CCSBT-OMMP/2306/05
(OMMP Agenda item 1.3)

Update of CPUE abundance index using GAM for southern bluefin tuna in CCSBT up to the 2022 data

CCSBT のミナミマグロについての GAM を用いた CPUE 資源豊度指数の 2022 年データまでの更新

Tomoyuki ITOH and Norio TAKAHASHI
伊藤智幸・高橋紀夫
Fisheries Resources Institute, Japan Fisheries Research Agency
水産研究･教育機構 水産資源研究所

要約

2022 年 ESC27 では、2 段階のデルタログノーマルアプローチで GAM で CPUE を標準化して、面積重みづけする新たなミナミマグロの CPUE 資源量指数の作成方法に合意した。合意された方法に従い、CPUE 資源量指数の作成を 2022 年までの漁獲データに対して実施した。本文書ではベースケースの結果と、様々な感度分析を行った結果を示す。得られた 2022 年の指数値は前年から大きく上昇し、シリーズの最高値となった。モデル選択、レトロスペクティブ解析、船の ID、海域範囲の変更、対象年齢の変更、データおよびモデルの分解能の変更を含む様々な感度分析に対して資源量指数は頑健であった。

Summary

At ESC27 in 2022, the new calculation method of the abundance index of southern bluefin tuna, which standardized by generalized additive model in the two-step delta log normal approach with area weighting, was agreed. The CPUE abundance index was updated for fishery data up to 2022 according to the agreed methodology. This document presents the base case results as well as the results of various sensitivity tests. The index value for 2022 was a significant increase from the previous year and was the highest value in the series. The abundance index was robust for a variety of sensitivity analyses, including model selection, retrospective analysis, vessel ID, area range changes, age range changes, and data and model resolution changes.

1. Introduction

Stock assessment and stock management through the Management Procedure (MP) of southern bluefin tuna (*Thunnus maccoyii*; SBT) in CCSBT have historically been strongly relied on the abundance index obtained from the CPUE (number of fish / 1000 hooks) of the Japanese commercial longline fishery. In the old days, Nishida and Tsuji (1998) developed a model to calculate the abundance index by the generalized linear model (GLM). Since 2007, alternative abundance index was developed which called the core vessel CPUE standardized by GLM in response to the shrinking operating area in time and space and the problem of target fish species (ESC12 report, Itoh et al. 2008). The CPUE abundance index had been used as one of the main abundance indices in the two MPs of the Bali procedure used for the TAC calculation from 2012 to 2020 and the Cape Town Procedure (CTP) used for the TAC calculation from 2021.

It was recognized that the 2018 value of the CPUE abundance index by the core vessel CPUE was anomalously high in ESC24 held in 2019 (ESC24 Report). This prompted further investigation, which subsequently identified that this estimate was generated due to a prediction bias in the GLM standardisation method being used, which generated anomalously high estimates for cells with no effort. At ESC26 in 2021, it was agreed that a new CPUE abundance index should be prepared by May 2022 to assess its impact on MP (ESC26 report). Through the collaboration work between Japanese scientists and the consultant hired by CCSBT, as well as the discussion and suggestion of the CPUE working group, a new abundance index using CPUE standardised by the generalized additive model (GAM) was developed and agreed at ESC27 in 2022 (OMMP12 and ESC27 Reports).

This document presents the CPUE results obtained by updating the data to 2022 using the agreed GAM methodology not only for the base case but also for the various sensitivity analyses. Although the ESC has already agreed to use the GAM new series in 2022, the CPUE abundance index may change with additional updates of the data over a one-year period, we need to carefully consider whether we can lead the same conclusion again. We have done all the analyses which had done last year (Itoh and Takahashi 2022).

2. Materials and Methods

2-1. Dataset used

The dataset was extracted from logbook data for the Japanese longline fishery, which include the period from 1969 to the latest year (currently 2022). Following procedures for the conventional SBT CPUE abundance index, records in Statistical Areas between 4 and 9 and from April to September were selected. From the logbooks, year, month, latitude (in 1 degree), longitude (in 1 degree), vessel ID (available from 1979), number of hooks used, number of fish caught of SBT, bigeye tuna (*T. obesus*, BET), yellowfin tuna (*T. albacares*, YFT), albacore (*T. alalunga*, ALB) and swordfish (*Xiphias gladius*, SWO) data were used. In the development work in 2022, the number of hooks between floats (HBF; available since 1975) and other fish species (several species of marlines, and butterfly kingfish (*Gasterochisma melampus*; available since 1994)) were examined and decided not to use them so that these items were not included this year.

From the size data of the CCSBT database, the age composition of Japanese commercial catch was calculated and converted into the number of fish caught age-4 and older (age-4 plus). The age composition information was first applied to the fork length composition of 50 or more individuals measured in the same month, 5 degrees longitude, and 5 degrees latitude. At this stage, 97% of the number of SBT caught was incorporated and the ratio of age-4 plus was calculated. For records of the conditions for 50 or more

individuals were not met the time and space were gradually expanded to correspond to fork length composition, such as the same month - longitude 15 degrees - latitude 5 degrees, the same month - longitude 15 degrees - latitude 15 degrees, the same quarter - longitude 15 degrees - latitude 5 degrees, the same quarter - Statistical Area, and the same year - Statistical Area, and the same year. The fork length was converted to age by the age-length relationship used by CCSBT. Sensitivity analysis was conducted for age-5 plus and all ages.

The following records were eliminated: hooks 500 or less, hooks 4500 or more, CPUE 200 or higher. As a result of the examination, with the agreement in the CPUE working group discussion, the record of 50S (50S to 54S), which had a small number of data, were also eliminated.

2-2. Cluster analysis

A cluster analysis was performed to consider the target species of the fishing operations. The *clust_PCA_run* function of the R package *cpue.rfmo* was used. Cluster analysis was performed using the number of fish caught of five species, SBT, BET, YFT, ALB and SWO as data.

2-3. Standardization by GAM

Standardization by the generalized additive model (GAM) was carried out by use of the delta log normal approach. R was used for analysis. The *bum* function, which is suitable for large volumes of data, in the *mgcv* package was used. Based on the results of the study by the consultant, a binomial submodel (hereinafter referred to as BSM) and a positive catch submodel (hereinafter referred to as PCSM) were used, and gamma = 2, binomial distribution and gauss distribution were used respectively (Hoyle 2022). For the smoother, s (spline) was used for the offset term (hook logarithmic value), and ti (tenor product suitable when there was an interaction with the main effect) was used for the others. cs (cubic regression spline with shrinkage) was used for the basis function (bs) of ti. Gamma is a coefficient multiplied by EDF (described later) and promotes smoothing with values set to >1 (= 1.5 is common).

Binomial submodel

```
cpue > 0 ~ yf +ti(month) + ti(lon) + ti(lat) +
    ti(lon,  lat) + ti(month, lat) + ti(lon, month) + ti(year, lat) + ti(year, lon) + ti(year, month) +
    cl + s(log(hook))
```

Positive catch submodel

```
log(cpue) ~ yf +ti(month) + ti(lon) + ti(lat) +
    ti(lon,  lat) + ti(month, lat) + ti(lon, month) + ti(year, lat) + ti(year, lon) + ti(year, month) +
    ti(lat, month, year) + ti(lat,  lon, month) + ti(lat, lon, year) + ti(year, lon, month) +
    cl + s(log(hook))
```

where,

yf: year. In factor.

year: year. In number

month: month. In number

lat: Latitude in 5 degree. In number. Represented by the middle (e.g. -47.5 from 45.0S to 49.9S)

lon: Longitude in 5 degrees. In number. Represented by the middle (e.g. 32.5 for 30.0E to 34.9E). Convert to 360 degree while >240 was converted by -360 so that lon ranged from -22.5 to 187.5 continuously.

cl: Cluster. In factor. 1, 2, 3, and 4.

hook: Number of hooks used. In number.

R code actually used is as follows.

Binomial submodel

```
modA2 <- cpue > 0 ~ yf +
                    ti(month,       k=kA.month11,bs="cs")+
                    ti(lon,         k=kA.lon11,bs="cs")+
                    ti(lat,         k=kA.lat11,bs="cs")+
                    ti(lon,  lat,   k=c(kA.lon21,  kA.lat21), bs="cs")+
                    ti(month, lat,  k=c(kA.month22,kA.lat22), bs="cs")+
                    ti(lon, month,  k=c(kA.lon23,  kA.month23), bs="cs")+
                    ti(year, lat,   k=c(kA.year24, kA.lat24), bs="cs")+
                    ti(year, lon,   k=c(kA.year25, kA.lon25), bs="cs")+
                    ti(year, month, k=c(kA.year26, kA.month26), bs="cs")+
                    cl+
                    s(log(hook))
mgcv::bam(modA2, data =data, gamma = 2, method = 'fREML', family = binomial, discrete=F)
```

Positive catch submodel

```
modB3 <- log(cpue) ~ yf +
                    ti(month,           k=kB.month11,bs="cs")+
                    ti(lon,             k=kB.lon11,bs="cs")+
                    ti(lat,             k=kB.lat11,bs="cs")+
                    ti(lon,  lat,       k=c(kB.lon21,  kB.lat21), bs="cs")+
                    ti(month,lat,       k=c(kB.month22,kB.lat22), bs="cs")+
                    ti(lon,  month,     k=c(kB.lon23,  kB.month23), bs="cs")+
                    ti(year, lat,       k=c(kB.year24, kB.lat24), bs="cs")+
                    ti(year, lon,       k=c(kB.year25, kB.lon25), bs="cs")+
                    ti(year, month,     k=c(kB.year26, kB.month26), bs="cs")+
                    ti(lat,  month,year,k=c(kB.lat31, kB.month31, kB.year31), bs="cs")+
                    ti(lat,  lon, month,k=c(kB.lat32, kB.lon32, kB.month32), bs="cs")+
                    ti(lat,  lon, year, k=c(kB.lat33, kB.lon33, kB.year33), bs="cs")+
                    ti(year, lon, month,k=c(kB.year34, kB.lon34, kB.month34), bs="cs")+
                    cl+
                    s(log(hook))
```

mgcv::bam(modB3, data = data.positive, gamma = 2, method ="fREML", discrete=F)

The larger the k value (basis dimension for smoothing flexibility) of the interaction, the better, but the longer the calculation time (Wood, help of *choose.k* in mgcv). The effective degrees of freedom for a model term (EDF) value is calculated by the *k.check* function in mgcv package, and if EDF was close to k' (the maximum possible EDF for the term), "and" the p-value of k-index is < 0.05, a larger k value was set. The k values were determined by trial and error. Since the k value of the interaction is treated as the value of 2 multiplications (3 multiplications for 3 interactions), it is not necessary to set them separately, however, for the purpose of organizing the work, the k value of each variable in the interaction was set to the same value (i.e. k for year = 20 was used for all interaction terms which include year).

For the diagnosis of the GAM result, the fit was confirmed by the plot diagram (QQ plot, residual distribution) by the *gam.check* function of the mgcv package. AIC is calculated. The distribution of the residuals for each variable was examined. It was examined whether the predicted values were consistent with our knowledge of distribution of SBT and plausible trend of SBT stock abundance. We made a comprehensive judgment by looking at these information as well as AIC.

Calculation is performed by desktop PC (CPU = Intel (R) Core (TM) i9-10900T CPU @ 1.90GHz and 1.90 GHz, RAM = 64.0GB, 64 bit, Windows 10 Pro). The software R (R4.1.2) was used to make the dataset. Microsoft R Open 4.0.2 was used to calculate GAM.

2-4. Calculation of abundance index

After creating data with all combinations of year / month / latitude / longitude (using R's *expand.grid* function), we made a dummy data set limited to the month / latitude / longitude where the fishing was operated in the past. The predicted value was calculated for each submodel for the dummy data set, and the product was calculated. Since the expected value is biased when the lognormal distribution is restored, the predicted value was corrected by adding mean squared error (MSE) / 2 in the case of the positive catch submodel.

Furthermore, the area weighting coefficient was calculated in consideration of the fact that the distance of 1 degree of longitude differs depending on the latitude and the number of 1 degree squares that SBT have been caught in the past within the 5 degree x 5 degree squares. The abundance index can be calculated by the following formula.

Σ(predicted value of binomial submodel of dummy data set × predicted value of positive catch submodel of dummy data set × Area weighting coefficient) / Overall average value.

2-5. Sensitivity analysis

Various sensitivity analyses were performed along the way in selecting the datasets and methods. The following sensitivity analyses were performed at the final stage in 2022. We repeated the same procedure in 2023.

Model selection: In some cases, estimation did not converge, and in some cases, even if the AIC was low, the abundance index behaved significantly differently from the others, so a simple selection by AIC seemed inappropriate. For the binomial submodel, we tried the case where all the interactions were removed from the base case, the case where the two-way interaction was removed one by one, and the case where the three-way interaction was added one by one. For the positive catch submodel, we tried the case where all the interactions were removed from the base case, the case where the two-way interaction was removed one by one, and the case where the three-way interaction was removed one by one.

Retrospective analysis: Excludes data from the last year up to the past 10 years. It was also carried out in a part of the sensitivity analysis (i.e. latitude-longitude resolution in the model, core vessel).

Selection of k: Effect when k was increased by one step.

Effect including vessel ID: The effect of including vessel ID in each of BSM and PCSM.

Effect excluding 30S: We saw the effect excluding 30S from the dataset.

Effect of changing age: Age-4 plus used in the base case, but limited to age-5 plus, or all ages were tried.

Data resolution: The shot-by-shot data are used in the base case. Aggregated data by month, 1 degree latitude and 1 degree longitude, and aggregated data by month, 5 degrees latitude and 5 degrees longitude were tried.

Latitude-longitude resolution in the model: The latitude and longitude in the model is based on the 5-degree in the base case. We tried the effect when this was made into one degree resolution.

2 clusters: 4 clusters were the base case, but we tried clusters in 2 groups.

Core vessel: From the dataset prepared for this analysis (note that it differs from the traditional GLM dataset), we tried to select core vessels and create CPUE index. A core vessel is defined as a vessel that has been included in the top *xx* rank in terms of SBT catch in number of *yy* years.

3. Results

3-1. Dataset used

Data from 1969 to 2022 amounted to 797,416 records. Of these, 704,842 records included catch of SBT age-4 plus, accounting for 88% of the total. A fairly high positive catch rate is characteristic of this dataset. By year, the positive catch rate dropped to about 60% in the mid-1990s and around 2010, but otherwise remained above 80% (Fig. 1). The nominal CPUE of the positive catch dataset is high in the 1970s, low in the 1980s to 2000s, and high after 2010. The nominal CPUE in 2022 was the highest in the past 40 years.

Similar figures are shown for other variables (month, longitude, latitude, latitude and longitude maps) (Fig. 2 and Fig. 3). There is no strong tendency for the month and longitude. For latitude, positive rate and CPUE in the positive catch data was low at 30S, high up to 35S (CPUE) or 40S (positive rate), and 45S was similar to 40S. Data of 30S exists only in the Pacific Ocean (Area 4 and Area 5).

3-2. Cluster analysis

The cluster analysis was divided into four groups. Relevant figures are shown in Fig. 4 to Fig. 8. Since the eigenvalues are greatly reduced to 2 groups and the decrease to 4 groups is not so large, it may be appropriate to divide them into 2 groups. However, in the analysis of the data up to 2020, there was a problem that the BSM of GAM did not converge when divided into two groups (the data up to 2021 converged in a short time). Therefore, we decided to analyze in 4 groups. In addition, the case of 2 groups was carried out by sensitivity analysis.

The fish species included five species: SBT, BET, YFT, ALB and SWO. At the stage of trial and error in the 2022 work, we also tried 3 species (SBT, BET and YFT) and obtained the similar results as 5 species. But 3 species are few and cover all species that can be the main target of operation, it was decided by the CPUE working group to have 5 species (Itoh and Takahashi 2022).

The latitudes of the four clusters differed (Fig. 7), however, there were no noticeable trends in year,

month, longitude, number of hooks used, or hooks between floats (HBF). It was probable that HBF had a narrow range in the dataset and did not make a difference because it contained few data of deep longline targeting on BET. Such an effect may have been seen in the waters north of the Area 4-9. The main catch in the first cluster which locate southernmost was SBT. SBT and ALB were caught in the second cluster. The third cluster was a mixture of SBT, ALB and BET and the fourth cluster was a mixture of five species.

3-3. Standardization by GAM

For the binomial submodel, the model including all main effects and two-way interaction terms was selected mainly from AIC. There was a problem that the run did not converge when the three-way interaction term was included. For the positive catch submodel, the model including the main effect and all the two-way and three- way interaction terms was selected mainly from AIC.

The k value was examined independently for each submodel. The same sets used in the 2022 work were utilized (Table 1). Table 2 shows relevant statistics including the EDF value for k and the p value for k-index. There are cases where EDF is close to k' (e.g. positive catch submodel ti (lat)), but since the p value is well above 5% in that case, k is large enough and there is no problem.

The diagnosis results are shown in Table 3, Fig. 9, and Fig. 10. The binomial submodel explained 73.5% deviance, and the positive catch submodel explained 49.4%. For BSM, the QQ plot is generally good, although some parts do not fit at both ends. The residual histogram has a single peak and is skewed to near 0 residual. For PCSM, the QQ plot is generally good, and the residual has a single peak. In the plot of the fit value and the response variable, there is a roughly upward-sloping relationship. Both are judged to be not bad fit.

The residuals were further examined. Plots were made for year, month, latitude, and longitude (Figs. 11 and 12). Note that these figures are not from *gamVis*, which uses simulation. There was too much data and *gamVis* caused a memory over and couldn't get any results. These are simple box plots of residuals. For BSM, the median residuals were positively biased in 2004-2007 in the year. There was a slight positive bias for month. At latitude, the negative bias was large at 30S, a slight positive bias was seen at 35S, and the bias was small at 40S and 45S. At the western end of the longitude, there was a large negative bias.

