CCSBT-ESC/2208/BGD 03
(*Previously* CCSBT-OMMP/2206/08)
(ESC Agenda item 7.2)

Development of the new CPUE abundance index using GAM for southern bluefin tuna in CCSBT

CCSBT のミナミマグロについての GAM を用いた新たな CPUE 資源豊度指数の開発

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

2021 年の ESC26 において、OM および MP に含めるミナミマグロの新たな CPUE 資源量指数を作成することになった。方法論は CCSBT と契約したコンサルタントが検討したが、日本の延縄操業データは日本漁業者の知的財産であることから公開はできない。実データへの当てはめは、データへのアクセス権を有する日本科学者が担うこととなった。本文書ではその結果と、様々な頑健性試験を行った結果を示す。新たに GAM の 2 段階のデルタログノーマル法で CPUE を標準化して、面積重みづけして作成した資源量指数を得た。資源量指数は 2006 年を最低値として 2019 年まで多くの年で増加した。2020 年と 2021 年には 2015-2017 年水準に低下した。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。最近年のデータが追加されると過去の相対値が変化することが見られた。

Summary

At ESC26 held in 2021, it was decided to develop a new CPUE abundance index of southern bluefin tuna to be used in OM and MP. The methodology is examined by a consultant contracted with CCSBT, but Japanese fishermen's longline operation data cannot be disclosed because it is the intellectual property of Japanese fishermen. The application to actual data will be carried out by Japanese scientists who have access to the data. This document summerizes the results of the base case and the results of various robustness tests. CPUE was newly standardized by GAM with two-step delta log normal method, and the abundance index with area weighting was obtained. The abundance index was the lowest in 2006 and increased in most years until 2019. In 2020 and 2021, it fell to 2015-2017 levels. 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. It was observed that the relative values of the past changed when the data of the recent year was added.

1. Introduction

Stock assessment and stock management through MP of southern bluefin tuna (*Thunnus maccoyii*) 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 olden days, stock assessment was carried out using the abundance index obtained by the generalized linear model (GLM) developed by Nishida and Tsuji (1998). Since 2007, we have been using 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). For the area weighting, we assumed two types of hypotheses, Constant Squares (CS) that the distribution time and space of fish does not change over time depending on the stock abundance, and Variable Squares (VS) that the distribution changes in time and space, and used intermediate values (W0.8 and W0.5). The abundance index obtained from the Japanese longline has 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. Thus, the index directly affects the calculation of TAC.

It was recognized that the 2018 value of the CPUE abundance index was anomalously high in ESC24 held in 2019. The report of ESC26 stated in the paragraph 90 that ”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 (Report of OMMP 11, paras 11-24 and Report of ESC 25, paras 94-100). The ESC agreed that, even though the 2018 estimate was within the bounds of the range for which the MP had been tested and the immediate implications for the current TAC recommendation were small, this technical bias needed to be addressed through the development of a CPUE standardisation method that more effectively dealt with the spatial-temporal variation in the CPUE data.”

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). A consultant hired by CCSBT will consider the methodology. However, since the Japanese fishermen's longline operation data is the intellectual property of Japanese fishermen, it cannot be disclosed. The application to actual data will be carried out by Japanese scientists who have access to the data. This document summarises the results of the base case run and the results of various robustness tests.

2. Materials and Methods

2-1. Dataset used

The dataset was made from logbook data by Japanese longline fishery, which includes the period from 1969 to the latest year (currently 2021). Following the conventional CPUE abundance index, records in CCSBT statistical area between 4 and 9 and from April to September were selected. From the logbook data, year, month, latitude (in 1 degree), longitude (in 1 degree), vessel ID (available from 1979), number of hooks used, number of fish caught of southern bluefin tuna (SBT), bigeye tuna (*T. obesus*, BET), yellowfin tuna (*T. albacares*, YFT), albacore (*T. alalunga*, ALB) and swordfish (*Xiphias gladius*, SWO) were used. At the stage of trial and error, 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)) was also included.

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 over 4 years old (age-4 plus). The age composition was first applied to the fork length composition of 50 or more individuals measured at the same month, 5 degrees longitude, and 5 degrees latitude. At this stage, 97% of the number of SBT caught

was applicable, and the ratio of age-4 plus was calculated. If the conditions for 50 or more individuals are 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 moon - 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 CCSBT. Sensitivity analysis was conducted for age-5 plus and all ages.

The following records have been deleted: 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, was also deleted.

2-2. Cluster analysis

A cluster analysis was performed to consider the target species of fishing operation. 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 delta log normal. The mgcv package of R was used. The bum function, which is suitable for large volumes of data, is used. Based on the results of the study by consultant, a binomial submodel (hereinafter referred to as BSM) and a positive catch submodel (hereinafter referred to as PCSM) are used, and gamma = 2, binomial distribution and gauss distribution are 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 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 smooths) of the interaction, the better, but the longer the calculation time (Wood, help of choose.k in mgcv). The effective degree of freedom for a model term (EDF) value is calculated by the k.check function in mgcv package, and when EDF is 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 obtained 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. It was set separately, however, for the purpose of organizing the work, and the k value of each variable in the interaction was set to the same value (i.e. k for year = 20 is 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) and laptop PC (CPU = Intel (R) Core) (TM) i7-7500U CPU @ 2.70GHz and 2.90 GHz, RAM = 16.0 GB 64 bit, Windows 10 Pro). However, the making the dataset was limited to the desktop because it could not be made because the memory was over on the laptop PC. R (R4.0.5) 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 predict value was calculated for each submodel for the dummy data set, and the product was calculated. Since the expected value is biased when the logarithmic normal distribution is restored, the predicted value is 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 dataset and method. The following sensitivity analysis was performed at the final stage.

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 is 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 is 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: Basically, age-4 plus used, but limited to age-5 plus, or all ages were tried.

Data resolution: We used shot-by-shot data, aggregated data by month, 1 degree latitude and 1 degree longitude, and aggregated data by month, 5 degrees latitude and 5 degrees longitude.

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

2 clusters: 4 clusters were the basis, but I 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 a year for *yy* years.

3. Results

3-1. Dataset used

Data from 1969 to 2021 amounted to 794,481 records. Of these, 702,481 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% and has been increasing for the last five years (Fig. 1). The CPUE of the positive catch dataset is high in the 1970s, low in the 1980s to 2000s, and high after 2010.

Similar figures are shown for other variables (month, longitude, latitude, latitude and longitude maps) (Fig. 2 and Fog. 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. 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, 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 CPUEWG to have 5 species. We also tried to include several species of marlines and butterfly kingfish, but because the catch record was limited in recent years, the number of fish caught was smaller than the 5 species, which was not considered to be the main target, there was no substantial difference from the 5 species case, 5 species were selected.

The latitudes of the four clusters differed (Fig. 7). There were no noticeable trends in year, month, longitude, number of hooks used, or hooks between floats(HBF). It is 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, a 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, a 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. As a result of trial and error, the settings were set as shown in 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 73.2% deviance, and the positive catch submodel explained 49.1%. 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 expected and residual plots, the data up to 2020 showed a small bias at the left end, which was in the north area of of NZ, but now with the addition of 2021 data, the bias has been reduced. 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 simply 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 western 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 surrounding waters. It has been confirmed that the data in the area where these large residuals are seen has almost no effect on the abundance index (examined by the CPUEWG in October 2021).

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 2019. In 2020 and 2021, it fell to 2015-2017 levels.

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 the model with some terms removed from modA2 (Table 5). On the contrary, in the model (e.g. modA2.p11) to which one three-way interaction term was added, 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 that of modB3 without certain terms (Table 6). The difference between the models in the abundance index is small (Fig. 23 and Fig. 24). However, a relatively large difference was seen in modB3.no9 excluding ti(year, lon). Also, the difference in the final year (2021) seemed to be larger than in other years.

Retrospective analysis

Figure 25 shows the results of retrospective analysis of the base case model. Figure 26 shows the results for each submodel. Looking at the figure, it was overestimated in four years (2011, 2016, 2019, 2020) compared to the results in the dataset up to 2021. On the contrary, it was underestimated in two years (2012 and 2013). It was not change in four years (2014, 2015 yen, 2017, 2018). The impact before 2010 was small. There was no tendency to be over / under biased, suggesting that it is robust against data updates.

There was a concern that it was an overestimation in 2020, so we examined it further. First, for the data up to 2020, we compared the results of the data made in 2021 and that in 2022. That is, verification of the additional effect of data for 2020. As a result, no difference was seen between the two, and it was found that the difference caused was due to the addition of the 2021 data (Fig. 27). Next, we tried to identify the time and space of 2021 that had the effect of lowering the 2020 index value. As a result of arbitral trial and error, it was found that the 2020 index value increases when the CPUE in Area 8 in July and August 2021 is artificially lowered as an attempt. This time and space may not be the only influence, but it seems to be the main factor. By referring to the data of later years, it can be seen that the abundance index of the previous year is affected.

