameters (Table 6)
* The extracted yearly trend with its uncertainty (Table 7), comparing with the nominal CPUE (Fig. 6). If the trends between nominal and standardized CPUE are significantly different, explain the reasons whenever possible.
* Further discussion

**5. REFERENCES**

**APPENDICES**

* **Appendix I:** Checklist for the CPUE standardization protocol
* Further information in forms of description, figures, or table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table 1.** Catch and effort information by CPUE FLEET | | | | |  |  |
| Year | Number of observations | % Coverage of CPUE FLEET(catch ). | % Coverage of CPUE FLEET(effort ). | Total Catch CPUE FLEET (MT) | Total Effort for CPUE FLEET and unit | Percentage of overall catch by member (across all fleets/gears) |
| 2000 | 10000 | 100% | 100% | 99 | 100 vessel days | 10% |
| 2001 | 9000 | 90% | 90% | 120 | 101 vessel days | 9% |
| 2002 | 12500 | 80% | 85% | 110 | 102 vessel days | 11% |
| 2003 | 11000 | 89% | 89% | 229 | 103 vessel days | 1% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 2.** Filter "Rules" used on data for CPUE standardization and the effect on the overall sample size. | | | | | |
| Filter Applied |  | Number of Records Remaining | Number Removed |  | Number of Records With Chub Mackerel Catch >0 |
| Initial Data set |  | 12,345,678 | - |  | 10,000,000 |
| Remove Records outside convention area |  | 12,000,000 | 345,678 |  | 9,000,000 |
| Remove records <2°C & >28°C |  | 11,500,000 | 500,000 |  | 8,500,000 |
| Remove Records with <30% Chub Mackerel |  | 10,500,000 | 1,000,000 |  | 7,500,000 |
| Final Data Set |  | 10,500,000 | 1,000,000 |  | 7,500,000 |

**Table 3.** Summary of explanatory variables used in GLM\*.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variables | | Number of categories | Detail | Note |
| Year | *Year* | 25 | 1994–2018 |  |
| Month | *Month* | 5 | August–December |  |
| Fishing area | *Area* | 5 | I–V | *see* Fig. 1 |
| Vessel size | *Grt1* | 10 | *Grt*＜20, 20≦*Grt*＜40, ..., 180≦*Grt*＜200 tons | at intervals of 20 tons |
| *Grt2* | 5 | *Grt*＜40, 40≦*Grt*＜80, ..., 160≦*Grt*＜200 tons | at intervals of 40 tons |
| Sea surface temperature | *Sst1* | 12 | *Sst*＜10, 10≦*Sst*＜11, ..., 20 oC≦*Sst* | at intervals of 1oC |
| *Sst2* | 5 | *Sst*＜10, 10≦*Sst*＜13, ..., 19 oC≦*Sst* | at intervals of 3oC |

*\*All of the tables and figures in this document template are presented as an example.*

**Table 4.** Result of model selection\*

|  |  |  |  |
| --- | --- | --- | --- |
| No |  | Deviance explained  % | BIC |
| 1 |  | 13.9 | 211201 |
| 2 |  | 15.1 | 210828 |
| 3 |  | 18.7 | 210330 |
| 4 |  | 23.0 | **209391** |
| 5 |  | 19.3 | 209958 |

*β*0 – intercept, – coefficient of *i*-th year (*yeari*), – coefficient of *i*-th month (*monthi*), – coefficient of *i*-th unique ID of a vessel (*Idvesi*).



**Table 5.** Analysis of deviance table\*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | SS | Df | F | Pr(>F) | Signif. codes |
| Year | 453.8 | 24 | 39.35 | < 2.2e-16 | \*\*\* |
| Month | 117.3 | 1 | 244.06 | < 2.2e-16 | \*\*\* |
| Grt1 | 265.9 | 7 | 79.03 | < 2.2e-16 | \*\*\* |
| Sst2 | 51.1 | 4 | 26.60 | < 2.2e-16 | \*\*\* |
| Year:Month.int | 1067.4 | 72 | 30.85 | < 2.2e-16 | \*\*\* |
| Year:Area.int | 296.3 | 48 | 12.85 | < 2.2e-16 | \*\*\* |
| Year:Grt.int | 258.6 | 48 | 11.21 | < 2.2e-16 | \*\*\* |
| Month.int:Area.int | 45.4 | 6 | 15.734 | < 2.2e-16 | \*\*\* |
| Month.int:Grt.int | 33.5 | 6 | 11.624 | 4.74E-13 | \*\*\* |
| Area.int:Grt.int | 39.4 | 6 | 13.651 | 1.50E-15 | \*\*\* |
| Residuals | 19277.7 | 40113 |  |  |  |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | | | | |

**Table 6.** The estimated coefficients in the best models for the standardization of autumn survey CPUE (NOTE: not only estimated values but also SE of the estimates should also be shown ideally)

**グラフィカル ユーザー インターフェイス, テーブル

自動的に生成された説明**

**Table 7.** Nominal and standardized CPUEs of CPUE FLEET from 1994 to 2018.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | Nominal CPUE (metric ton / hauls) | Standardized CPUE | CV (%) | 95% CI | |
| Lower | Upper |
| 1994 | 5.38 | 2.93 | 3.53 | 2.74 | 3.14 |
| 1995 | 4.41 | 2.16 | 6.53 | 1.90 | 2.44 |
| 1996 | 2.40 | 1.62 | 4.69 | 1.48 | 1.77 |
| 1997 | 4.77 | 3.58 | 12.93 | 2.79 | 4.63 |
| 1998 | 1.44 | 1.02 | 3.86 | 0.94 | 1.09 |
| 1999 | 1.45 | 0.75 | 3.97 | 0.70 | 0.81 |
| 2000 | 2.18 | 1.37 | 4.38 | 1.26 | 1.49 |
| 2001 | 3.18 | 2.06 | 5.64 | 1.84 | 2.32 |
| 2002 | 1.93 | 1.15 | 5.66 | 1.02 | 1.29 |
| 2003 | 3.21 | 2.17 | 4.27 | 2.01 | 2.37 |
| 2004 | 3.65 | 2.51 | 3.95 | 2.33 | 2.71 |
| 2005 | 6.63 | 4.38 | 4.05 | 4.03 | 4.72 |
| 2006 | 6.03 | 3.93 | 4.30 | 3.61 | 4.28 |
| 2007 | 7.81 | 4.05 | 4.31 | 3.73 | 4.40 |
| 2008 | 7.81 | 4.93 | 4.06 | 4.56 | 5.31 |
| 2009 | 4.60 | 3.58 | 4.43 | 3.29 | 3.92 |
| 2010 | 2.73 | 1.49 | 3.66 | 1.37 | 1.59 |
| 2011 | 4.45 | 2.36 | 4.01 | 2.19 | 2.55 |
| 2012 | 3.65 | 2.31 | 4.31 | 2.12 | 2.52 |
| 2013 | 3.04 | 1.43 | 3.88 | 1.33 | 1.55 |
| 2014 | 5.42 | 2.49 | 3.64 | 2.32 | 2.67 |
| 2015 | 2.65 | 1.34 | 4.43 | 1.23 | 1.46 |
| 2016 | 2.82 | 1.50 | 5.94 | 1.33 | 1.68 |
| 2017 | 1.40 | 1.08 | 4.23 | 1.00 | 1.17 |
| 2018 | 2.96 | 1.40 | 3.91 | 1.30 | 1.52 |

グラフィカル ユーザー インターフェイス, アプリケーション, Word

自動的に生成された説明

**Fig. 1a.** Spatio-temporal distribution of the total catch of CPUE FLEET (metric tons). Rectangles with bold black lines are the area stratification determined by delta-GLM-tree. (NOTE: This example figure shows effort distribution because there was no good example. In actual document, catch should be shown)