For PCSM, the bias was small by year and month. At latitude, the range was large at 30S. The bias of the longitude was small, but a negative bias was seen only at the eastern end. When made into a map, the area with zero residuals was greatly expanded in both submodels (Fig. 13). In some places, large residuals may occur in the peripheral waters. It has been confirmed in the 2022 work that the data in the area where these large residuals are seen has almost no effect on the abundance index.

Box plots of predicted values for variables (year, month, latitude, longitude, latitude x longitude) are shown (Fig. 14, Fig. 15, Fig. 16, Fig. 17 and Fig. 18). No inconsistency was found in comparison with the current knowledge of the distribution of SBT and changes in the abundance. The high predicted values in the southeastern waters of Australia (35S, 140E) are interesting (Fig. 18). Currently, there is no fishing operation in this area, but it was confirmed in the data that the fishing was operated in this area in the 1970s and 1980s.

3-4. Calculation of abundance index

The predicted value of the dummy data set was weighted by the area factor and normalized by the average value to obtain the abundance index. To see the effect of area weighting, we compared it with a simple unweighted average (Fig. 19). As a result, it was found that they are similar to each other and the influence of weighting is small. Since this method includes the interaction of years in the model, it is no longer necessary to obtain the conventional Constant / Variable square hypothesis and its intermediate

index (see Hoyle (2022) for details).

Figure 20 shows the obtained abundance index. The values are shown in Table 4. It increased in many years from 2006 to 2012. In 2022, it increased largely from previous year and reached the highest value in the series.

3-5. Sensitivity analysis

Model selection

For BSM, a model (modA2) containing all two-way interactions was selected as the base case. Its AIC was lower than any other model with one term removed from modA2 (Table 5). On the contrary, in the model to which one three-way interaction term was added (e.g. modA2.p11), the AIC was low, but there was a problem that it did not converge. The effect on the abundance index was small in both models (Fig. 21 and Fig. 22). Therefore, it is considered appropriate to use modA2 as the base case.

For PCSM, a model (modB3) containing all the two-way and three-way interaction terms was selected as the base case. Its AIC was lower than any other model with one term removed from modB3 (Table 6). The difference between the models in the abundance index is small (Fig. 23 and Fig. 24). A relatively large difference was seen in modB3.no9 which excluding ti(year, lon).

Retrospective analysis

Figure 25 shows the results of retrospective analysis of the base case model. Figure 26 shows the results by each submodel. Looking at the figure, it was overestimated in four years (2015, 2018, 2019, 2020) compared to the results in the dataset up to 2022. On the contrary, it was underestimated in three years (2012, 2013 and 2021). It was not change in three years (2014, 2016, 2017). The impact before 2011 was small. There was no tendency to be over / under biased, suggesting that it is robust against data updates.

We examined further for the 2021 value. We compared the results of the data up to 2021 made in 2022 and that in 2023. That is, verification of the additional effect of data for 2021. As a result, no significant difference was observed between the two (Fig. 27).

Selection of k

For BSM, we examined the effect of adding +1 to k of the month, +5 to k of the year, and +5 to k of the longitude. The latitude has already reached the maximum value (k = 4). For PCSM, we examined the effect of increasing the year k by +5 and the longitude k by +5. The month and latitude are already at their maximum.

As a result, there was very little effect on BSM (Fig. 28 and Fig. 29). It is suggested that k was large enough. For PCSM, there was a noticeable change when k.year34 (ti (year, lon, month)) or k.year33 (ti (lat, lon, year)) was changed from 20 to 25 (Fig. 30, Fig. 31). It might be better to consider increasing the k-values associated with year in future.

Effect of including vessel ID

The vessel ID was included in the logbook data after 1979. The analysis was limited to the data in the years after that and records that the vessel ID was not missing. We tried both cases where the vessel ID was included as a fixed effect and where it was included as a random effect. Both cases took a long time to calculate. Whereas the base case took 30 minutes, the random term took more than 120 minutes, and the fixed effect took more than 155 minutes.

The results are shown in Fig. 32. The trajectories of the three abundance indices were similar to each other, which suggests the effect including the vessel ID was small. However, the behavior of recent year differs depending on the model and index with vessel ID had larger values in 2022. Due to the long run time, it was not possible to conduct a more detailed examination by trial and error.

Effect of excluding 30S

In the dataset used, 30S existed only in the Pacific Ocean in Statistical Area 4 and 5. As a test, data in 30S was eliminated from the dataset for the base case. The cluster analysis has not been redone. The abundance index is shown in Fig. 33. The abundance index is slightly lower in 2016-2018 and higher in 2021 and 2022. That in 2022 is significantly higher than the base case.

Effect of changing age range

The results are shown for the base case of age-4 plus, limited to age-5 plus (Fig. 34), and for all ages (Fig. 35). At the age-5 plus, there was no significant difference. For all ages, the values for 1990-1994 were high and those for 2010-2012 were low, but the overall trajectory was similar.

This sensitivity analysis is related to release and discard. When fish is released and discarded from longline vessels, it is often a small fish, age-3 or age-4. The proportion of released fish will depend not only on the vessel's IQ utilization strategy but also on the cohort strength. If the proportion of released fish changes in a certain year in the future, the effect can be examined by calculating the abundance index for those ages other than 4 and comparing it with the abundance index for those age-4 plus.

Resolution of data

The base case was obtained from shot-by-shot data. On the other hand, Fig. 36 shows the comparison with the aggregated data by month in 5 degrees x 5 degrees, and Fig. 37 shows the comparison with the aggregated data by month in 1 degree x 1 degree. The abundance index in the 5-degree aggregated data behaved differently from that in the shot-by-shot data after 2014. This is probably due to differences in the amount of data and effort. In this analysis, CPUE is taken as a response variable, and the number of hooks used is included as an offset term by taking a logarithm. For this reason, in the 5-degree aggregate data, even if there are multiple operations, they are treated as one record, and the effect of the number of operations (number of hooks) is treated only in logarithmic. Shot-by-shot data influences the results according to the number of operations. That is, aggregated data underestimates the variance in CPUE fluctuations. The abundance index of 1-degree aggregated data was an intermediate characteristic between shot-by-shot and 5-degree aggregated data.

The residuals were plotted against four variables (year, month, latitude, longitude) (Fig. 38). A significant improvement was seen in the homoscedasticity of the residuals by using shot-by-shot at latitude. From these facts, the resolution of the data has a strong influence on the result, and it is considered appropriate to use shot-by-shot data as agreed in 2022.

Resolution in the model in latitude and longitude

The latitude and longitude in the model uses a 5-degree resolution, but we tried the effect of using this as a 1-degree resolution. The runtime has then doubled. The abundance index was higher in the 1-degree model in 2010-2015 and 2022 (Fig. 39). When retrospective analysis was performed, the 1-degree model was more deviated than the base case (Fig. 40). By submodel, the bias of underestimation was remarkable in BSM (Fig. 41).

2 clusters

The base case used clusters in four groups, but clusters in two groups were tried. SBT was abundant in the first cluster, and five species were mixed in the second cluster. The first cluster was located to the south (Fig. 42, Fig. 43, Fig. 44, Fig. 45). The abundance index of two clusters was similar to that of four clusters, although larger in 2019 and lower in 2022 (Fig. 46). The 2022 work showed that the four clusters are considered more robust and suitable, as the dataset up to 2020 caused the problem of GAM convergence and there was no significant change in the abundance index.

Core vessel

The core vessel is that included in the top *xx* vessels in terms of SBT catch in number of a certain year and has been included for *yy* years. The conventional core vessel CPUE is selected with xx = 56 and yy = 3. The core vessel data set was obtained by setting xx and yy in various ways (Table 7). Data is limited to 1979 and later with vessel ID. The number of vessels selected has decreased from 3% to 17%. The abundance index is shown in Fig. 47. The behavior that deviated greatly was shown in the case of data in which the number of vessels was greatly compressed to <7%. The index value for 2022 was lower on the core vessel.

Retrospective analysis was performed for the cases of core vessel data xx = 56 and yy = 3 (Fig. 48, Fig. 49). It was roughly robust, but underestimated in 2012-2013 and overestimated in 2020, less robust than the all-vessel dataset. Since the data is limited after 1979 because the vessel ID is required, and the robustness is inferior, it is considered that the significance of using the core vessel is small.

3-6. Comparison of abundance indices

We compared the newly created abundance index (GAM_new) with the core vessel index by the conventional GLM and the one obtained by GAM used for the 2020 stock assessment (GAM11) (Fig. 50). The overall trends were similar to each other. Compared to the other two series, the new GAM series had lower values from 1970 to 1990. Also, the value in 1993-1994 was high, and it was high in 2019 and 2022. In 2022, although all indices increased from 2021 the increase in GAM_new was significantly large. It should be noted that the high value of 2018 in the core vessel by GLM, which was a problem in the 2020 assessment, is no longer seen due to the subsequent addition of data.

4. Discussion

The 2022 fishing data was added to the dataset. The method using GAM agreed at the 2022 ESC was able to obtain a convergent solution without changing the settings such as the k parameter. For various sensitivity analyses, we were able to obtain results similar to those in the 2022 work. The distribution of the residuals for each variable and the overall fitting of the residuals were the same as in the 2022 work, and no problems were observed. The results of the sensitivity analysis were similar to those of the 2022 work.

The 2022 index value increased significantly from 2021 and was the highest value in the series since 1969. This result is not surprising as the nominal CPUE also increased in 2022. However, the GLM index from the core vessels in 2022 is significantly lower than those in 2018-2020. It reminds us that the sharp rise in the core vessel CPUE index in 2018 triggered the CPUE redevelopment, so further examination may be warranted on this point. The 2018 spike was due to overestimation of the non-fishing spatio-temporal values, but the current GAM values for 2022 show no such sign. More specific analysis results

are presented in another document (Itoh and Takahashi 2023).

5. References

CCSBT (2007) Report of the Twelfth Meeting of the Scientific Committee. 10 - 14 September 2007. Hobart, Australia. 80pp.

CCSBT (2019) Report of the Twenty Fourth Meeting of the Scientific Committee. 7 September 2019. Cape Town, South Africa. 121pp.

CCSBT (2021) Report of the Twenty Sixth Meeting of the Scientific Committee. 31 August 2021. Online. 93pp.

CCSBT (2022a) Report of the Twelfth Operating Model and Management Procedure Technical Meeting. 20-24 June 2022. Hobart, Australia and Online. 33pp.

CCSBT (2022b) Report of the Twenty Seventh Meeting of the Scientific Committee. 29 August-5 September 2022. Online. 114pp.

Hoyle, S. (2022) Validating CPUE model improvements for the primary index of southern bluefin tuna abundance. CCSBT-OMMP/2206/04.

Itoh, T., E. Lawrence, and J. G. Pope (2008) The development of new agreed CPUE series for use in future MP work. CCSBT-ESC/0809/09.

Itoh, T. and N. Takahashi (2022) Development of the new CPUE abundance index using GAM for southern bluefin tuna in CCSBT. CCSBT-OMMP/2206/08.

Itoh, T. and N. Takahashi (2023) Further examination of CPUE abundance index using GAM for southern bluefin tuna based on predicted values. CCSBT-OMMP/2306/09.

Nishida T. and S. Tsuji (1998) Estimation of abundance indices of southern bluefin tuna (*Thunnus maccoyii*) based on the coarse scale Japanese longline fisheries data (1969-97). CCSBT/SC/9807/13.

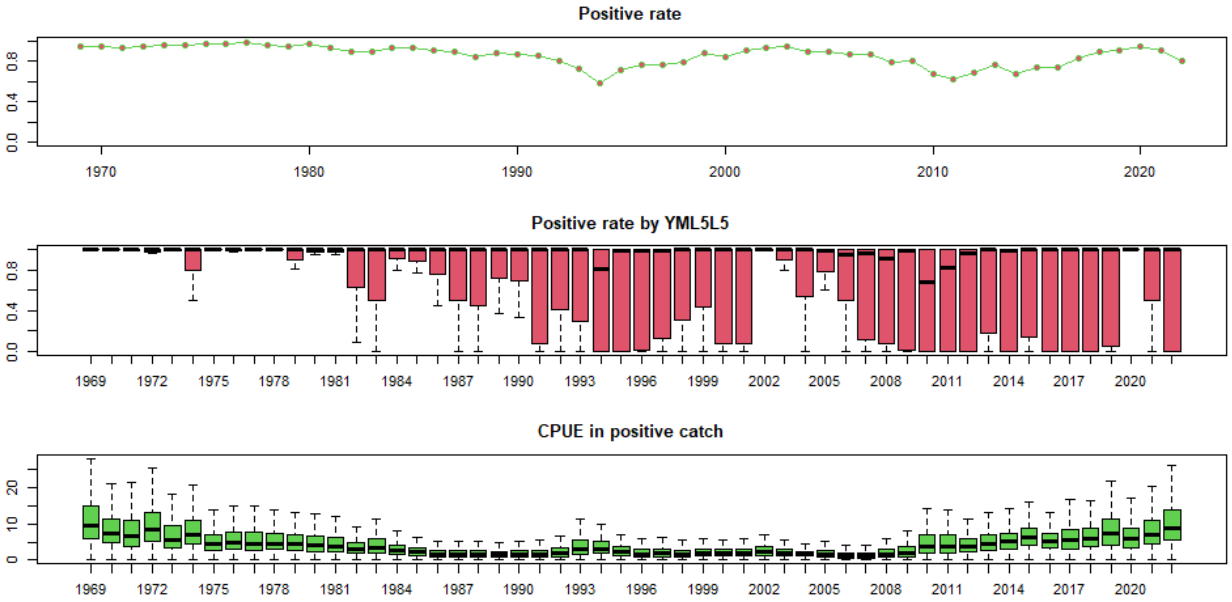

Fig. 1. Nominal value of positive catch rate and CPUE by year.
Upper panel is the positive rate which is the total number of positive catch operations / the total number of all records. Middle panels is boxplot based on the positive catch rate by year, month, 5 degree latitude and 5 degree longitude. Lower panel is CPUE in positive catch records.

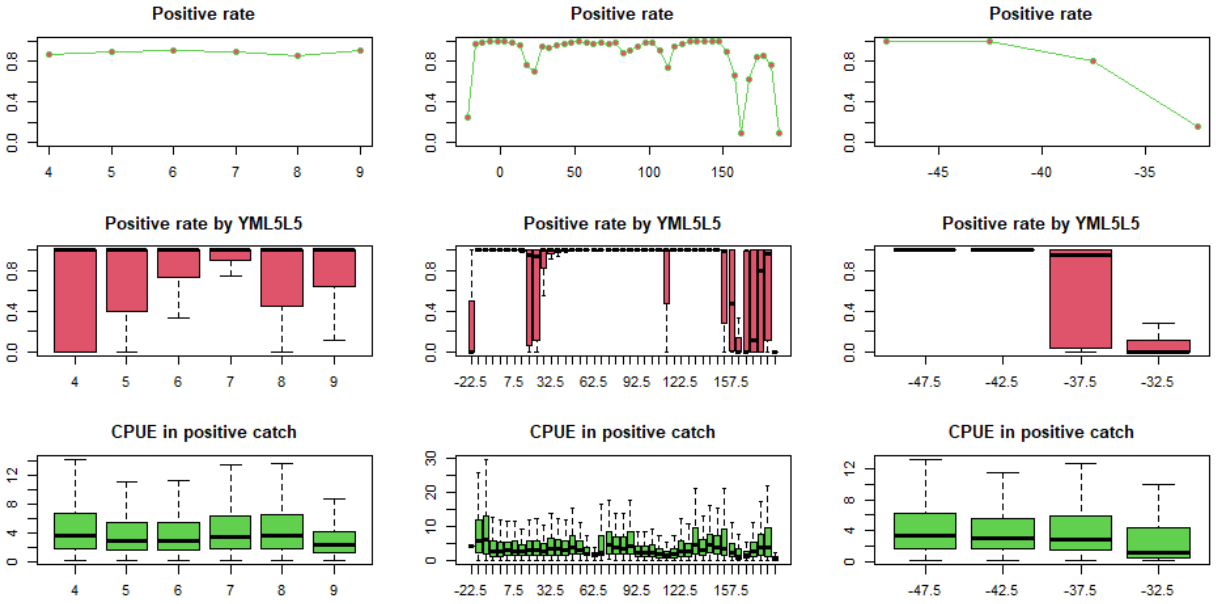

Fig. 2. Nominal value of positive catch rate and CPUE by month, longitude and latitude.
Upper panel is the positive rate which is the total number of positive catch operations / the total number of all records. Middle panels is boxplot based on the positive catch rate by year, month, 5 degree latitude and 5 degree longitude. Lower panel is CPUE in positive catch records.

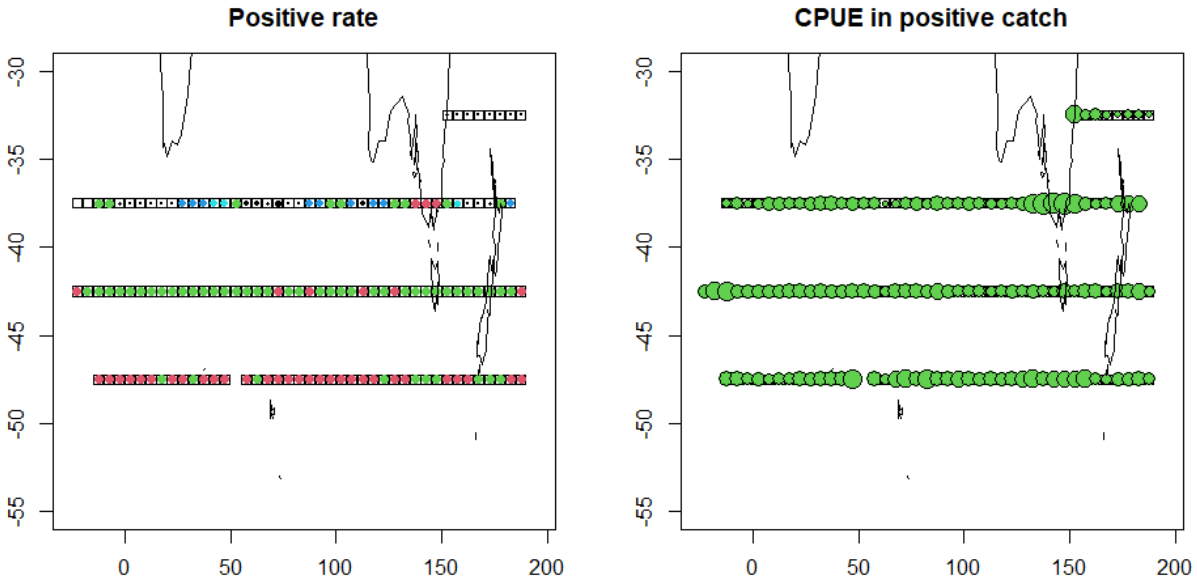

Fig. 3. Nominal value of positive catch rate and CPUE in map.
Left panel is the positive rate. Right panel is CPUE in positive catch records. Red is the higher value, followed by green, blue and white in the positive catch rate panel.