Selection of k

For BSM, we saw 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 BSM, we saw 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). There were no noticeable changes in these

data sets up to 2020. The fit may have changed as the number of years increased by 2021. It may 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 index 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. 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. 30S was excluded from the data set 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 2012-2014 and slightly higher in 2018-2021, but the difference is small.

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, the values for 2019-2021 were slightly higher, but no major changes have occurred. For all ages, the values for 1993-1994 and 2007-2008 were high, and the values for 2010-2011 and 2015-2017 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 result with the aggregated data by month in 5 degrees x 5 degrees, and Fig. 37 shows the comparison result 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 property 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.

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 2018-2021 (Fig. 39). When retrospective analysis was performed, the 1-degree model was less robust due to the bias of one-sided underestimation since 2014 (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 showed similar trajectories for the two clusters, although the fluctuations were larger (Fig. 46). 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 2021 was rising 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 2018-2019 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.

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 (Fig. 51).

4. Discussion

We were able to create a robust abundance index using GAM. Robustness was also shown in various sensitivity analyses. In the future, when the number of years increases, it may be better to increase the

value of k related to the year.

It is a concern that the 2020 index value declined when the 2021 data was added. It is uncomfortable to increase or decrease the abundance index in the years after setting TAC by MP (such cases have not been considered in MP testing), and stakeholders may question it. But standardization should be recognized as such. As with GLM 2018, as with GAM series 2020, later data additions will have an impact. Although the robustness was verified by retrospective analysis as much as possible, it was suggested that the behavior in actual data may exceed that. These uncertainties, for which the CTP seems to have already taken into account quite a bit, will be verified through OMMP meetings and ESC.

5. Acknowledgement

Thanks to Simon Hoyle for his consultant and significant contribution to method selection and technical advice to us. We would like to thank the members of the CPUE Working Group who contributed to the development of the index in constructive discussions and Jim Ianelli who led it.

6. Reffenrences

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.

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.

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.

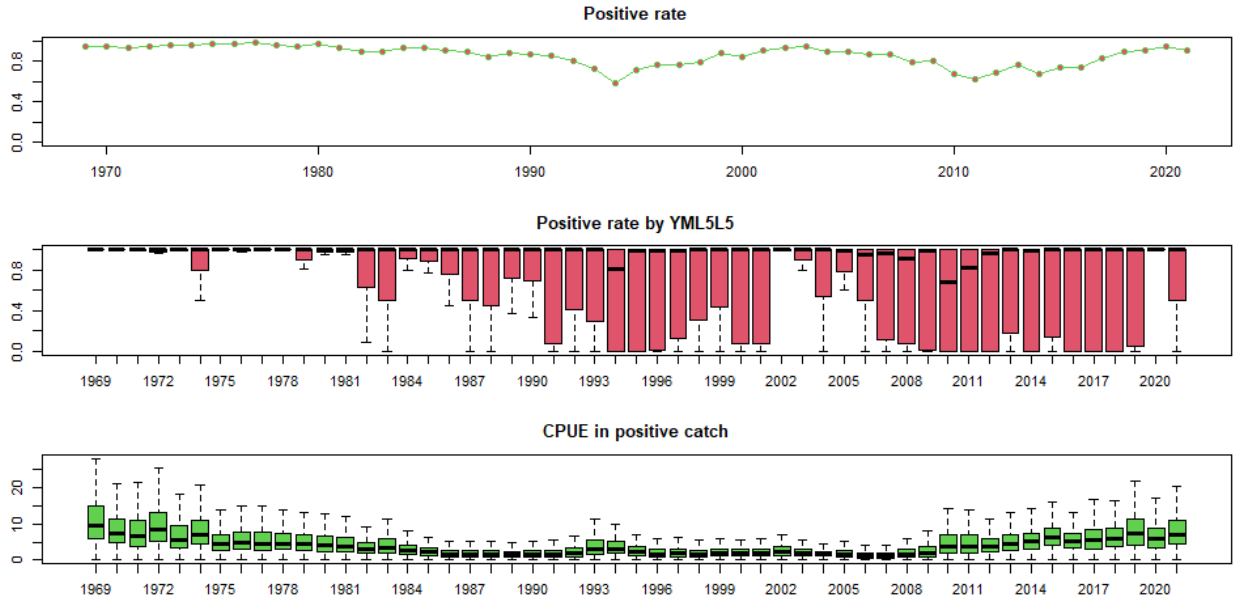

Fig. 1. Nominal value of positive catch rate and CPUE by year.
Upper panel is the positive rate which is the total number of positive catch operations / the total number of all records. Middle panels is boxplot based on the positive catch rate by year, month, 5 degree latitude and 5 degree longitude.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

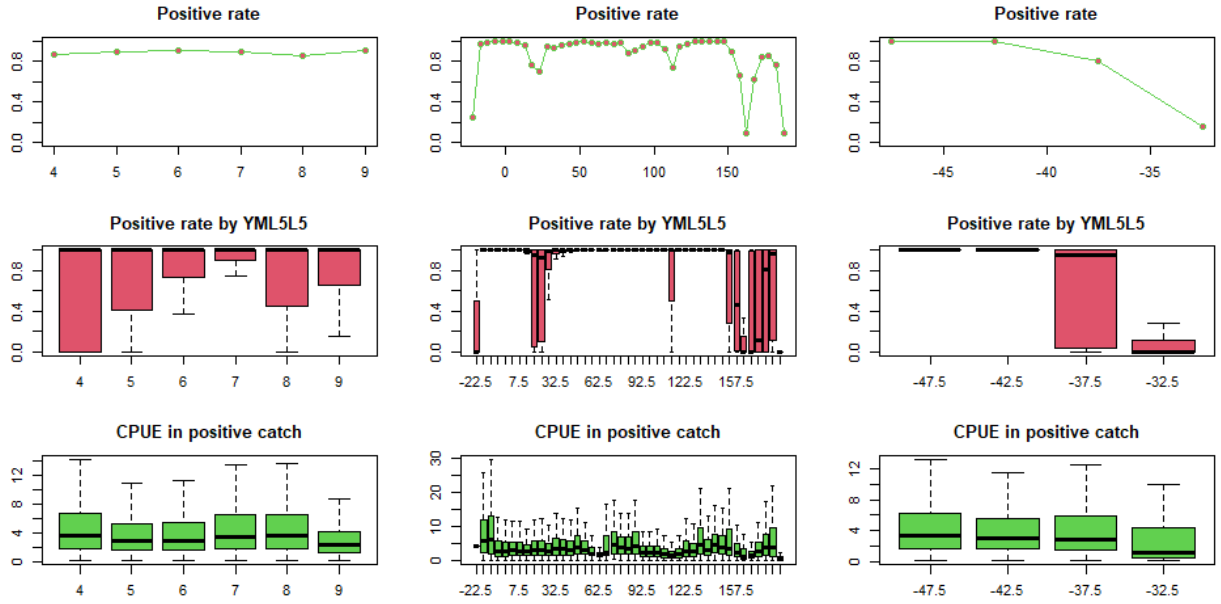

Fig. 2. Nominal value of positive catch rate and CPUE by month, longitude and latitude.
Upper panel is the positive rate which is the total number of positive catch operations / the total number of all records. Middle panels is boxplot based on the positive catch rate by year, month, 5 degree latitude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

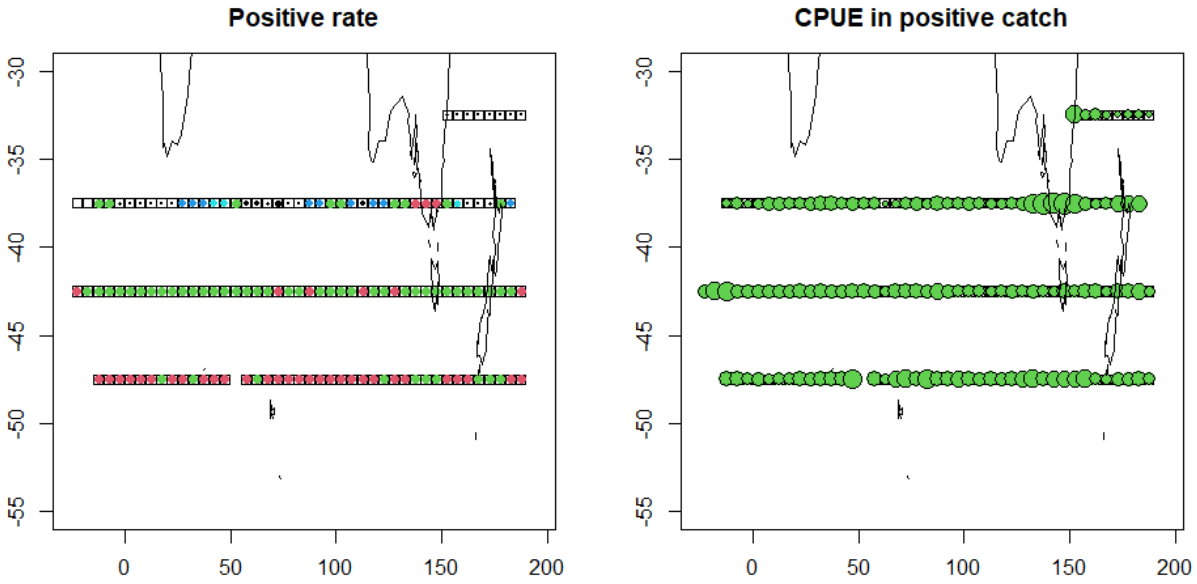

Fig. 3. Nominal value of positive catch rate and CPUE in map.
Left panel is the positive rate. Right panel is CPUE in positive catch records. Red is the higher value, followed by green, blue and white in the positive catch rate panel.