グラフィカル ユーザー インターフェイス, アプリケーション, Word

自動的に生成された説明

**Fig. 1b.** Spatio-temporal distribution of efforts by CPUE FLEET (hours of towing).

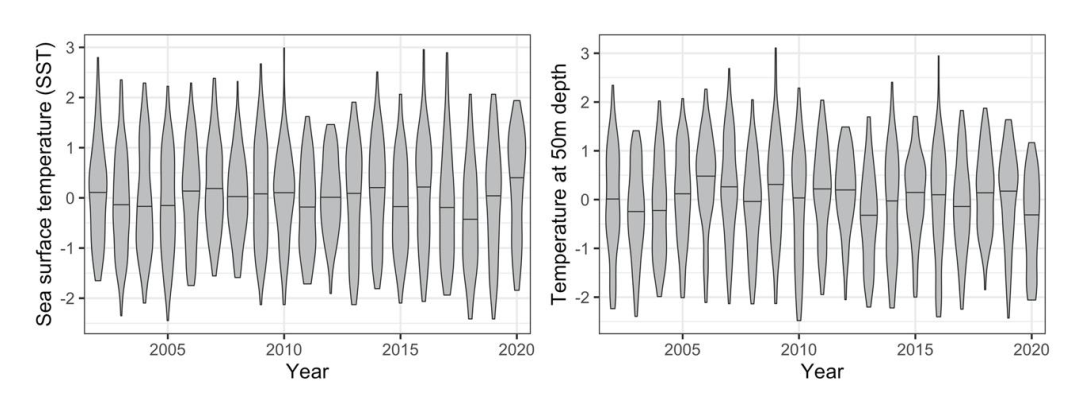
Rectangles with bold black lines are the area stratification determined by delta-GLM-tree.

グラフィカル ユーザー インターフェイス, アプリケーション, Word

自動的に生成された説明

**Fig. 1c.** Spatio-temporal distribution of CPUE of CPUE FLEET (number per hour). Rectangles with bold black lines are the area stratification determined by delta-GLM-tree.

(a)

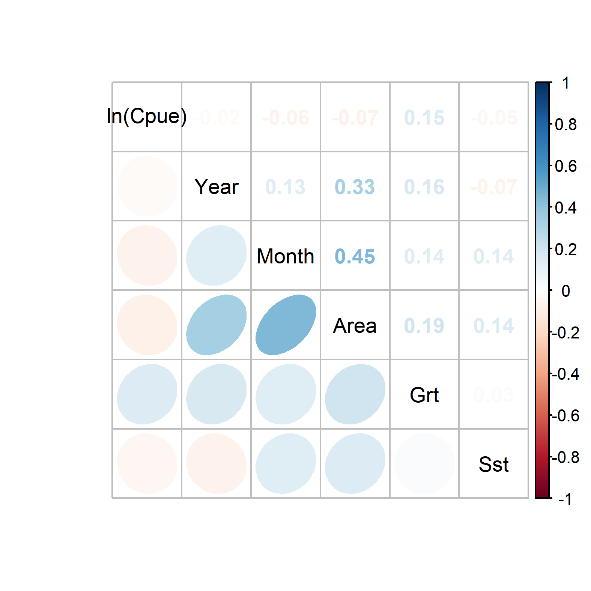


(b)

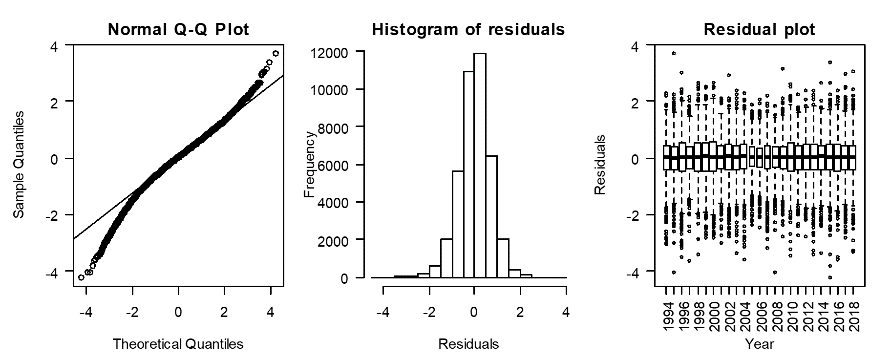
散布図 が含まれている画像

自動的に生成された説明

**Fig. 2.** Plots of explanatory variables of sea surface temperature (SST) and T50 by year (a) and scatter plots between CPUE and temperatures (b). (NOTE: How many graphs should be shown depends on explanatory variables considered.)