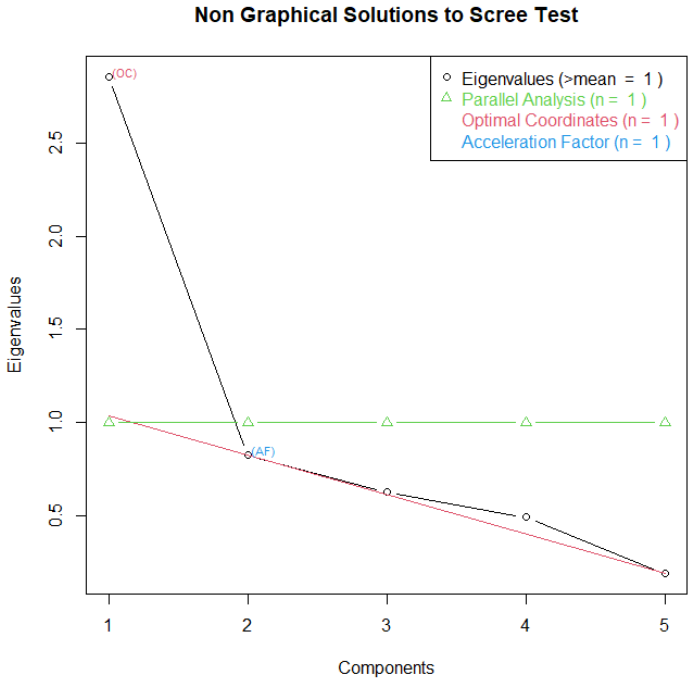

Fig. 4. Eigen values for the number of components in cluster analysis.

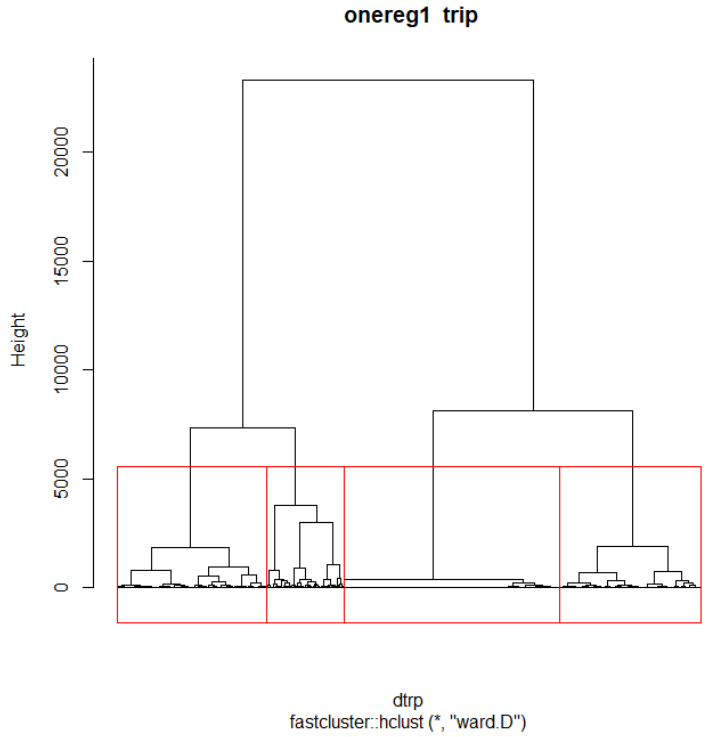

Fig. 5. Dendrogram of the cluster analysis.

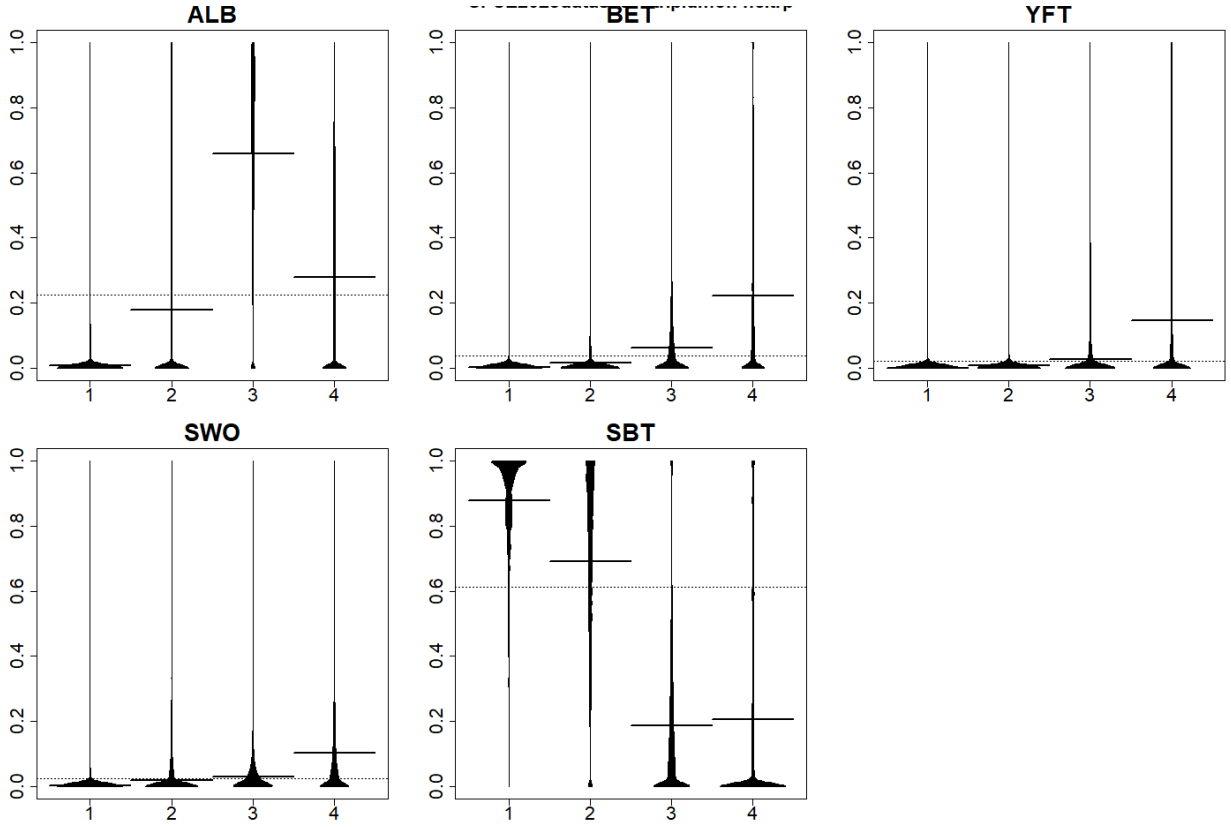

Fig. 6. Occurrence by species in each group in cluster analysis.
ALB is albacore, BET is bigeye tuna, YFT is yellowfin tuna, SWO is swordfish and SBT is southern bluefin tuna.

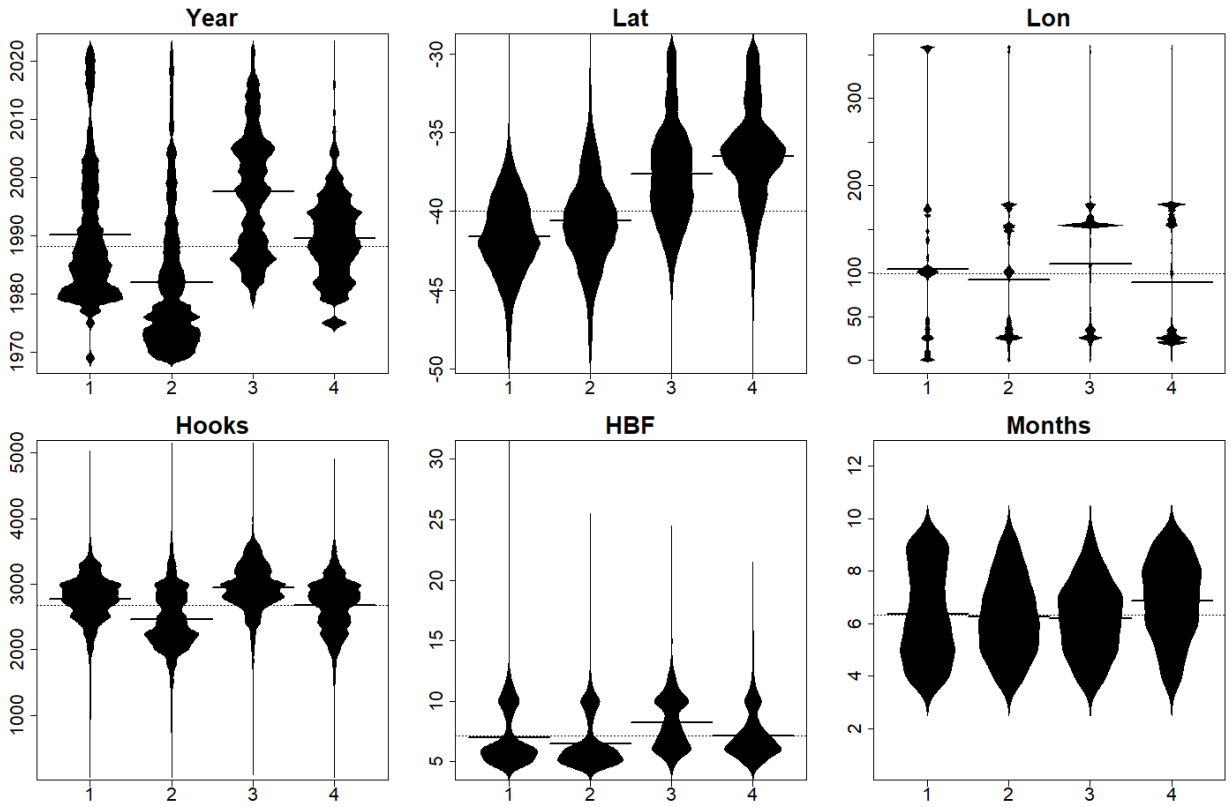

Fig. 7. Occurrence by variables of each group in the cluster analysis.

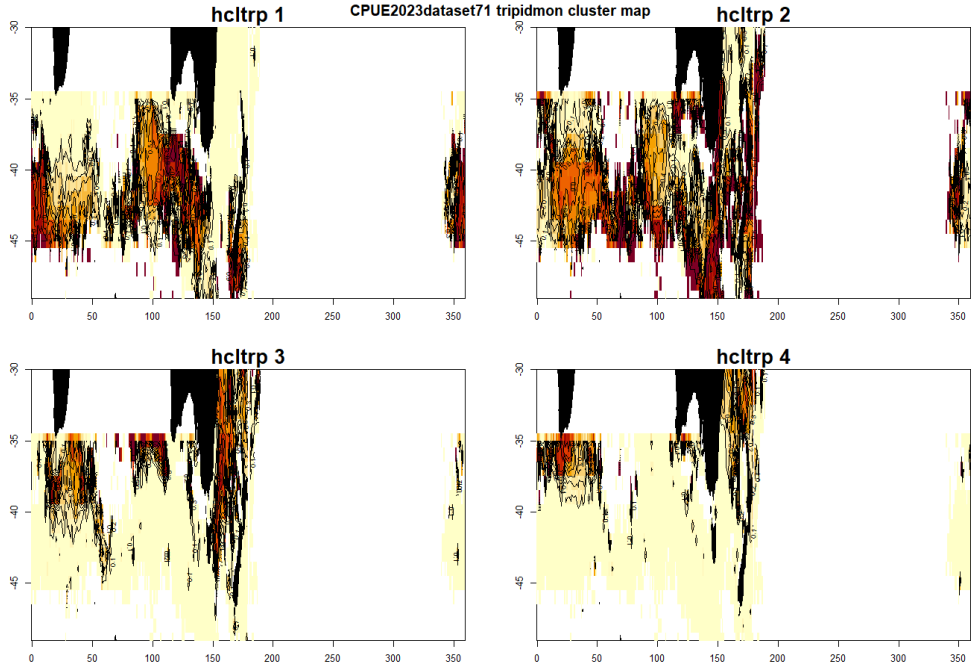

Fig. 8. Occurrence on map by group in the cluster analysis.

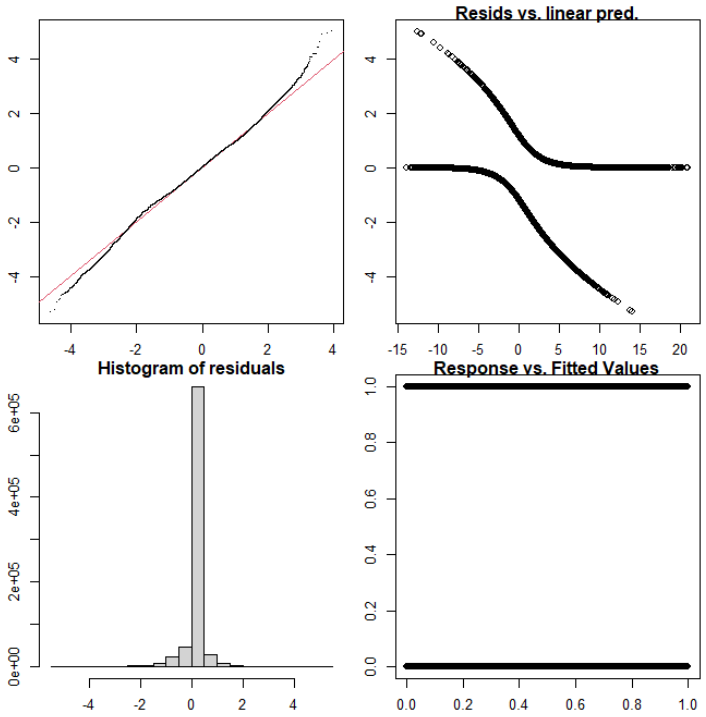

Fig. 9. Diagnostic plots for the binomial sub-model in the base case run.

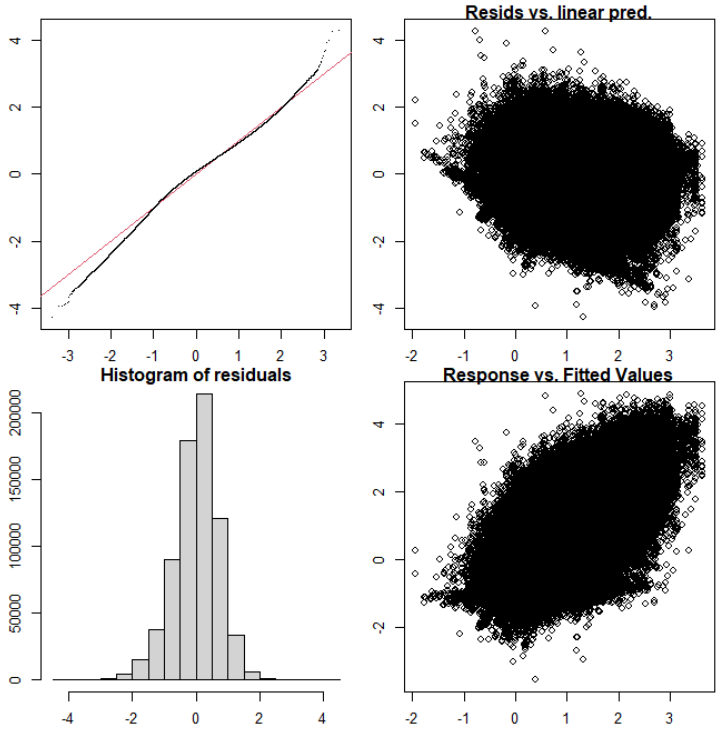

Fig. 10. Diagnostic plots for the positive catch sub-model in the base case run.

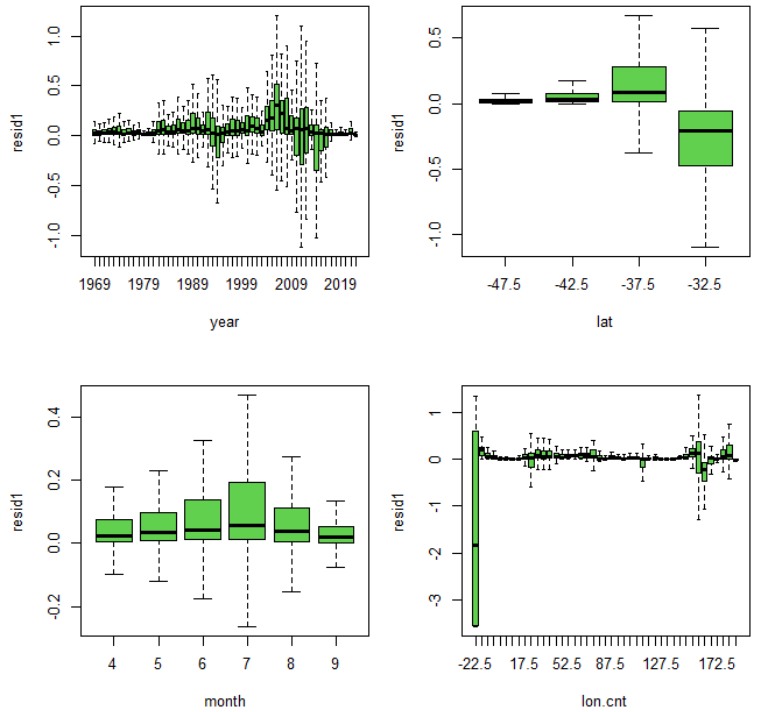

Fig. 11. Residuals by variable in the binomial sub-model in the base case run.