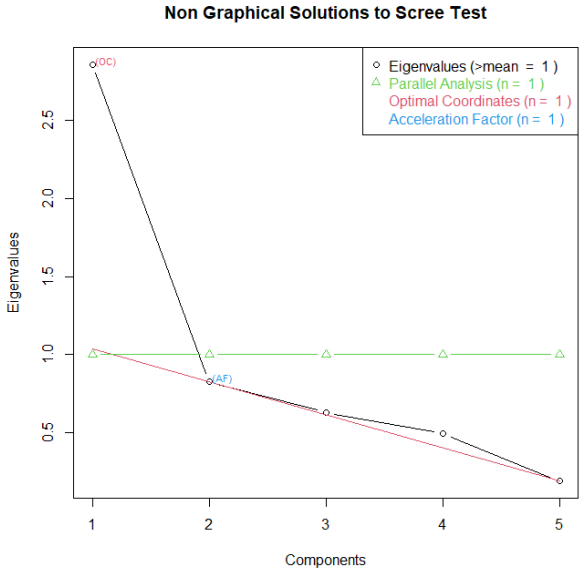

Fig. 4. Eigen values for the number of components in cluster analysis.

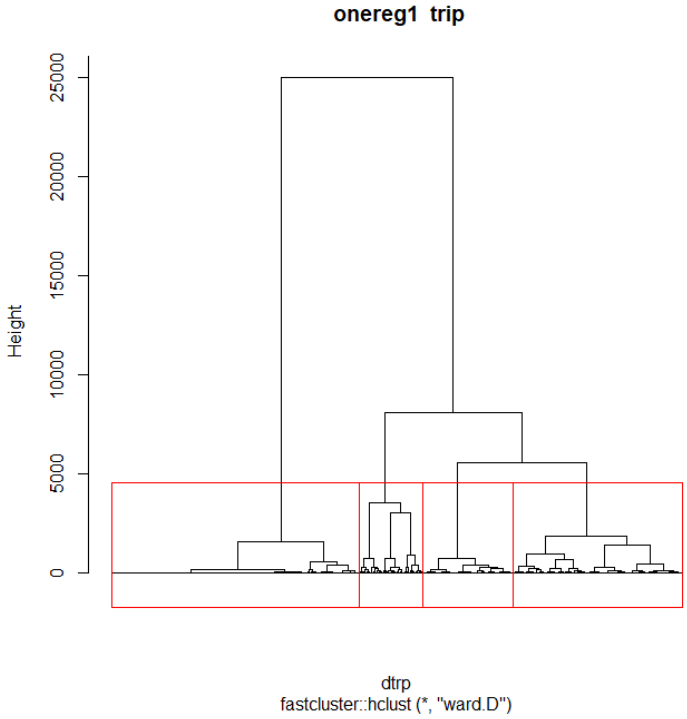

Fig. 5. Dendrogram of the cluster analysis.

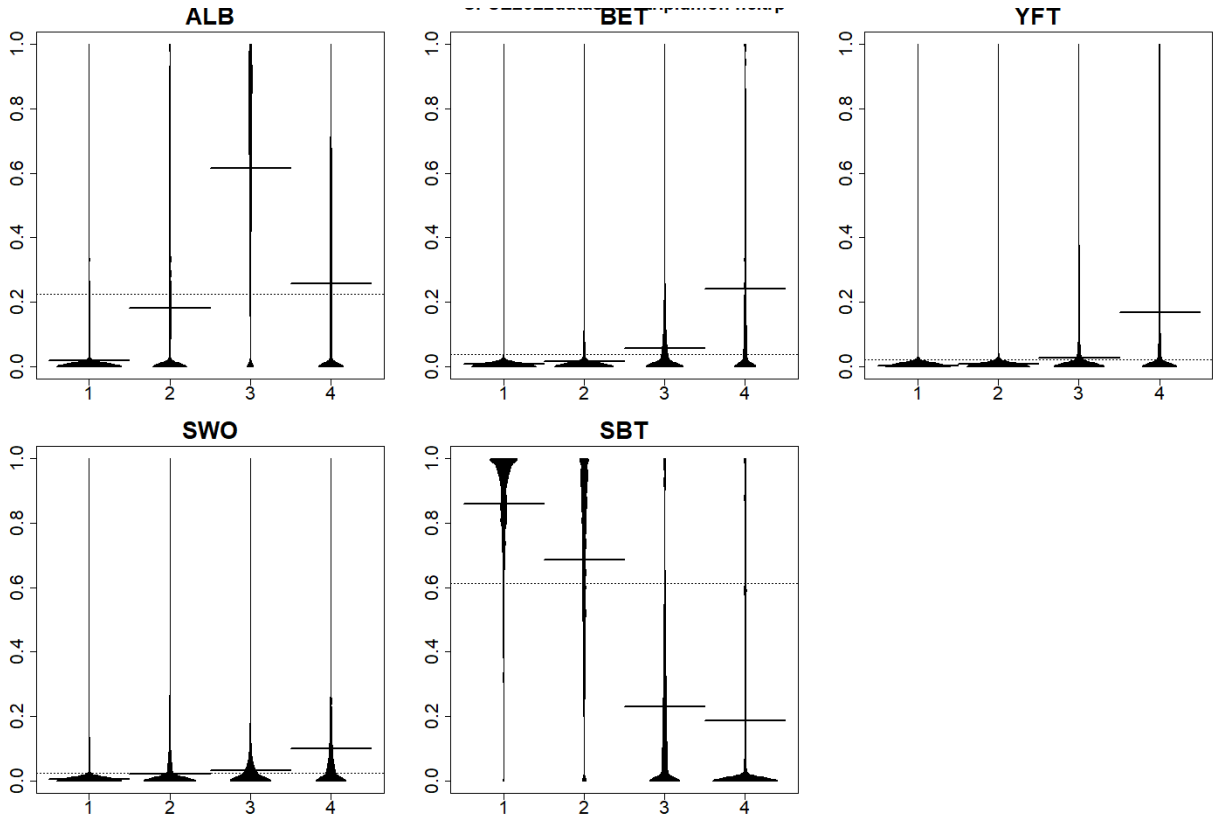

Fig. 6. Occurrence by species in each group in cluster analysis.
ALB is albacore, BET is bigeye tuna, YFT is yellowfin tuna, SWO is swordfish and SBT is southern bluefin tuna.