****

**Fig. 3.** Correlation matrix of explanatory variables used in the analysis.



**Fig. 4.** Q-Q plot, histogram of residuals and residual plots across years for the best GLM. (NOTE: The residual plots should be shown by important explanatory variables such as the right panel of Fig. 4.)

(A)

グラフ

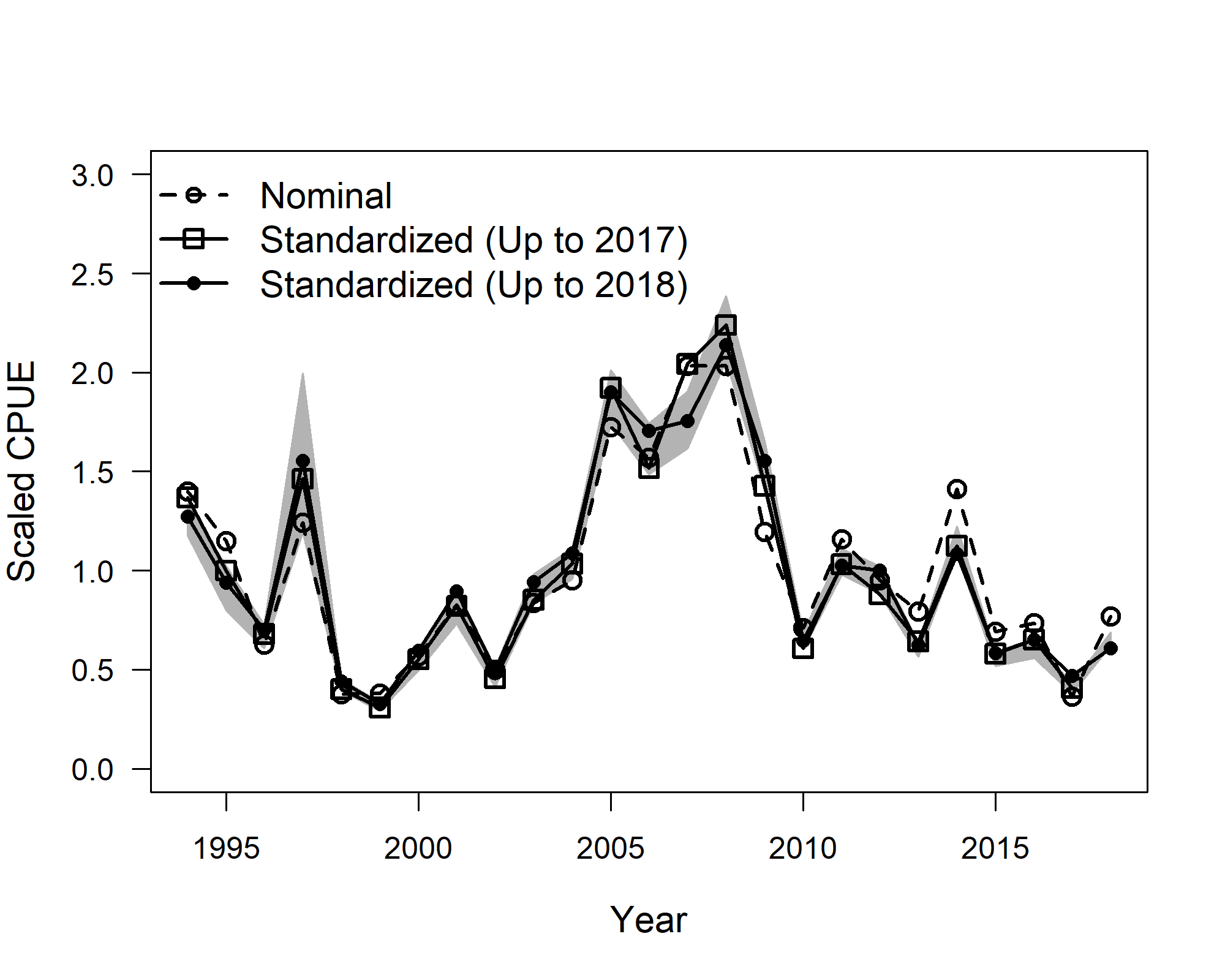
自動的に生成された説明

(B)

グラフ, 棒グラフ, ヒストグラム

自動的に生成された説明

**Fig. 5.** Estimated relationships between response and explanatory variables. (A) An example showing nonlinear relationships between response variables and SST. Points are observed values, and black lines are estimated relationship. (B) An example when the explanatory variable is categorical.