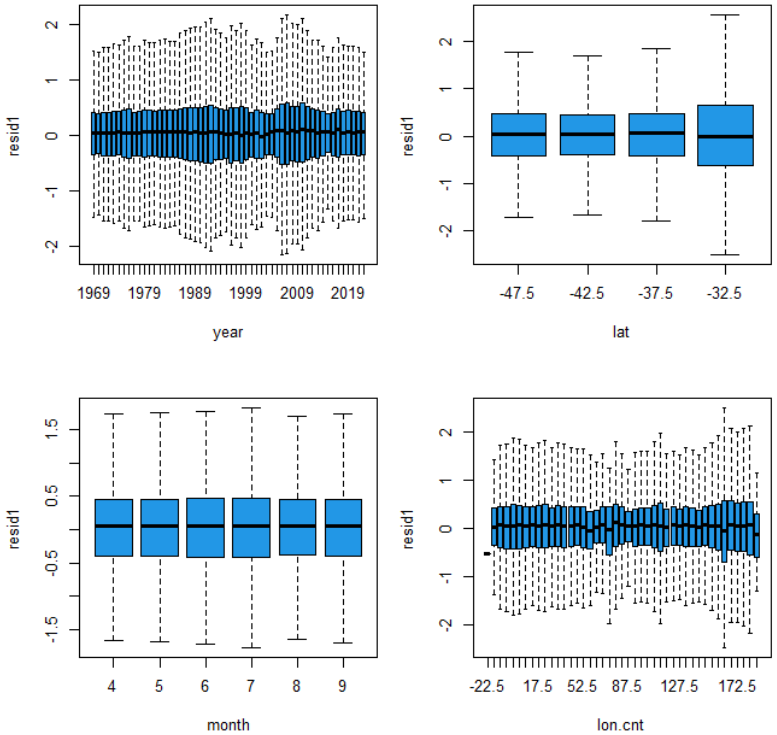

Fig. 12. Residuals by variable in the positive catch sub-model in the base case run.

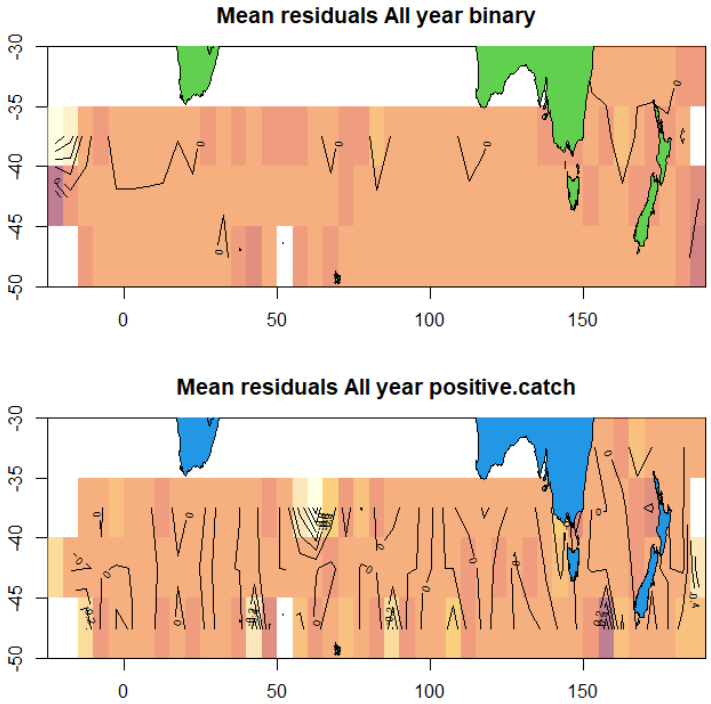

Fig. 13. Residual on maps for both sub-models in the base case run.

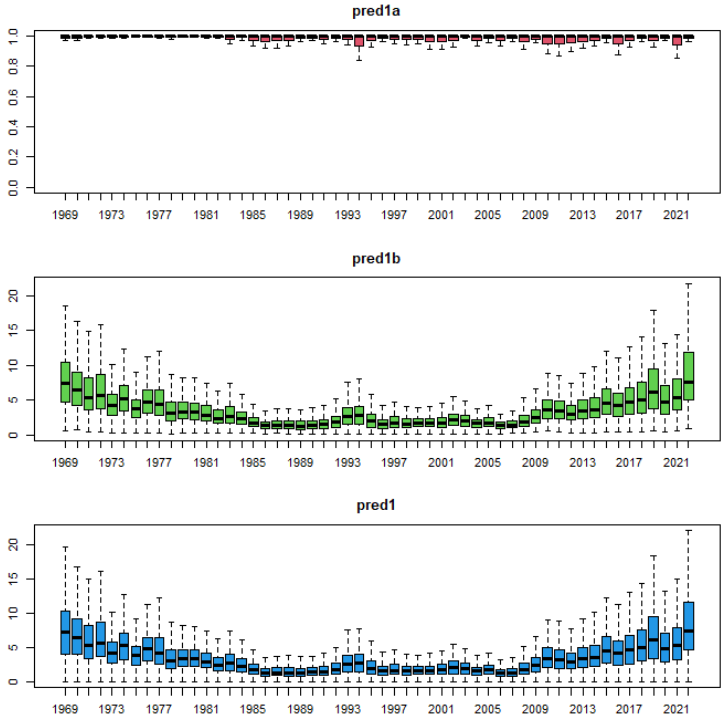

Fig. 14. Predicted value by year in the base case run.

Upper panel is the positive rate obtained from the binomial sub-model. Middle panels is CPUE obtained from the positive catch sub-model. Lower panel is product of the two.

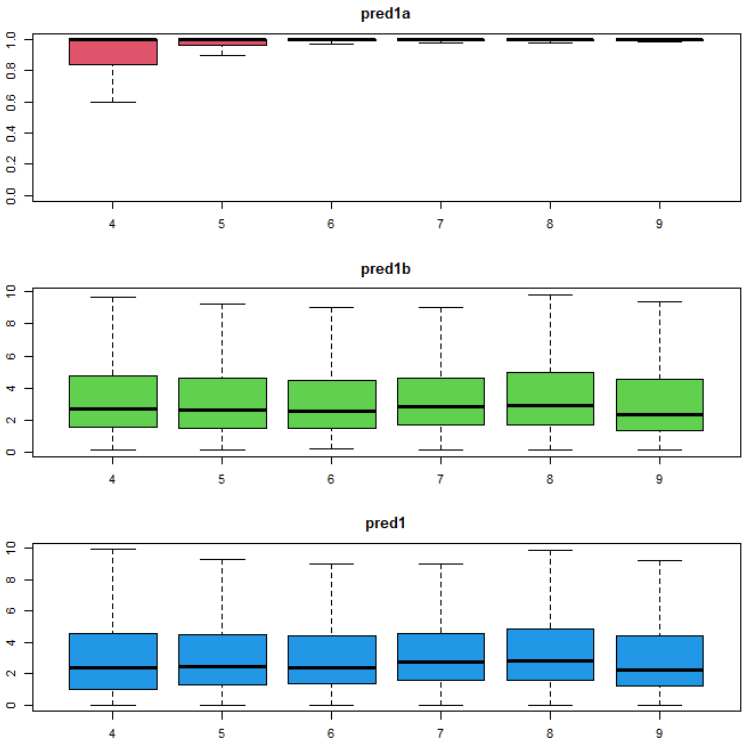

Fig. 15. Predicted value by month in the base case run.
See Fig. 14.

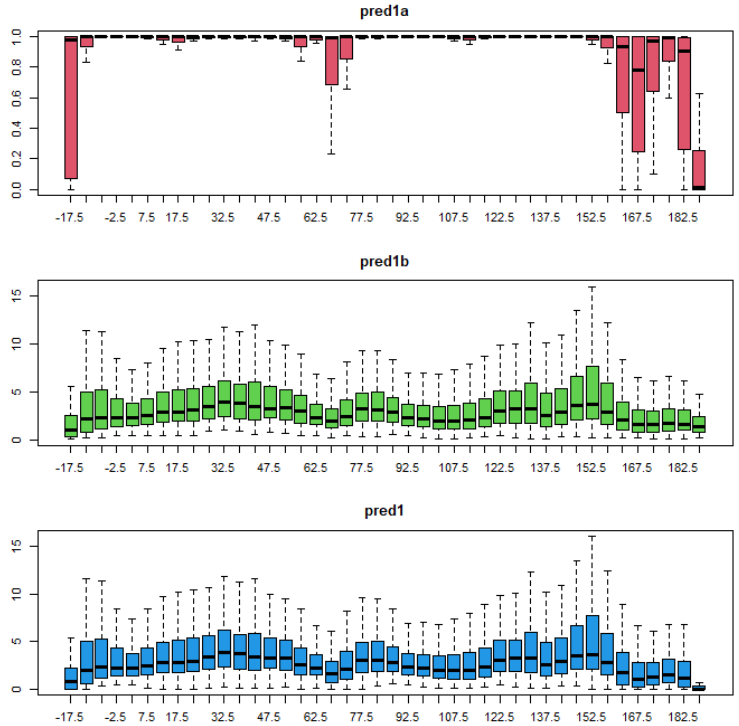

Fig. 16. Predicted value by longitude in the base case run.
See Fig. 14.

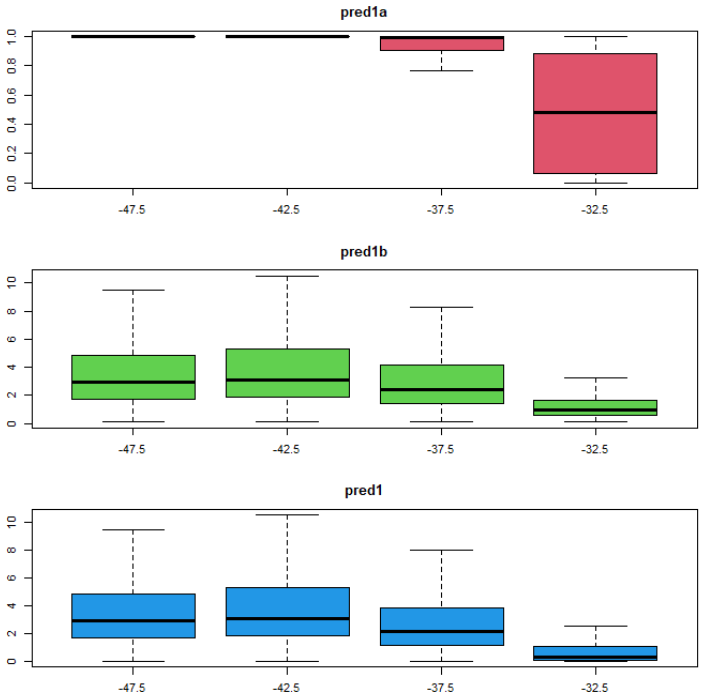

Fig. 17. Predicted value by latitude in the base case run.
See Fig. 14.

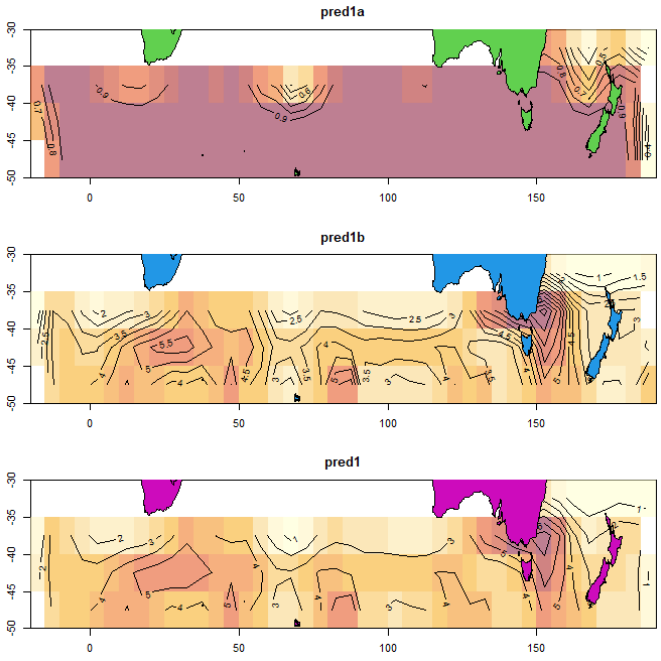

Fig. 18. Predicted value on map in the base case run.
See Fig. 14.

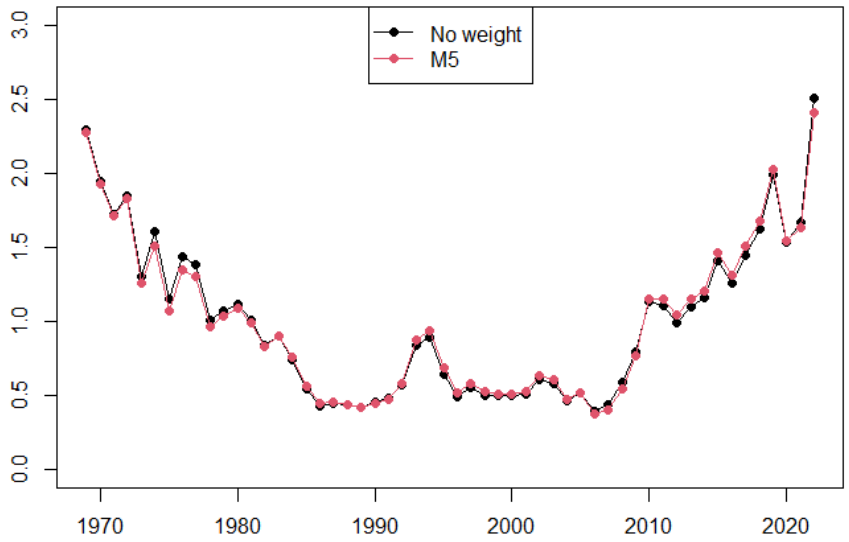

Fig. 19. Comparison of area weighted abundance indices in the base case run.
Red (M5) is area weighted abundance index which taking into account that the longitude length change over latitude and the number of 1x1 degree squares ever fished in a 5x5 degrees square. Black is the abundance index which weighting was not considered.

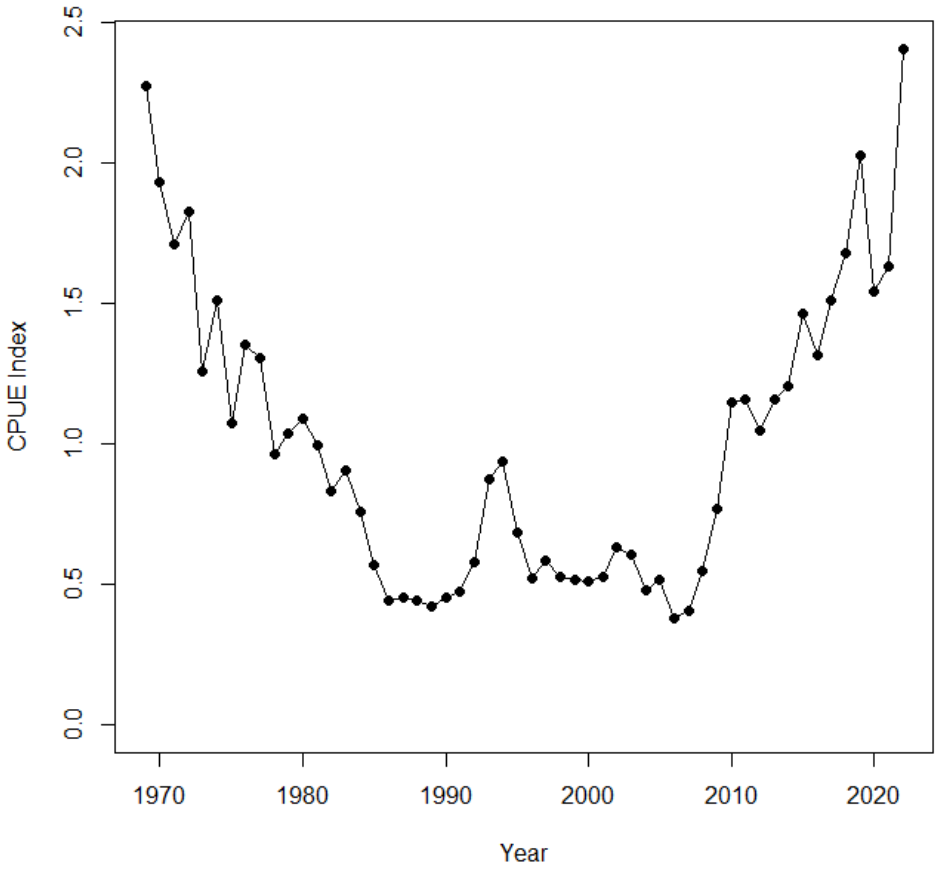

Fig. 20. CPUE abundance index for the base case.