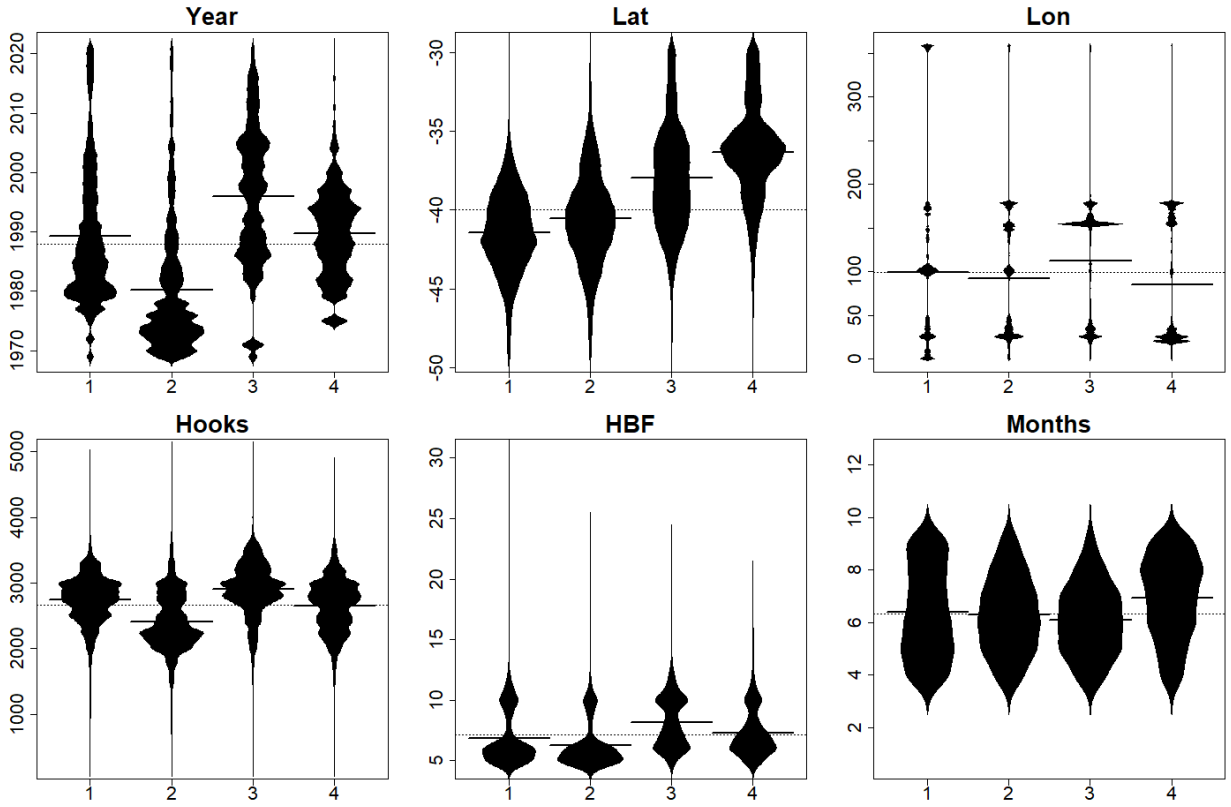

Fig. 7. Occurrence by variables of each group in the cluster analysis.

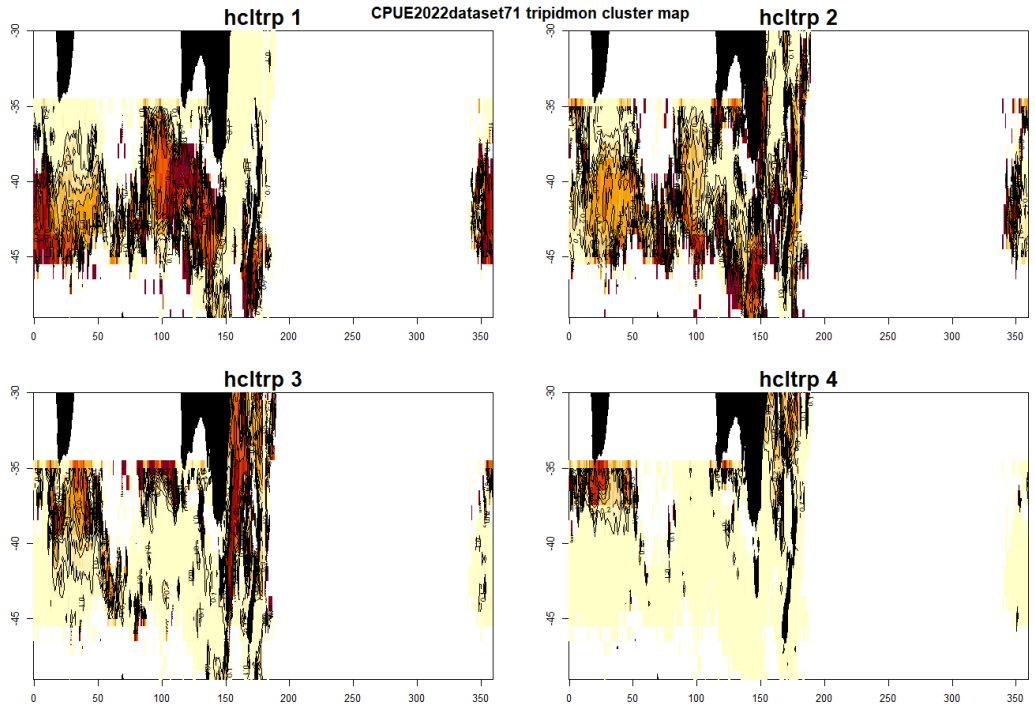

Fig. 8. Occurrence on map by group in the cluster analysis.

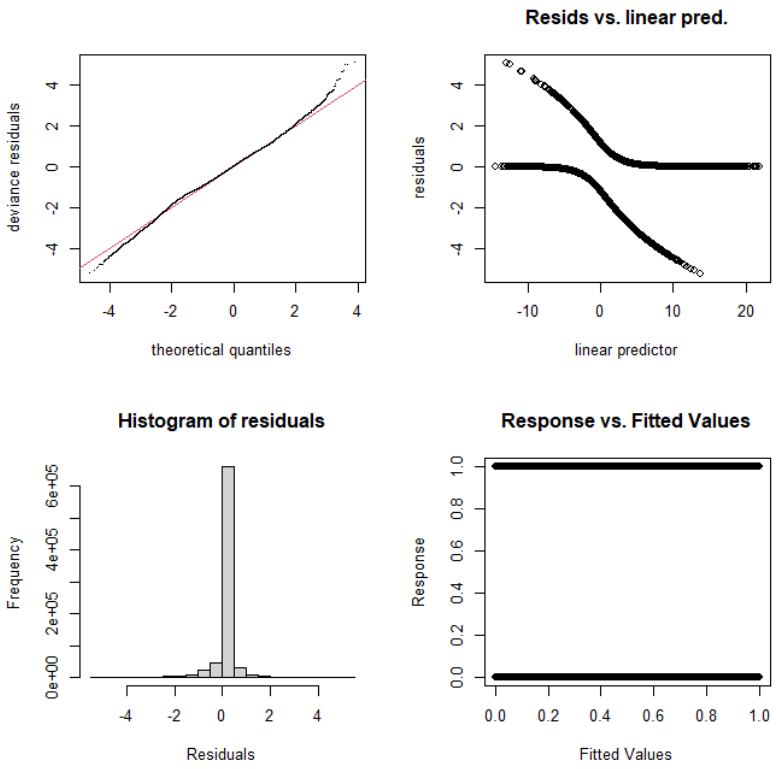

Fig. 9. Diagnostic plots for the binomial sub-model.

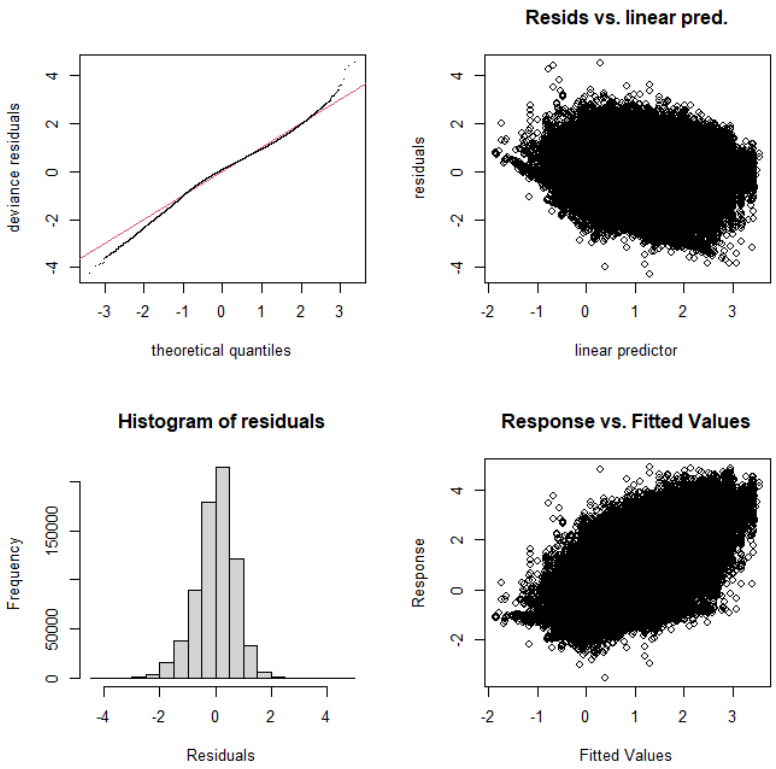

Fig. 10. Diagnostic plots for the positive catch sub-model.

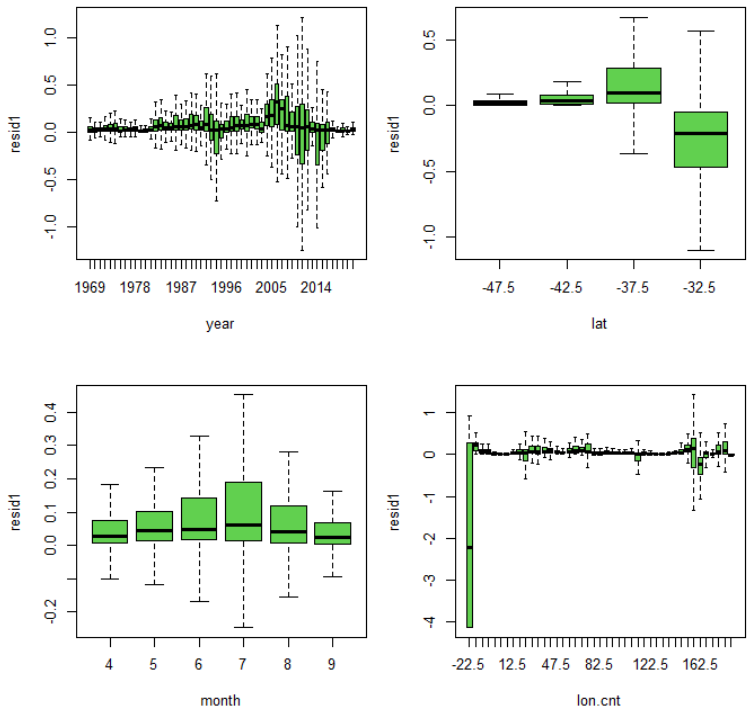

Fig. 11. Residuals by variable in the binomial sub-model.