**Fig. 6.** A scaled nominal CPUE series and two scaled standardized CPUE series with catch and effort data up to 2017 and 2018. Gray zone indicates 95% confidence band for the standardized CPUE up to 2018.

**AppendICES**

**Appendix I.** Checklist for the CPUE standardization protocol

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Step-by-step protocols | yes/no | Note |
| 1 | Provide a description of the type of data (logbook, observer, survey, etc. ), and the "resolution" of the data (aggregated, set-by-set etc..). This description should also include the representativeness of the data in two tables: (1st table) Number of observations, % Coverage of CPUE fleet (catch), % Coverage of CPUE fleet (effort), Total Catch CPUE fleet (mt), Total Effort CPUE fleet, Percentage of overall catch by member (across all fleets/gears); and (2nd table) Number of records remaining, Number removed, Number of records with chub mackerel catch >0; |  | See *2.1* ([INSERT page number(s)]) and Tables 1, [INSERT page number(s)] and 2, [INSERT page number(s)] |
| 2 | Conduct a thorough literature review to identify potential explanatory variables (i.e., spatial, temporal, environmental, and fisheries variables) that may influence CPUE values; |  | See 1. Background ([INSERT page number(s)]) |
| 3 | Plot annual/monthly spatial catch, effort and nominal CPUE distributions and determine temporal and spatial resolution for CPUE standardization; |  | See Fig. 1, [INSERT page number(s)] |
| 4 | Make scatter plots (for continuous variables) and/or box plots (for categorical variables) and present correlation matrix if possible to evaluate correlations between each pair of those variables; |  | See Figs 2, [INSERT page number(s)] and 3 [INSERT page number(s)] |
| 5 | Describe selected explanatory variables based on (2)-(4) to develop full model for the CPUE standardization; |  | See *2.2.* ([INSERT page number(s)]) and Table 3, [INSERT page number(s)] |
| 6 | Specify model type and software (packages) and fit the data to the assumed statistical models (i.e., GLM, GAM, Delta-lognormal GLM, Neural Networks, Regression Trees, Habitat based models, and Statistical habitat based models); |  | See *2.2.* ([INSERT page number(s)]) |
| 7 | Evaluate and select the best model(s) using methods such as likelihood ratio test, information criterions, cross validation etc.; |  | See Table 4, [INSERT page number(s)] and related descriptions in the main text |
| 8 | Provide diagnostic plots to support the chosen model is appropriate and assumption are met (QQ plot and residual plots along with predicted values and important explanatory variables, etc.); |  | See Table 5, [INSERT page number(s)] and Fig. 4, [INSERT page number(s)] |
| 9 | Present estimated values of parameters and uncertainty in the parameters in table; |  | See Table 6, [INSERT page number(s)] |
| 10 | Present relationship between dependent variable and independent variables. Check whether it is interpretable. |  | See Fig. 5, [INSERT page number(s)] |
| 11 | Extract yearly standardized CPUE and standard error by a method that is able to account for spatial heterogeneity of effort, such as least squares mean or expanded grid. If the model includes area and the size of spatial strata differs or the model includes interactions between time and area, then standardized CPUE should be calculated with area weighting for each time step. Model with interactions between area and season or month requires careful consideration on a case by case basis. Provide details on how the CPUE index was extracted. |  | See 2.3. ([INSERT page number(s)]) |
| 12 | Calculate uncertainty (SD, CV, CI) for standardized CPUE for each year. Provide detailed explanation on how the uncertainty was calculated; |  | See Table 7, [INSRT page number(s)] and Fig. 6, [INSERT page number(s)] |
| 13 | Provide a table and a plot of nominal and standardized CPUEs over time. When the trends between nominal and standardized CPUE are largely different, explain the reasons (e.g. spatial shift of fishing efforts), whenever possible. |  |

1. Members must present at least one of Table 4 and Table 5, but it is not obligatory to present both. [↑](#footnote-ref-2)