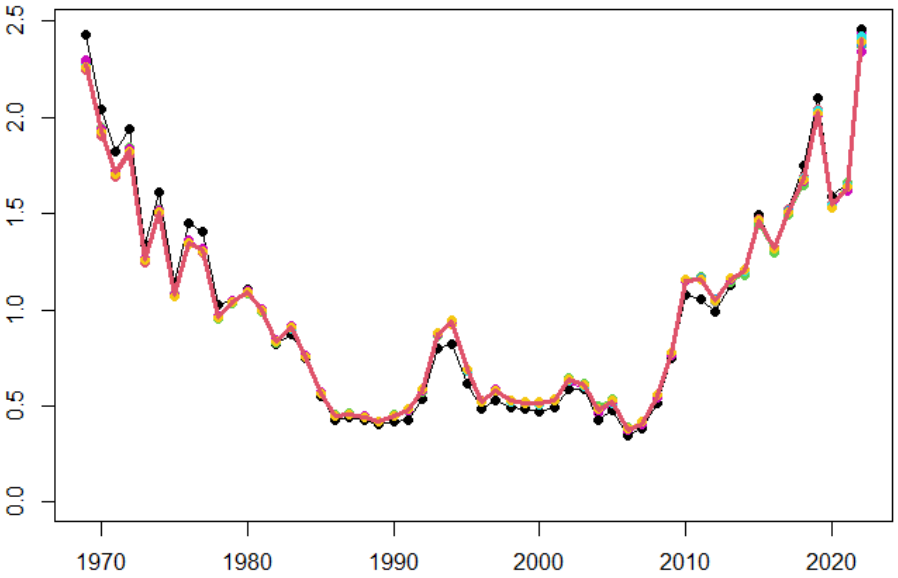

Fig. 21. Sensitivity analysis of model selection in the binomial sub-model for all runs.

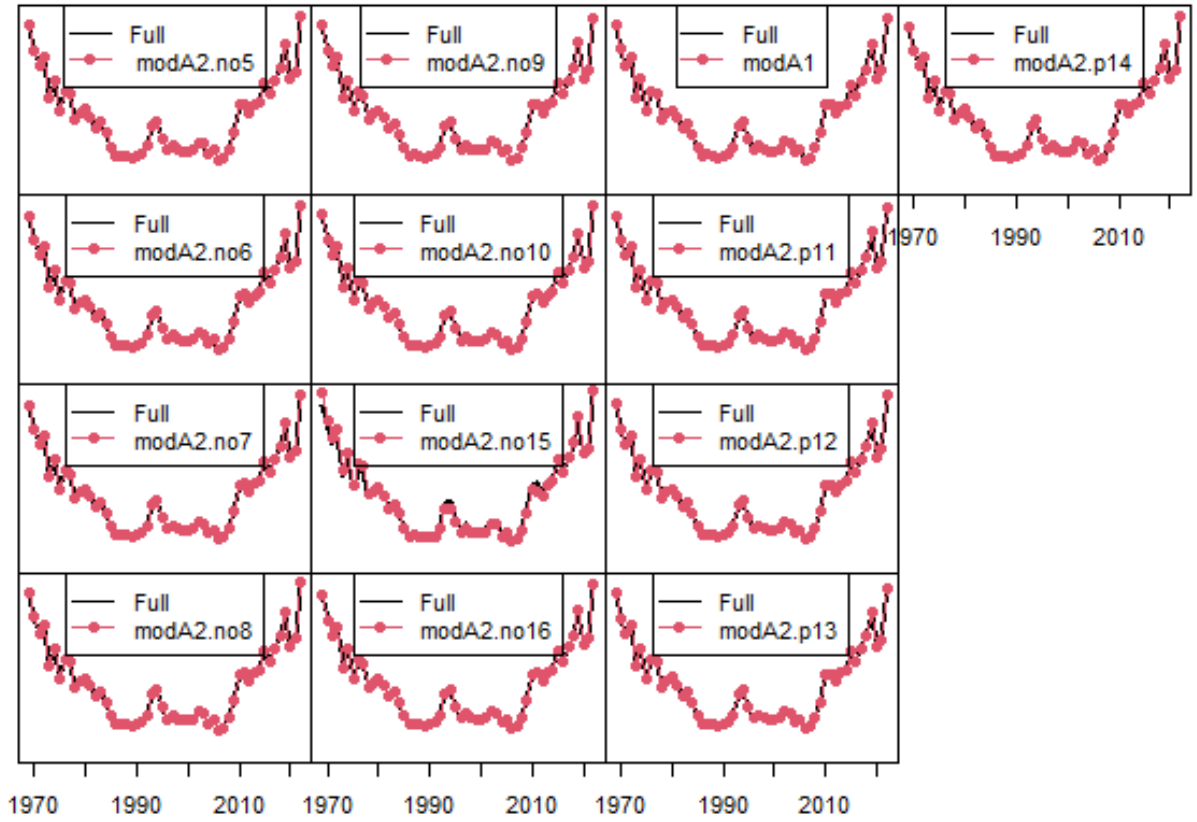

Fig. 22. Sensitivity analysis of model selection in the binomial sub-model for each run.

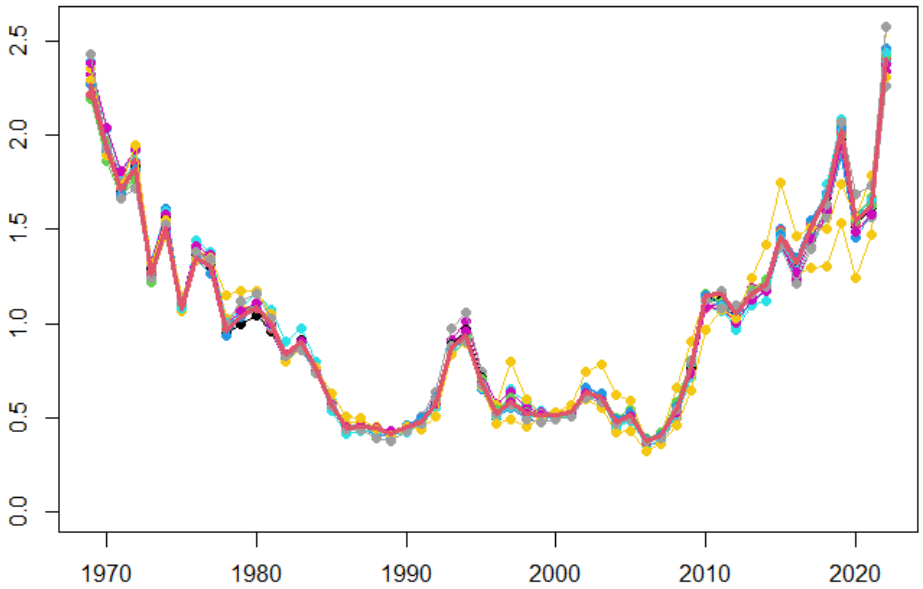

Fig. 23. Sensitivity analysis of model selection in the positive catch sub-model for all runs.

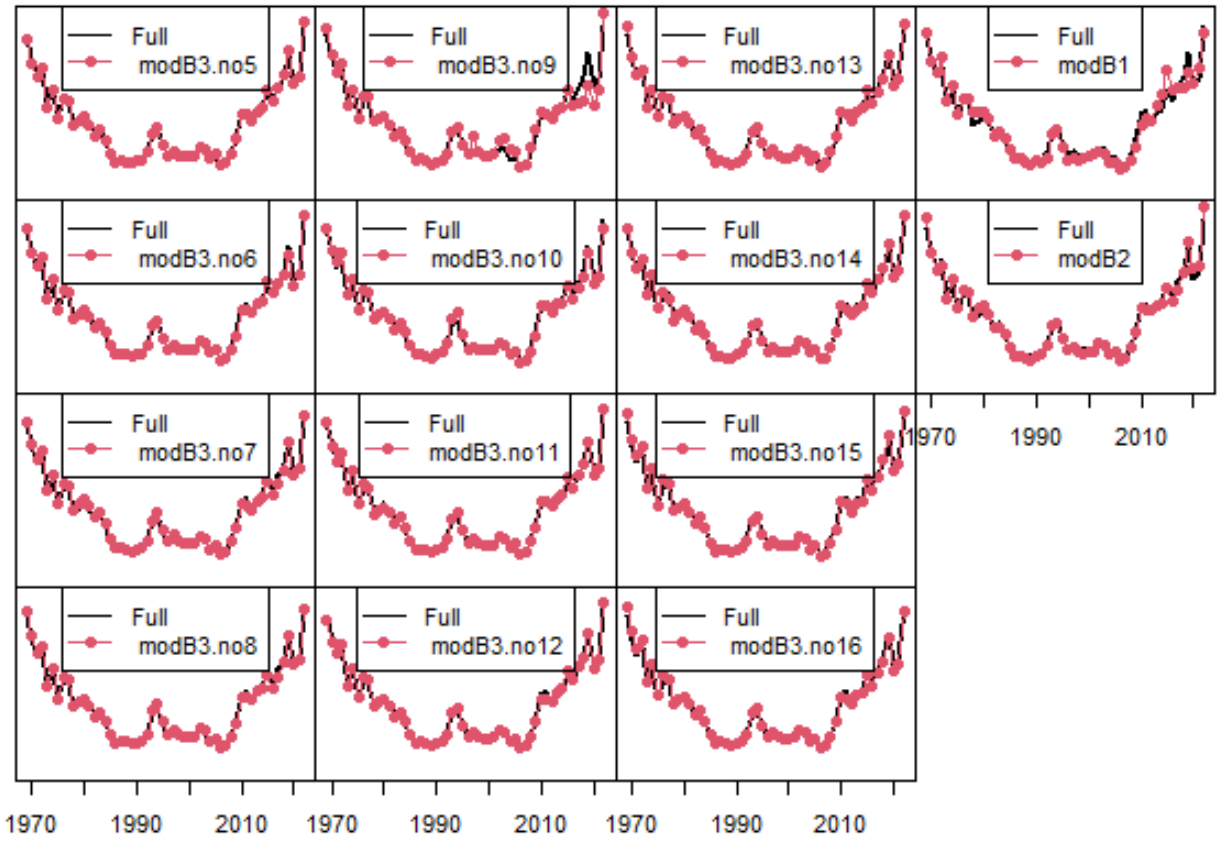

Fig. 24. Sensitivity analysis of model selection in the positive catch sub-model for each run.

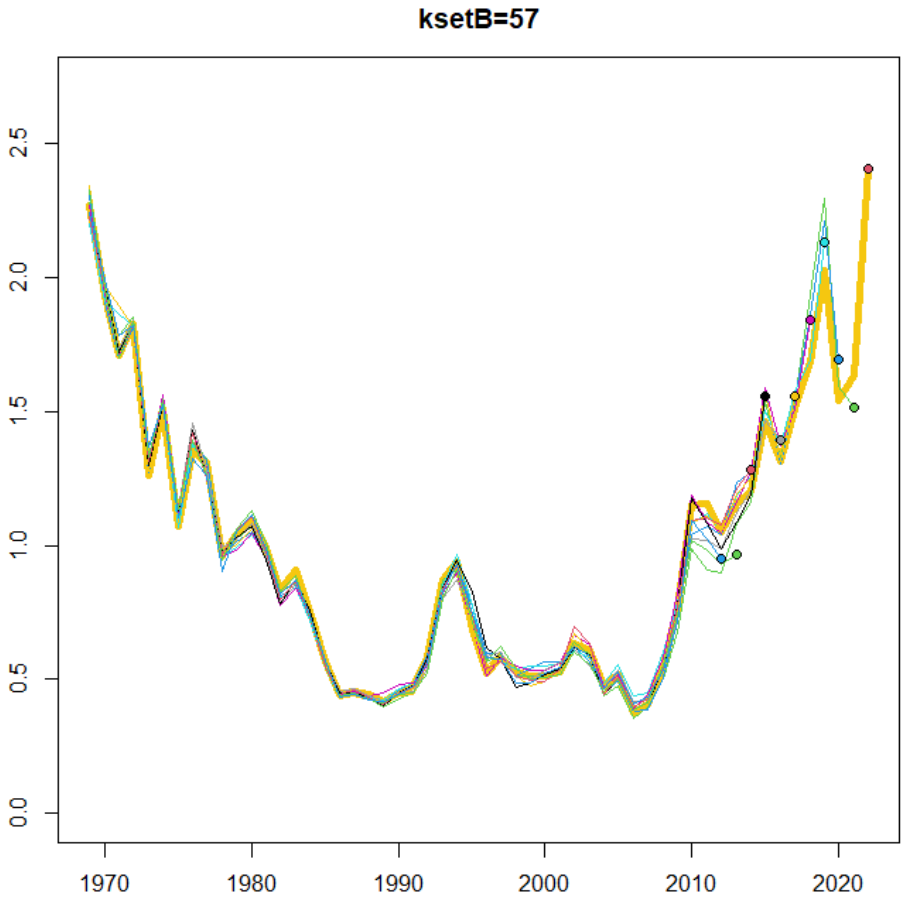

Fig. 25. Retrospective analysis for the base case model.

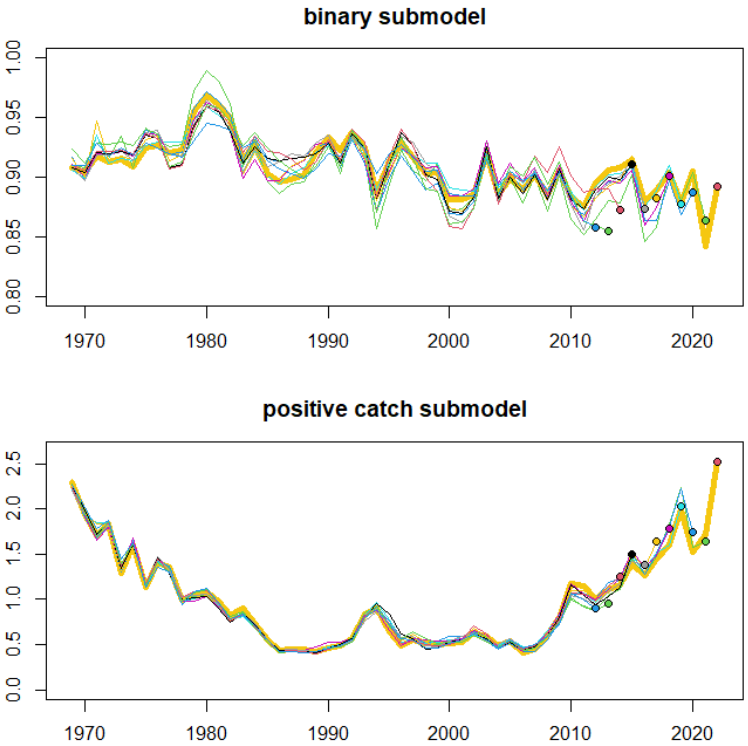

Fig. 26. Retrospective analysis for the base case model by sub-model.
Upper panel is by binomial submodel and lower panels is by positive catch submodel.

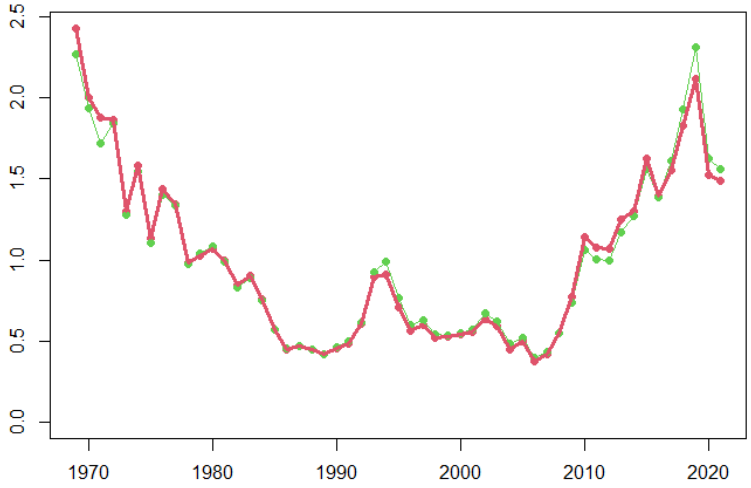

Fig. 27. Two indices from datasets up to 2021 by the datasets made in 2022 (red) and 2023 (green).

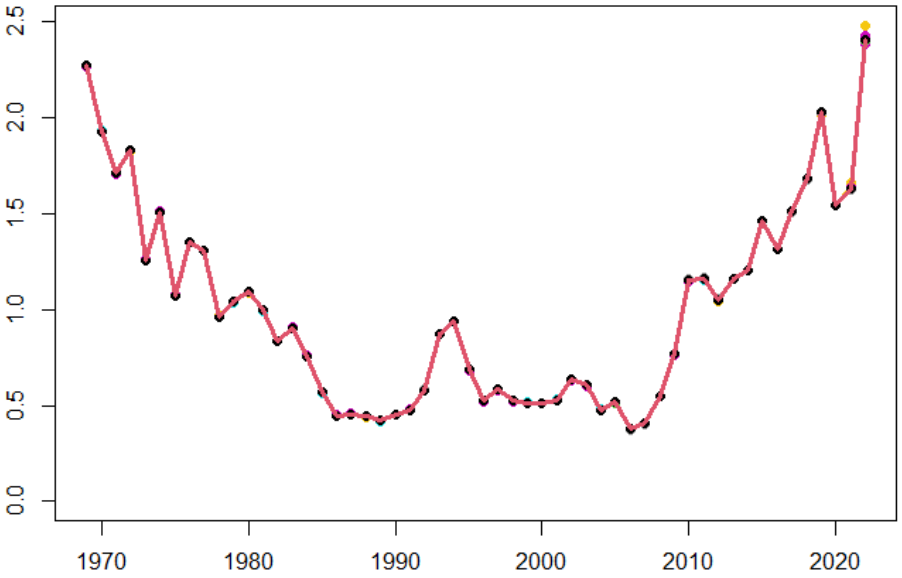

Fig. 28. Sensitivity analysis of k-value in the binomial sub-model for all runs.