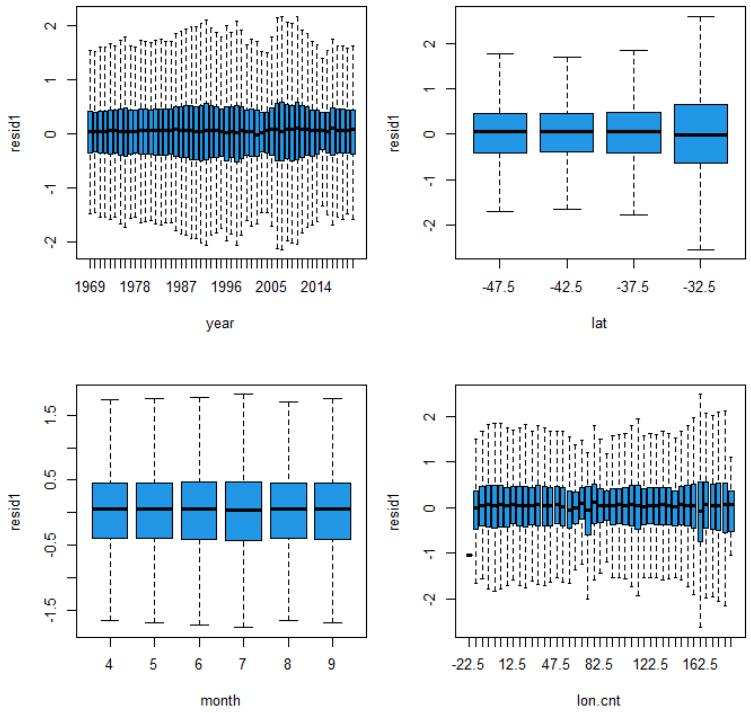

Fig. 12. Residuals by variable in the positive catch sub-model.

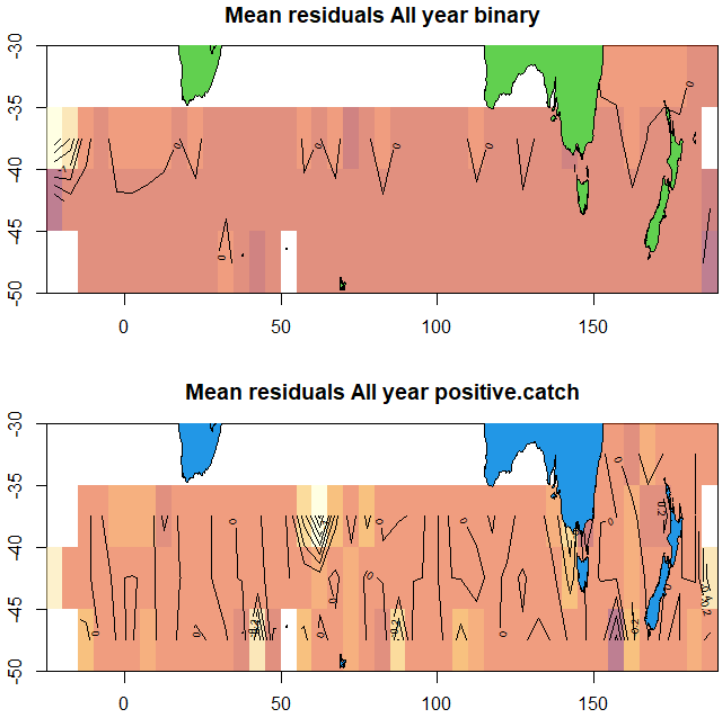

Fig. 13. Residual on maps for both sub-models.

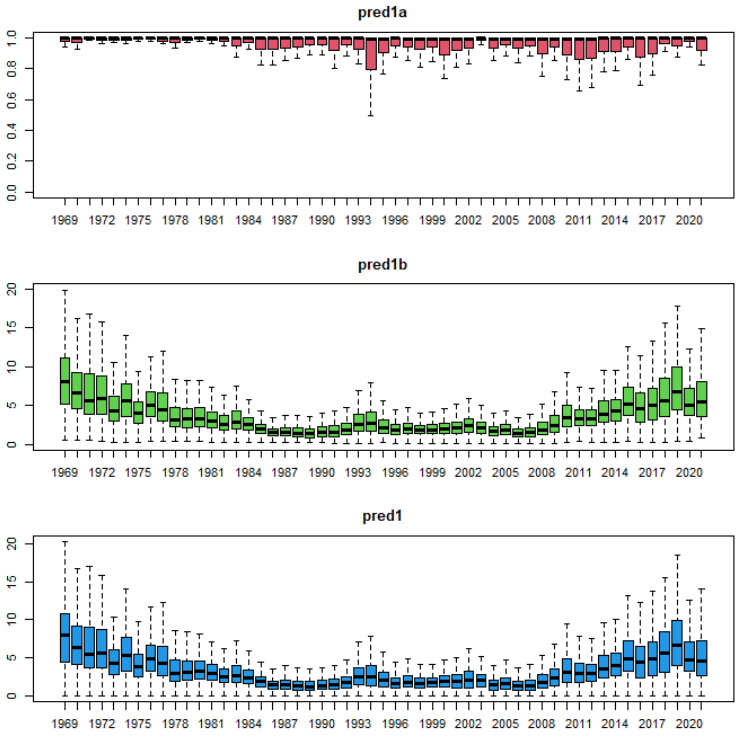

Fig. 14. Predicted value by year.

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.

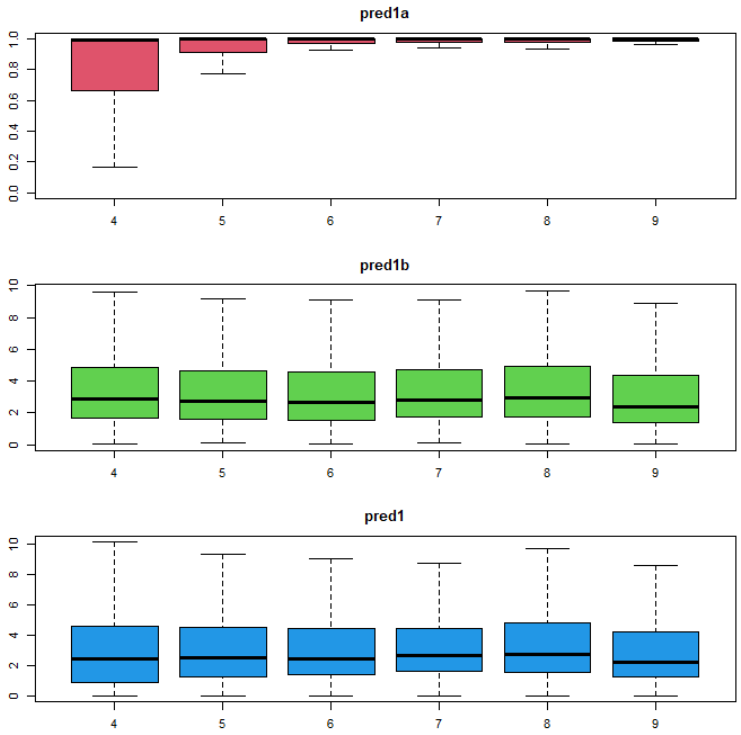

Fig. 15. Predicted value by month.
See Fig. 14.

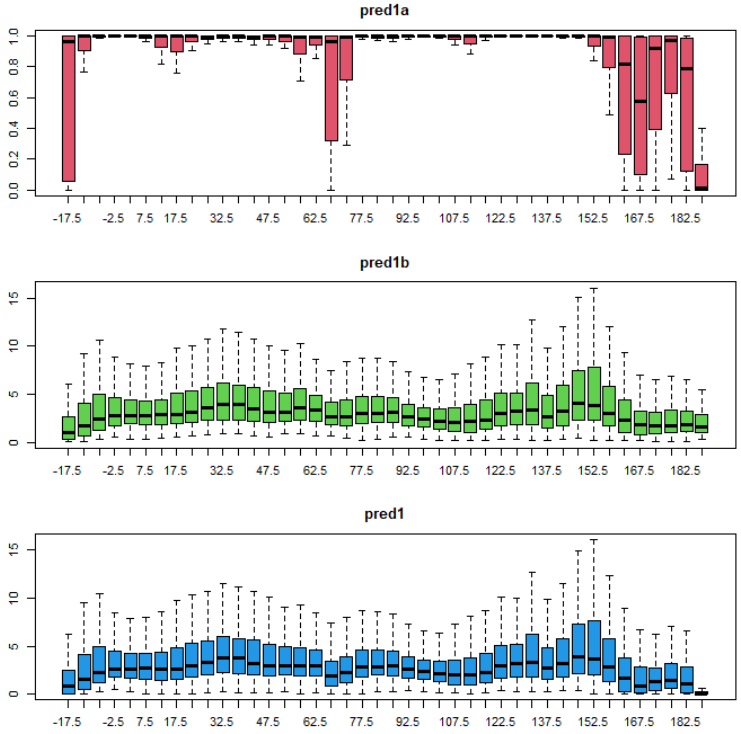

Fig. 16. Predicted value by longitude.
See Fig. 14.