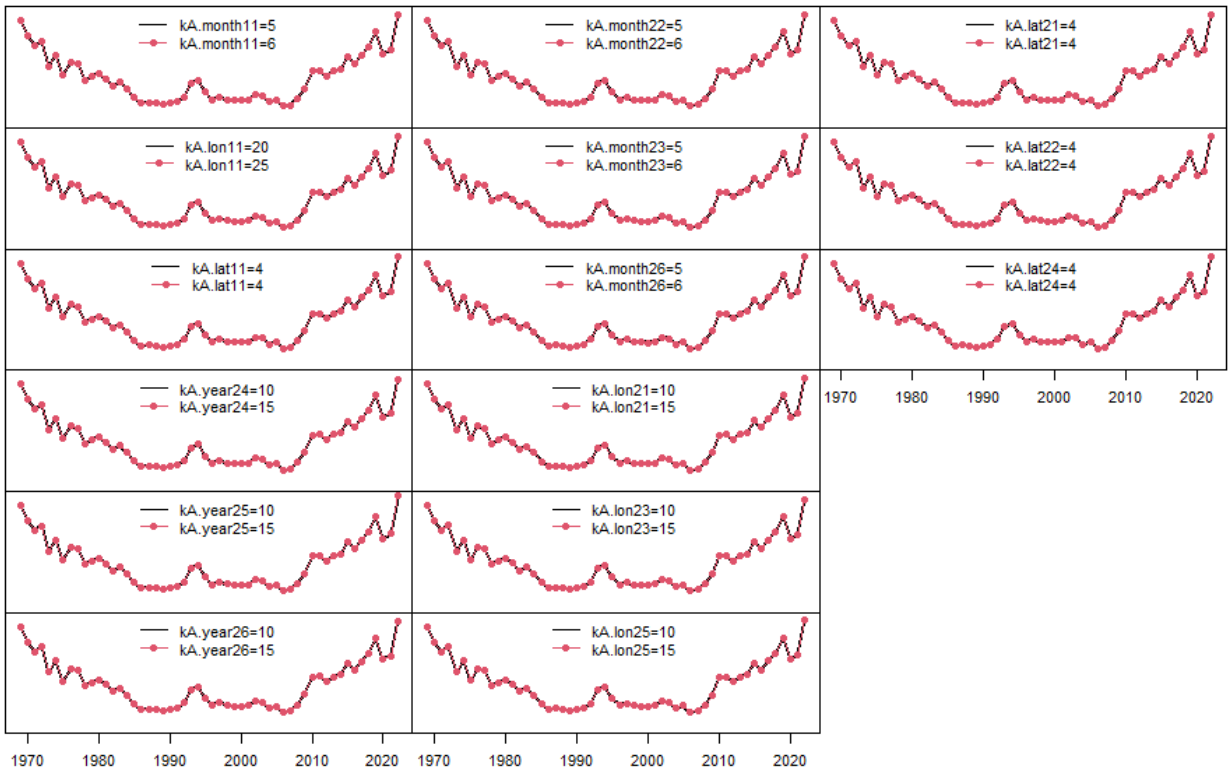

Fig. 29. Sensitivity analysis of k-value in the binomial sub-model for each of run.

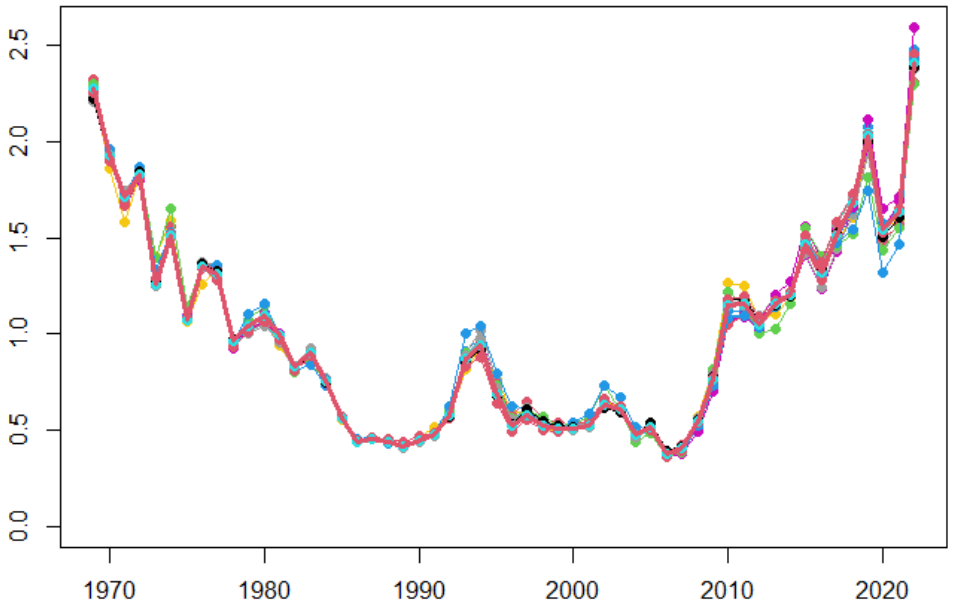

Fig. 30. Sensitivity analysis of k-value in the positive catch sub-model for all runs.

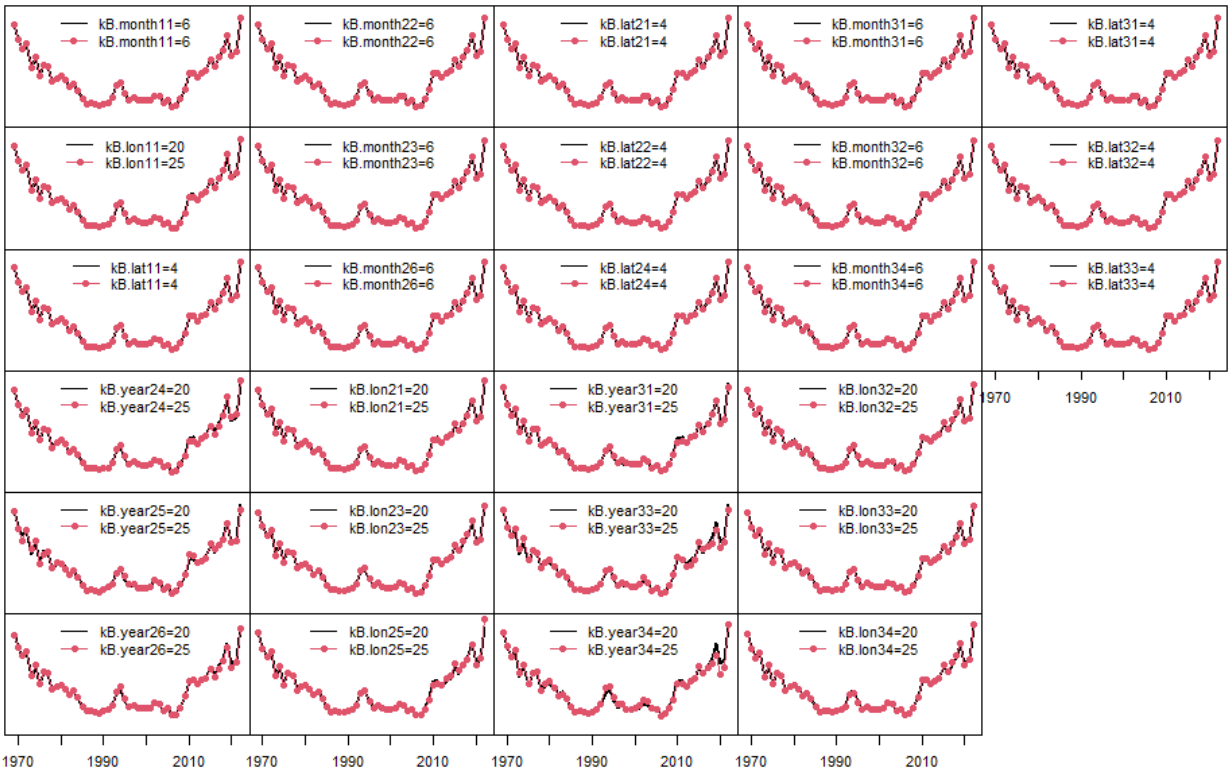

Fig. 31. Sensitivity analysis of k-value in the positive catch sub-model for each of run.

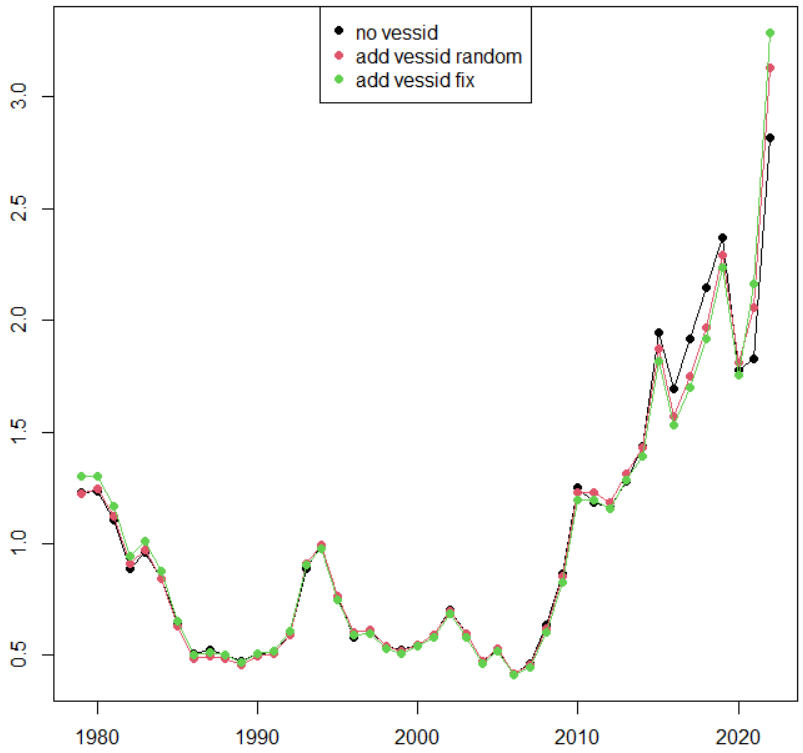

Fig. 32. Sensitivity analysis for the effect of vessel ID.

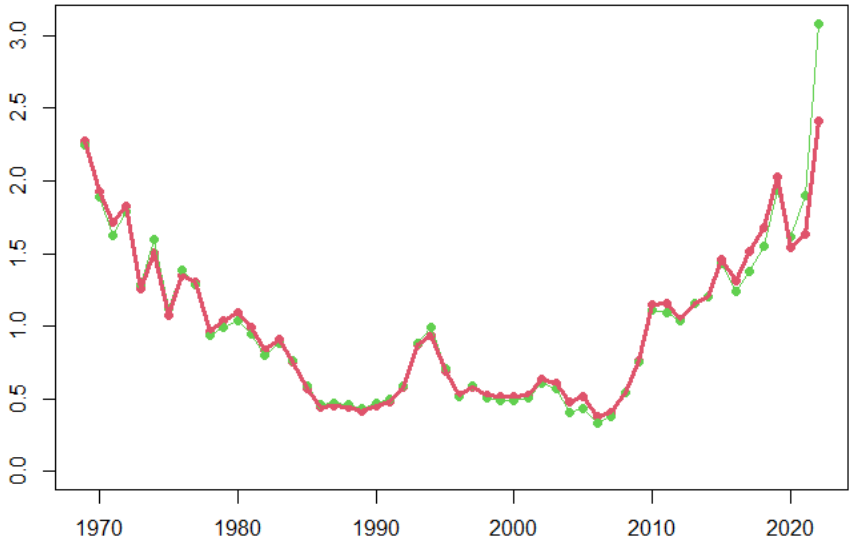

Fig. 33. Sensitivity analysis for the effect of eliminating 30S from the data.
Red is the base case, and green is the sensitivity run (eliminate 30S).

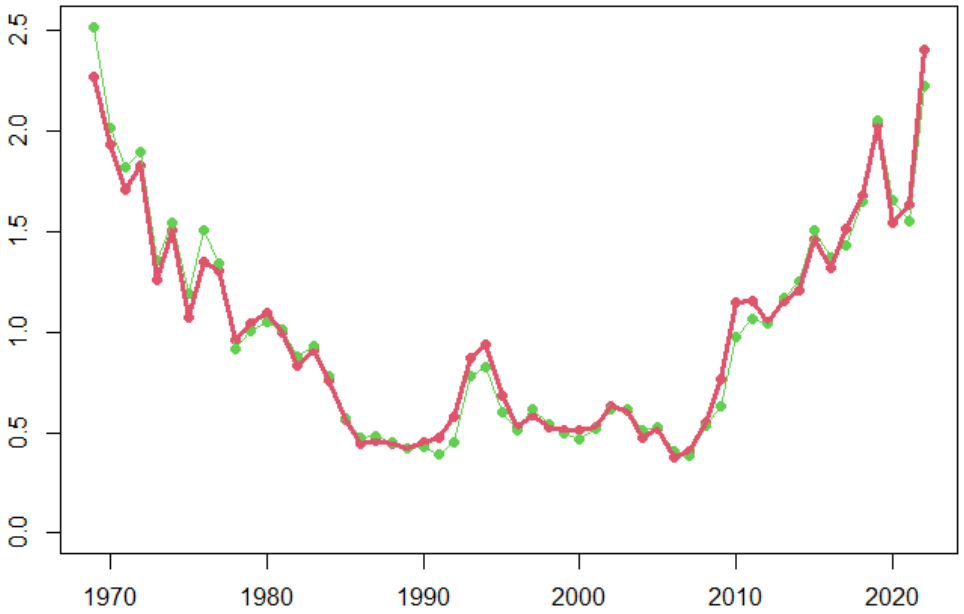

Fig. 34. Sensitivity analysis for the effect of age-5 plus instead of age-4 plus.
Red is the base case, and green is the sensitivity run (age-5 plus).

Fig. 35. Sensitivity analysis for the effect of of all ages instead of age-4 plus.
Red is the base case, and green is the sensitivity run (all ages).

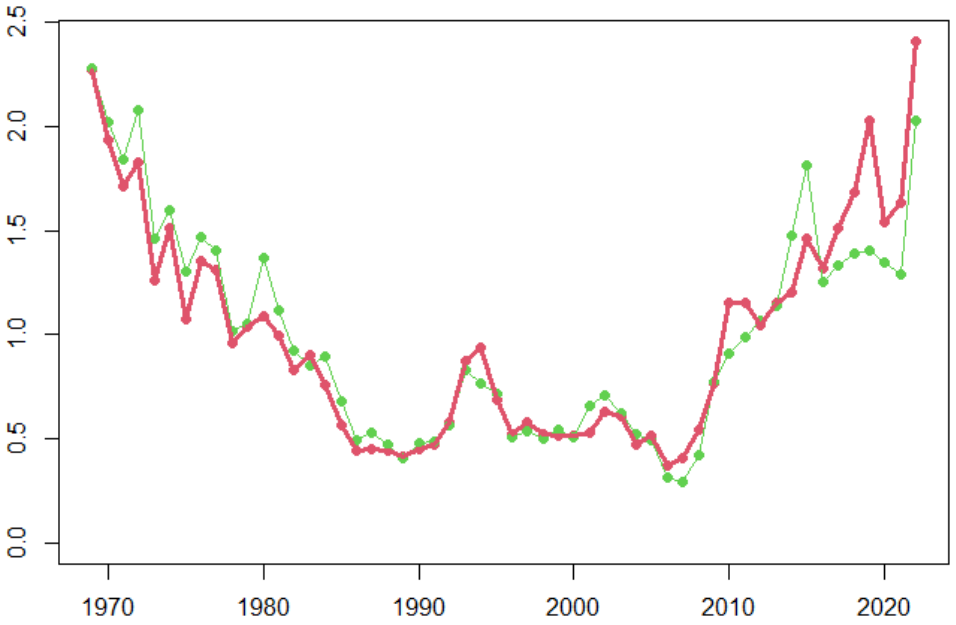

Fig. 36. Sensitivity analysis for the effect of data resolution in 5 degree.
Red is the base case, and green is the sensitivity run (aggregated by month and 5 degree latitude and 5 degree longitude).

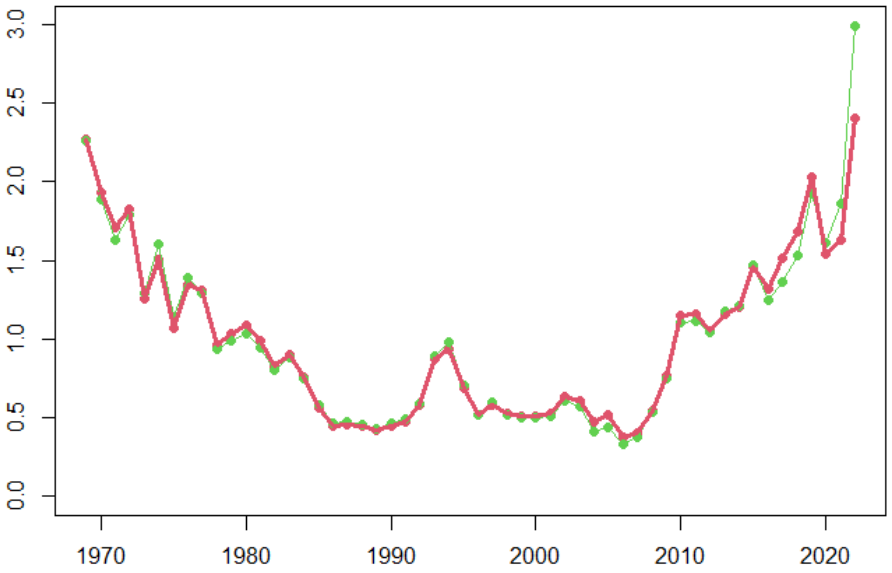

Fig. 37. Sensitivity analysis for the effect of data resolution in 1 degree.
Red is the base case, and green is the sensitivity run (aggregated by month and 1 degree latitude and 1 degree longitude).

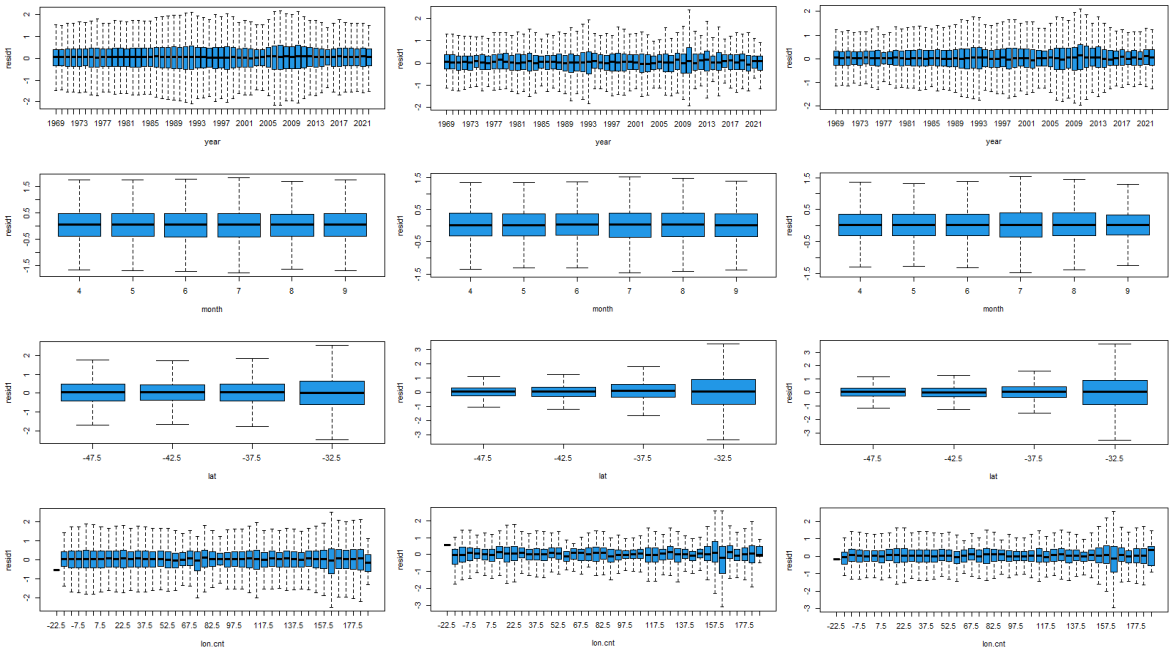

Fig. 38. Sensitivity analysis for the effect of data aggregation by residuals in the positive catch sub-model.
Left is shot-by-shot data, middle is 5 degree data and right is 1 degree data.

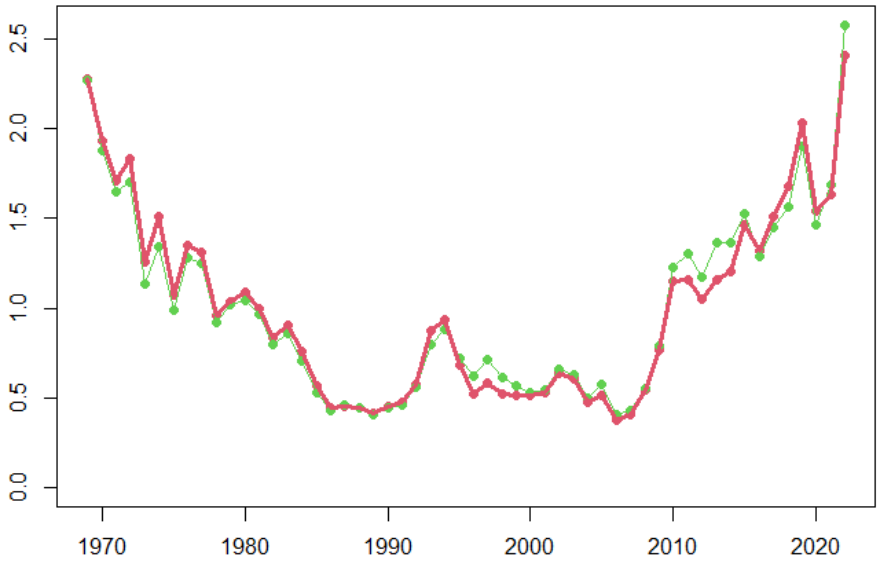

Fig. 39. Sensitivity analysis for the effect of resolution in model.
Red is the base case, where 5 degree for latitude and longitude, and green is the sensitivity run (1 degree in the model).

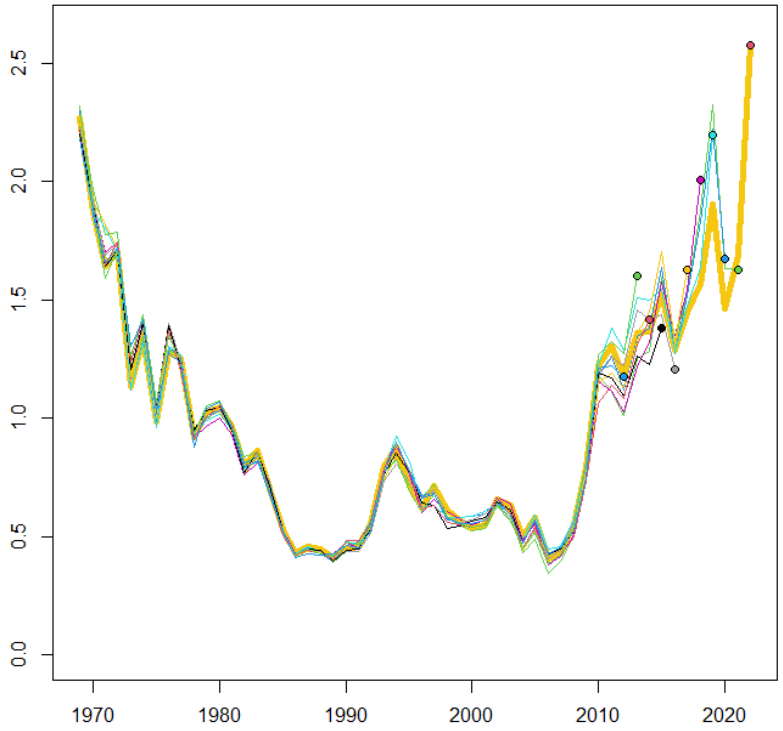

Fig. 40. Retrospective analysis in the sensitivity analysis of 1 degree model.

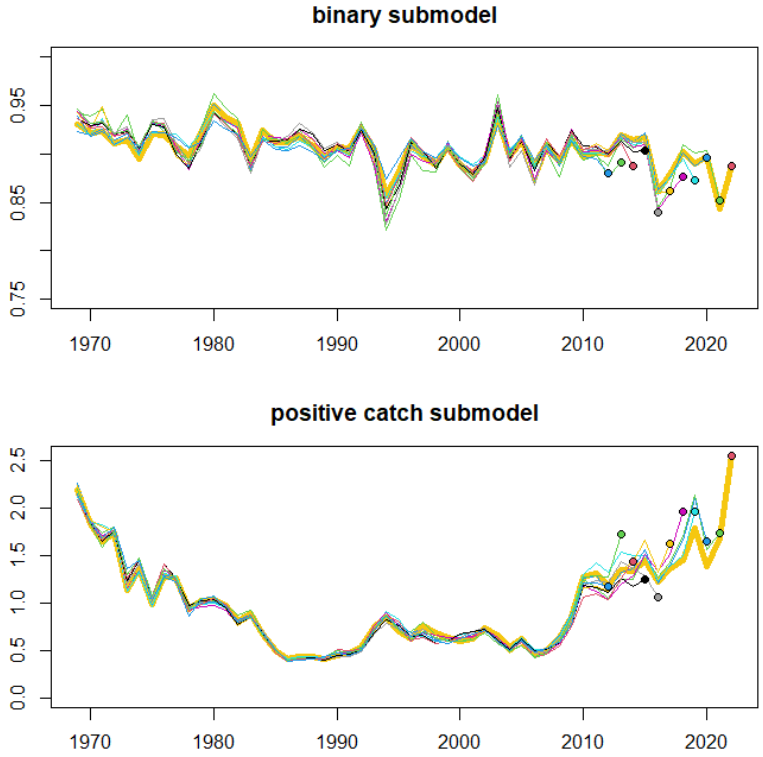

Fig. 41. Retrospective analysis in the sensitivity analysis of 1 degree model by sub-model.
Upper panel is by binomial submodel and lower panels is by positive catch submodel.

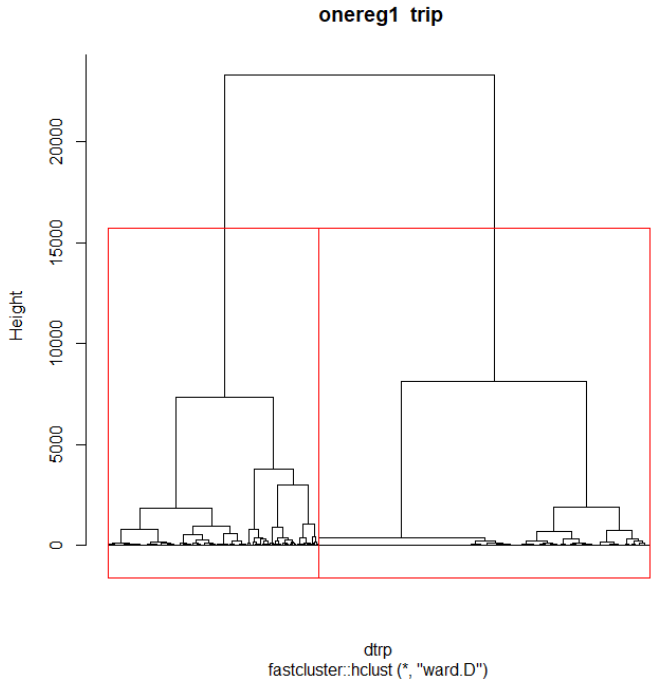

Fig. 42. Dendrogram in the cluster analysis of sensitivity analysis for the effect of 2 clusters.

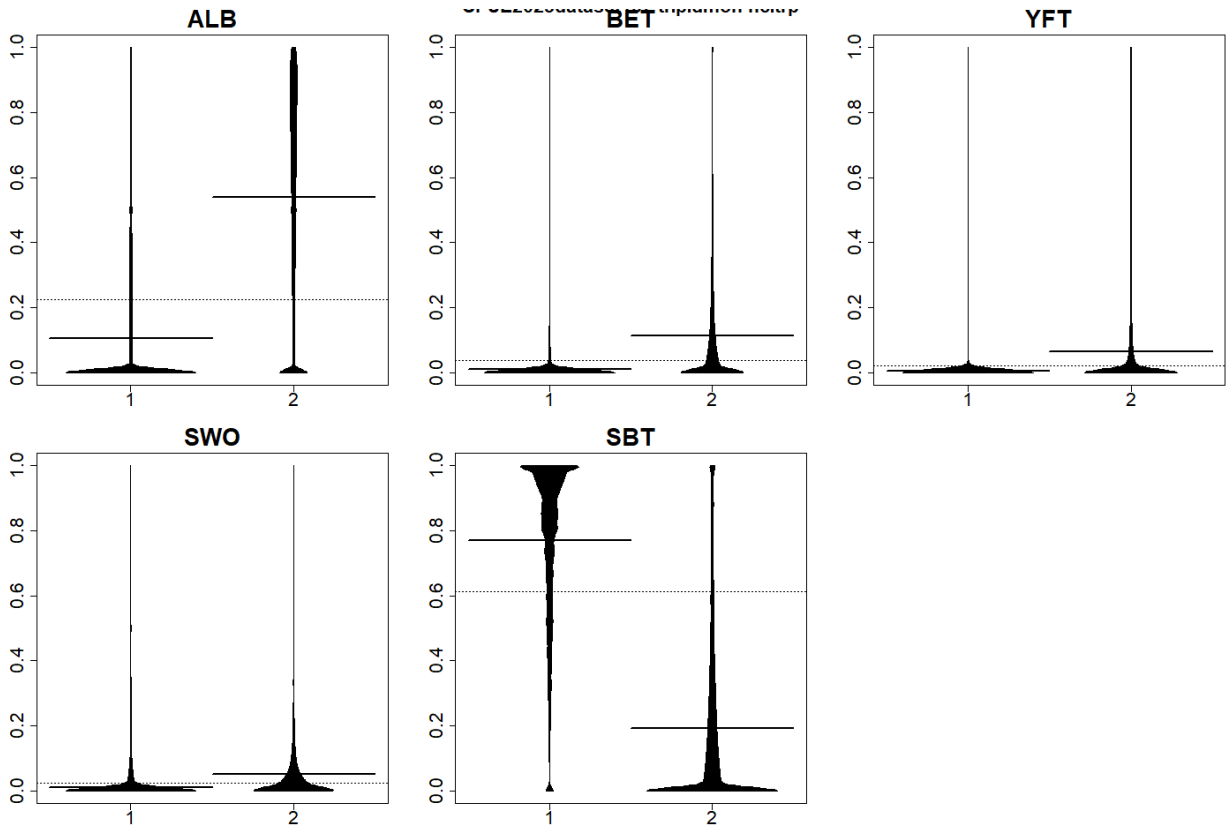

Fig. 43. Occurrence by species in the group in the 2 cluster analysis as a sensitivity analysis.

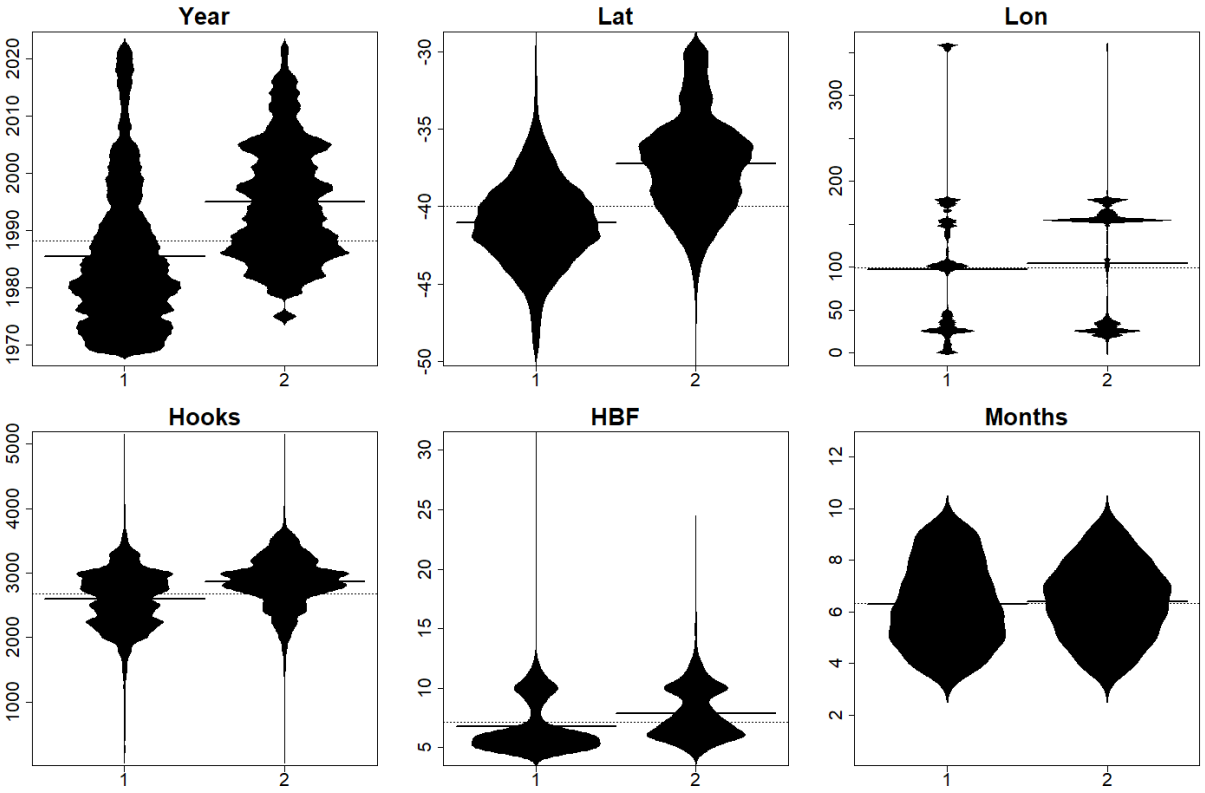

Fig. 44. Occurrence by variables in the group in the 2 cluster analysis as a sensitivity analysis.

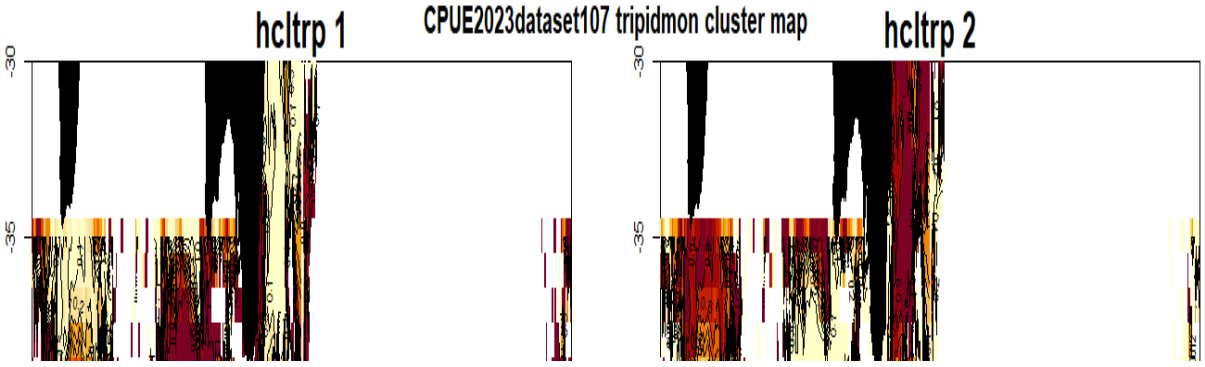

Fig. 45. Occurrence on map in the group in the 2 cluster analysis as a sensitivity analysis.