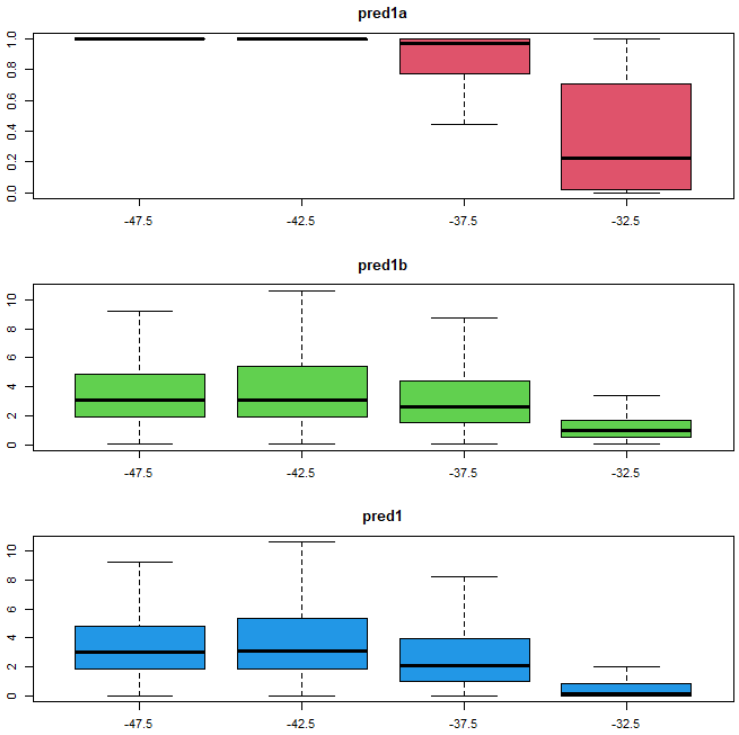

Fig. 17. Predicted value by latitude.
See Fig. 14.

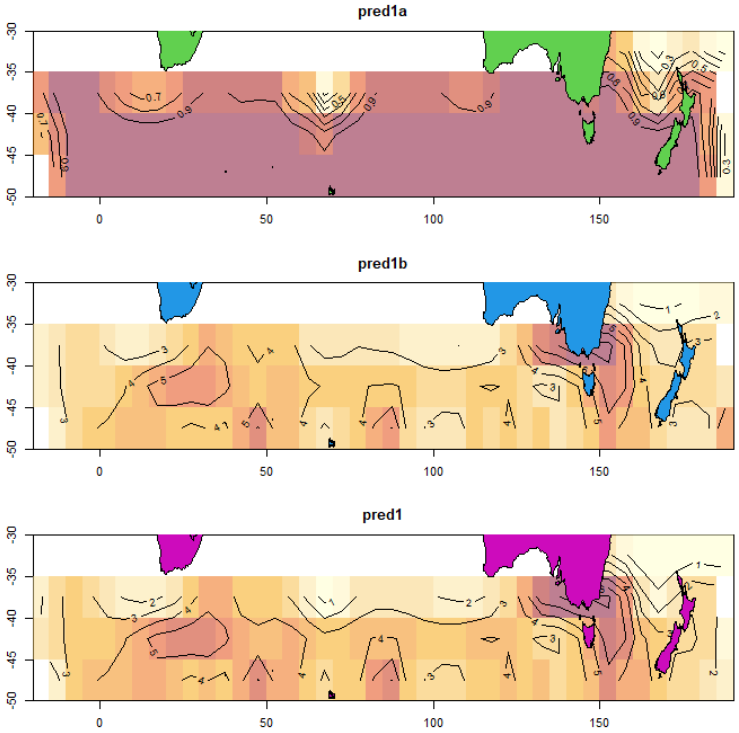

Fig. 18. Predicted value on map.
See Fig. 14.

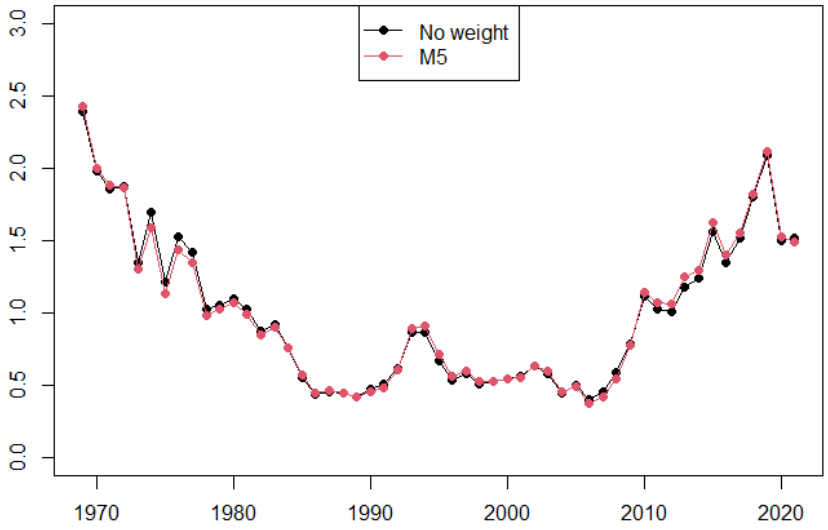

Fig. 19. Comparison of area weighted abundance indices.
Red (M5) is area weighted abundance index which taking into account that the longitude length change over latitude and the number of 1x1 degree squares ever fished in a 5x5 degrees square. Black is the abundance index which weighting was not considered.

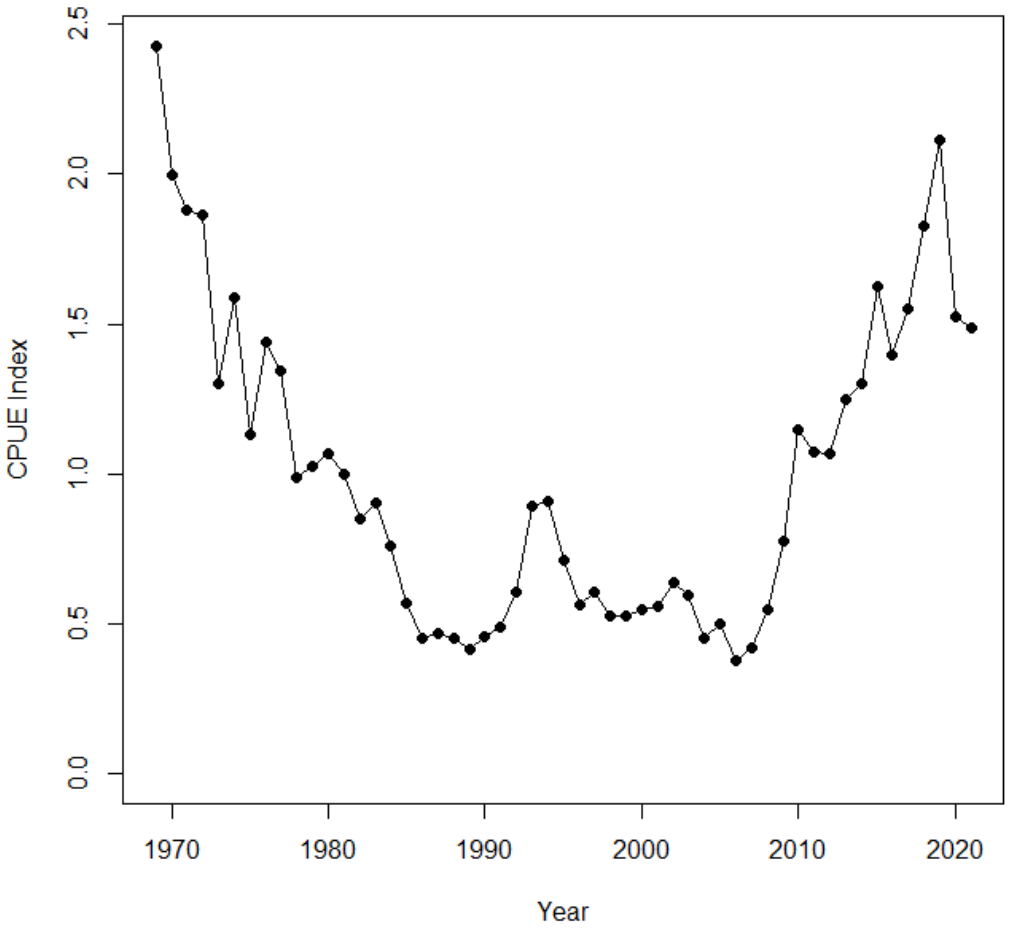

Fig. 20. Abundance index for the base case.