Fig. 46. Sensitivity analysis for the effect of 2 clusters instead of 4 clusters in the abundance index.
Red is the base case, and green is the sensitivity run (2 clusters).

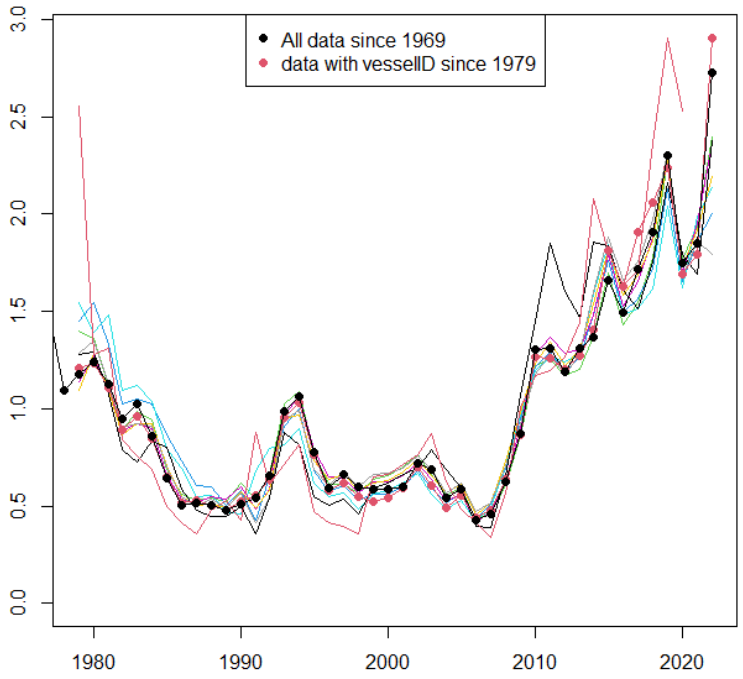

Fig. 47. Sensitivity analysis of core vessels selection.
Black and red lines with dots are the index from all vessels since 1969 and 1979, respectively. Black line is adjusted to the mean since 1979. Indices from different core vessels data are shown in different colors.

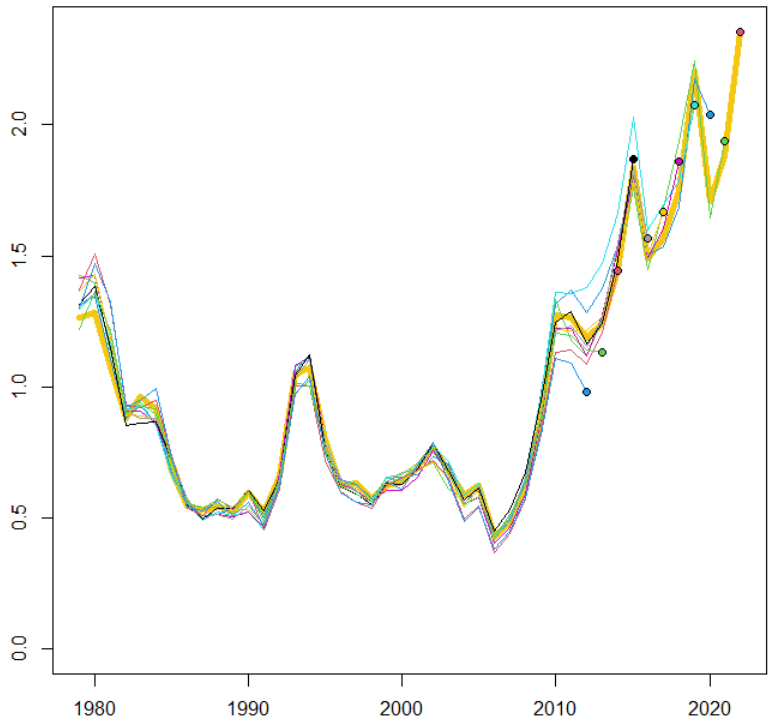

Fig. 48. Retrospective analysis in the sensitivity analysis of the core vessel.

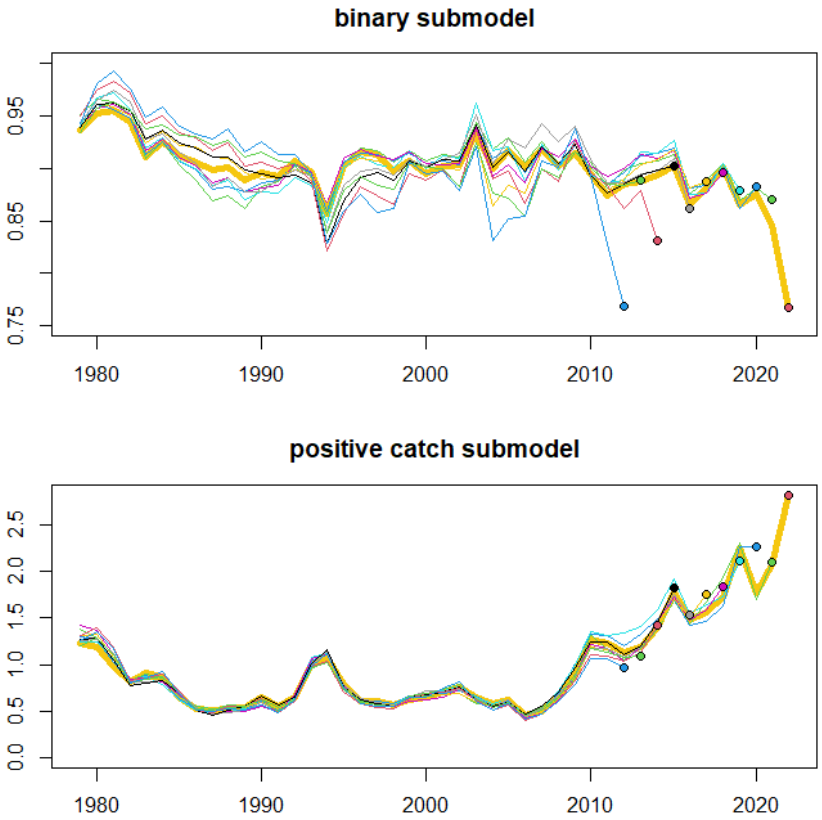

Fig. 49. Retrospective analysis in the sensitivity analysis of the core vessel by sub-model.
Upper panel is by binomial submodel and lower panels is by positive catch submodel.

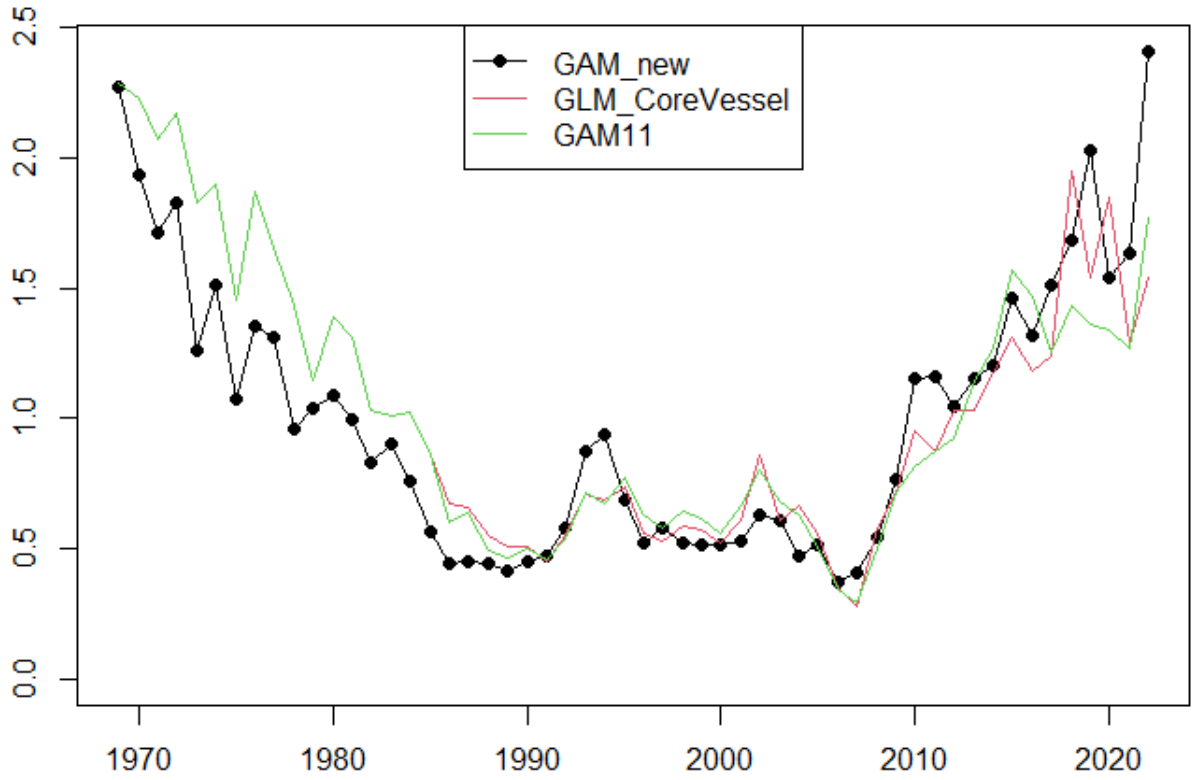

Fig. 50. Comparison of three abundance indices.

GLM_CoreVessel is the index by core vessel data with GLM base model in W0.8. GAM11 is the GAM model W0.8 used for the stock assessment in 2020.

Table 1. The k values selected for each of sub-model.

Submodel	BSM	PCSM
k.month11	5	6
k.lon11	20	20
k.lat11	4	4
k.year24	10	20
k.year25	10	20
k.year26	10	20
k.month22	5	6
k.month23	5	6
k.month26	5	6
k.lon21	10	20
k.lon23	10	20
k.lon25	10	20
k.lat21	4	4
k.lat22	4	4
k.lat24	4	4
k.year31		20
k.year33		20
k.year34		20
k.month31		6
k.month32		6
k.month34		6
k.lon32		20
k.lon33		20
k.lon34		20
k.lat31		4
k.lat32		4
k.lat33		4

Table 2. Statistics of choosing k values in the two sub-models of GAM.

BSM

Term	k'	edf	k-index	p-value
ti(month)	4	3.83	0.9800	0.19
ti(lon.cnt)	19	18.63	1.0078	0.86
ti(lat)	3	2.17	0.9365	0.00
ti(lon.cnt,lat)	27	18.53	0.9906	0.51
ti(month,lat)	12	8.89	0.9754	0.20
ti(lon.cnt,month)	36	31.36	0.9500	0.04
ti(year,lat)	27	22.04	0.9716	0.29
ti(year,lon.cnt)	81	73.51	0.9225	0.00
ti(year,month)	36	34.41	0.9449	0.05
s(log(hook))	9	8.28	0.9890	0.42

PCSM

Term	k'	edf	k-index	p-value
ti(month)	5	4.67	1.0122	0.84
ti(lon.cnt)	19	18.18	1.0016	0.52
ti(lat)	3	2.96	1.0024	0.56
ti(lon.cnt,lat)	43	33.71	0.9778	0.09
ti(month,lat)	14	12.15	1.0090	0.74
ti(lon.cnt,month)	92	71.14	0.9944	0.37
ti(year,lat)	57	46.19	0.9855	0.19
ti(year,lon.cnt)	334	248.38	0.9493	0.00
ti(year,month)	95	79.49	0.9748	0.04
ti(lat,month,year)	154	114.56	0.9717	0.05
ti(lat,lon.cnt,month)	86	65.33	0.9956	0.43
ti(lat,lon.cnt,year)	283	231.33	0.9658	0.02
ti(year,lon.cnt,month)	781	591.42	0.9448	0.00
s(log(hook))	9	7.57	1.0104	0.79

Table 3. Diagnostic statistics of GAM.

Sub-model	BSM	PCSM
n.data	797,416	704,842
dev.expl	73.52%	49.37%
AIC	152,230	1,495,091
residual.df	797,137	703,258
REMLscore	1,805,725	376,806

Table 4. Abundance index as the base case.

Year	Index
1969	2.27111
1970	1.92963
1971	1.71171
1972	1.82762
1973	1.25950
1974	1.50915
1975	1.07418
1976	1.35037
1977	1.30776
1978	0.96276
1979	1.03904
1980	1.09260
1981	0.99476
1982	0.83460
1983	0.90588
1984	0.75795
1985	0.56722
1986	0.44570
1987	0.45633
1988	0.44435
1989	0.41971
1990	0.45119
1991	0.47718
1992	0.58001
1993	0.87227
1994	0.93826
1995	0.68636
1996	0.52374
1997	0.58351
1998	0.52701
1999	0.51495
2000	0.51410
2001	0.52909
2002	0.63492
2003	0.60740
2004	0.47783
2005	0.51838
2006	0.37786
2007	0.40762
2008	0.54852
2009	0.76714
2010	1.15062
2011	1.15718
2012	1.04882
2013	1.15712
2014	1.20437
2015	1.46256
2016	1.31625
2017	1.51112
2018	1.68101
2019	2.02776
2020	1.54265
2021	1.63403
2022	2.40723

Table 5. Results of sensitivity analysis for model selection in the binomial sub-model.

name	term	dev.expl	AIC	residual.df	REMLscore	dAIC
modA2	Main and 2 way interactions	73.52%	152,230	797,137	1,805,725	2,618
modA2.no5	-ti(lon, lat)	73.04%	154,893	797,155	4,290,855	5,281
modA2.no6	-ti(month, lat)	73.26%	153,675	797,145	4,273,036	4,063
modA2.no7	-ti(lon, month)	73.22%	153,837	797,168	1,848,045	4,225
modA2.no8	-ti(year, lat)	72.84%	156,045	797,157	2,686,977	6,434
modA2.no9	-ti(year, lon)	72.09%	160,236	797,211	1,169,996	10,624
modA2.no10	-ti(year, month)	72.52%	157,875	797,171	1,435,820	8,264
modA2.no15	-cl	70.96%	166,867	797,137	2,485,989	17,255
modA2.no16	-s(log(hook))	73.10%	154,600	797,145	1,859,854	4,989
modA1	Main effects	66.88%	189,860	797,325	1,364,506	40,248
modA2.p11	+ti(lat, month, year)	73.90%	150,060	797,112	11,281,266	449
modA2.p12	+ti(lat, lon, month)	73.78%	150,753	797,109	199,913,668	1,142
modA2.p13	+ti(lat, lon, year)	73.74%	151,009	797,109	3,217,093	1,397
modA2.p14	+ti(year, lon, month)	73.99%	149,612	797,087	1,235,815	0

Table 6. Results of sensitivity analysis for model selection in the positive catch sub-model.

name	term	dev.expl	AIC	residual.df	REMLscore	dAIC
modB3	Full model	49.37%	1,495,091	703,258	376,806	0
modB3.no5	-ti(lon, lat)	49.35%	1,495,453	703,250	376,946	362
modB3.no6	-ti(month, lat)	49.34%	1,495,514	703,253	376,934	422
modB3.no7	-ti(lon, month)	49.31%	1,495,979	703,215	377,224	888
modB3.no8	-ti(year, lat)	49.35%	1,495,386	703,264	376,865	294
modB3.no9	-ti(year, lon)	49.18%	1,497,867	703,231	377,608	2,776
modB3.no10	-ti(year, month)	49.36%	1,495,280	703,261	376,858	189
modB3.no11	-ti(lat, month, year)	49.32%	1,495,654	703,326	376,854	563
modB3.no12	-ti(lat, lon, month)	49.30%	1,496,060	703,284	377,012	969
modB3.no13	-ti(lat, lon, year)	49.03%	1,499,547	703,434	377,616	4,456
modB3.no14	-ti(year, lon, month)	48.34%	1,508,165	703,806	379,244	13,073
modB3.no15	-cl	48.95%	1,500,923	703,253	378,282	5,832
modB3.no16	-s(log(hook))	49.33%	1,495,683	703,262	376,955	592
modB1	Main effects	41.33%	1,595,954	704,750	399,338	100,862
modB2	Main and 2 way interactions	47.33%	1,521,069	704,212	381,789	25,978

Table 7. Sensitivity analysis of core vessels chosen by various definition.

			Data records			Number of vessels		
Run	xx	yy	Original	core vessel data	%	Original	core vessel data	%
1	9999	1	600,977	600,977	100%	1,680	1,680	100%
2	56	3	600,977	233,037	39%	1,680	285	17%
3	56	5	600,977	164,880	27%	1,680	165	10%
4	56	7	600,977	121,012	20%	1,680	107	6%
5	50	3	600,977	216,063	36%	1,680	261	16%
6	40	3	600,977	182,759	30%	1,680	215	13%
7	30	3	600,977	143,502	24%	1,680	158	9%
8	20	3	600,977	91,271	15%	1,680	93	6%
9	20	5	600,977	57,220	10%	1,680	49	3%

			SBT catch			Run time in minutes			
Run	xx	yy	Original	core vessel data	%	BSM	PCSM	Total	%
1	9999	1	16,618,810	16,618,810	100%	3.78	7.75	11.54	53%
2	56	3	16,618,810	9,754,282	59%	1.41	3.41	4.83	22%
3	56	5	16,618,810	7,406,156	45%	1.07	1.71	2.78	13%
4	56	7	16,618,810	5,824,825	35%	0.84	1.40	2.23	10%
5	50	3	16,618,810	9,343,440	56%	1.40	2.68	4.08	19%
6	40	3	16,618,810	8,368,641	50%	1.17	2.16	3.33	15%
7	30	3	16,618,810	7,056,117	42%	0.94	1.74	2.68	12%
8	20	3	16,618,810	5,182,329	31%	0.60	0.90	1.50	7%
9	20	5	16,618,810	3,542,045	21%	0.38	0.61	0.99	5%