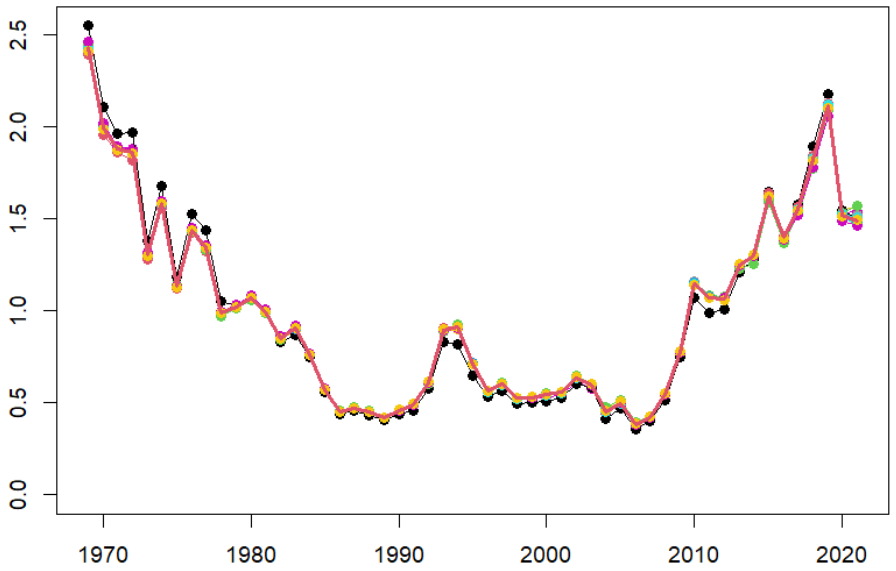

Fig. 21. Sensitivity analysis of model selection in the binomial sub-model for all runs.

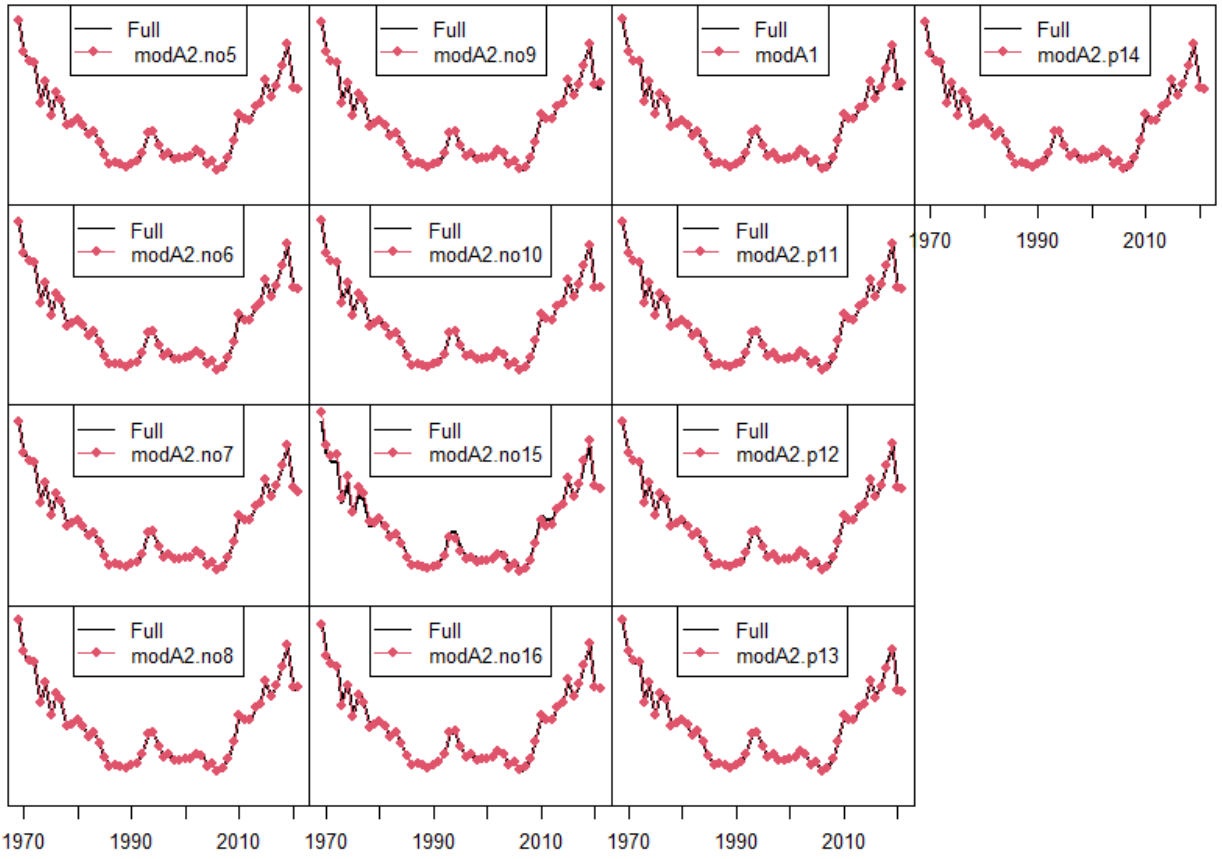

Fig. 22. Sensitivity analysis of model selection in the binomial sub-model for each run.

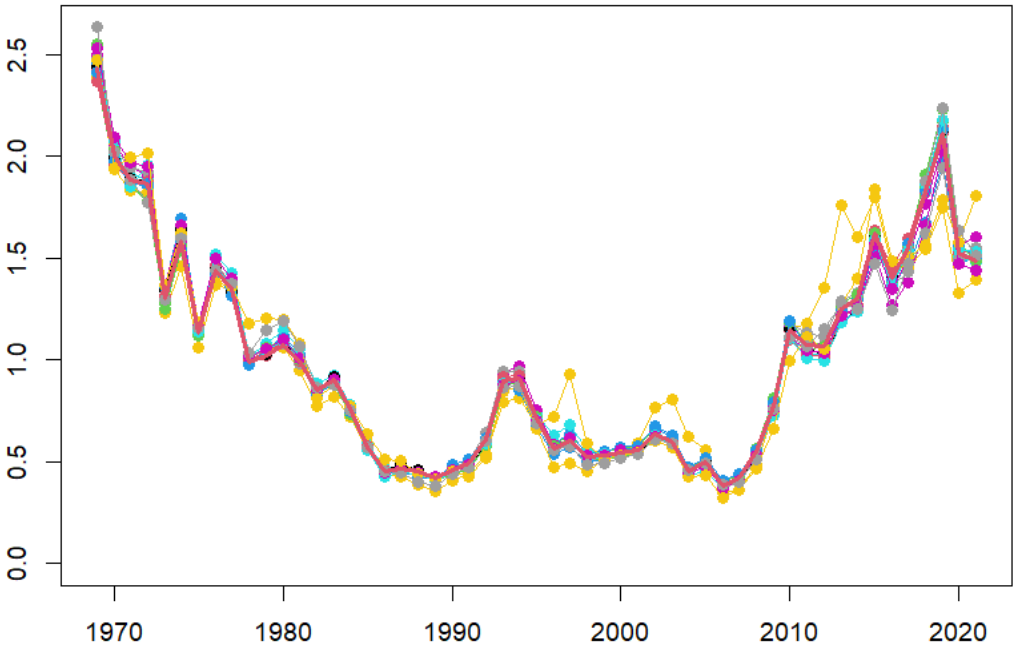

Fig. 23. Sensitivity analysis of model selection in the positive catch sub-model for all runs.

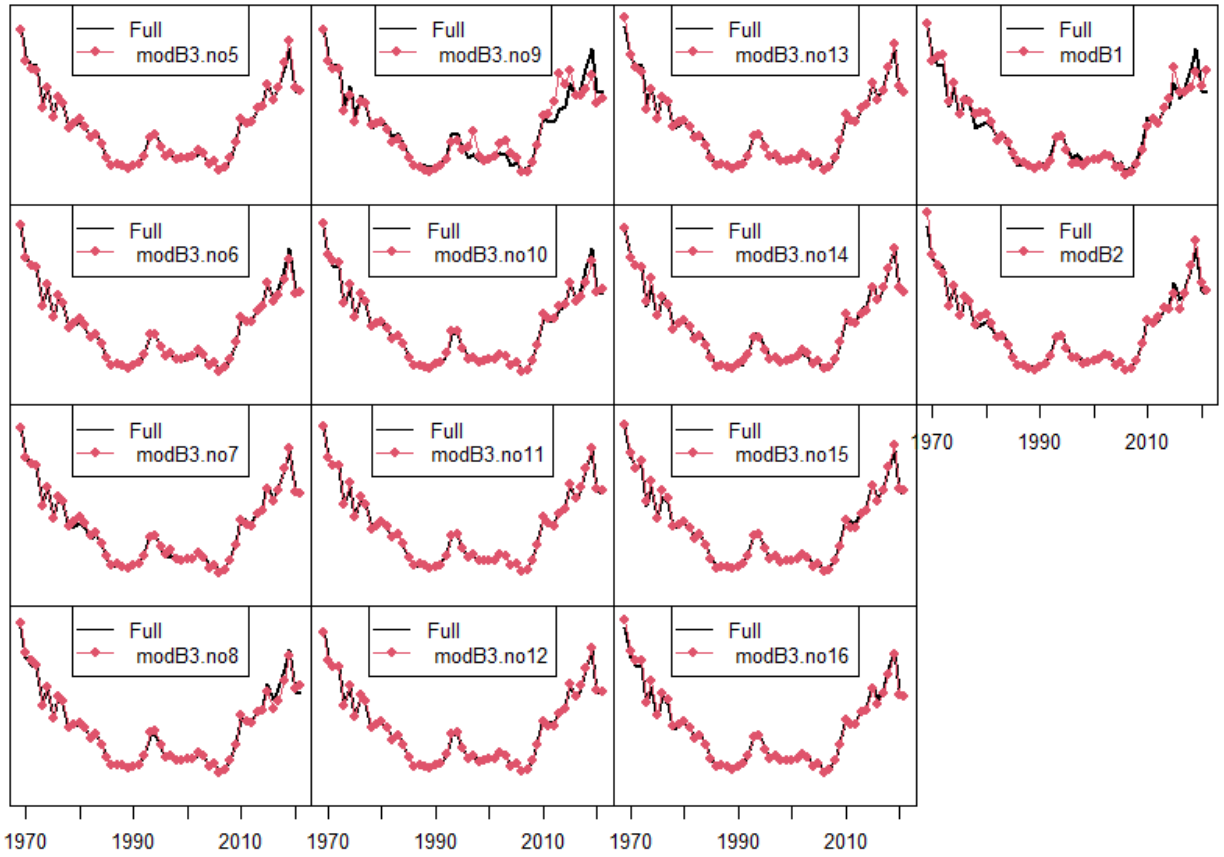

Fig. 24. Sensitivity analysis of model selection in the positive catch sub-model for each run.

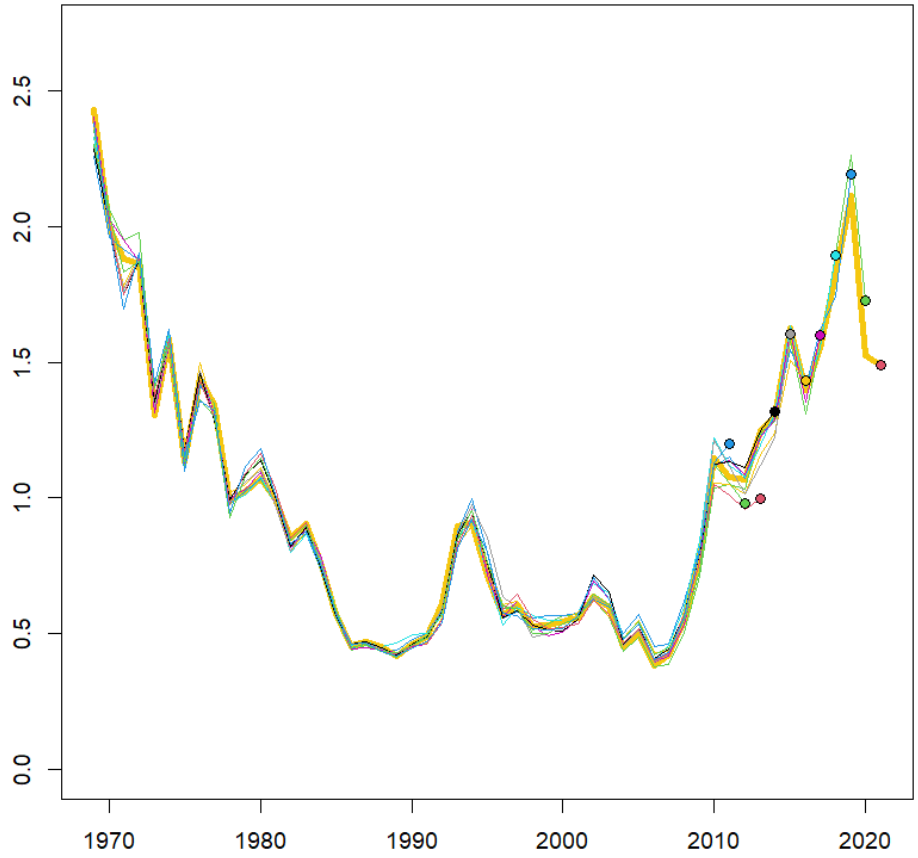

Fig. 25. Retrospective analysis for the base case model.

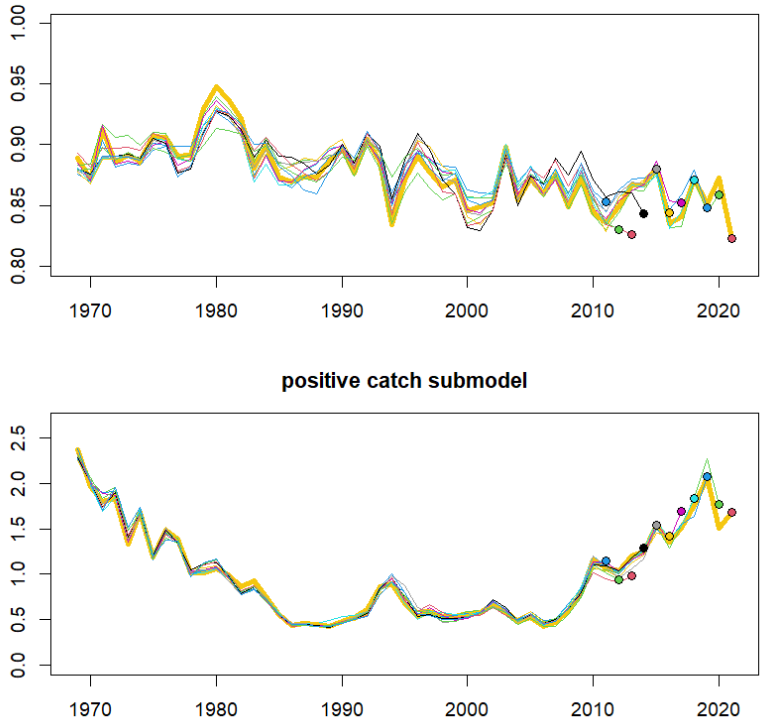

Fig. 26. Retrospective analysis for the base case model by sub-model.
Upper panel is by binomial submodel and lower panels is by positive catch submodel.

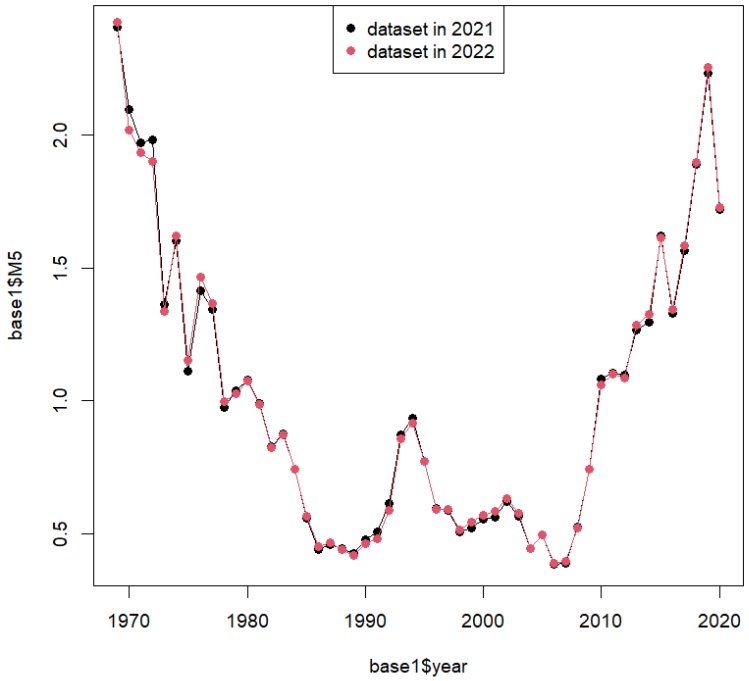

Fig. 27. Two indices from datasets up to 2020 by the datasets made in 2021 (black) and 2022 (red).

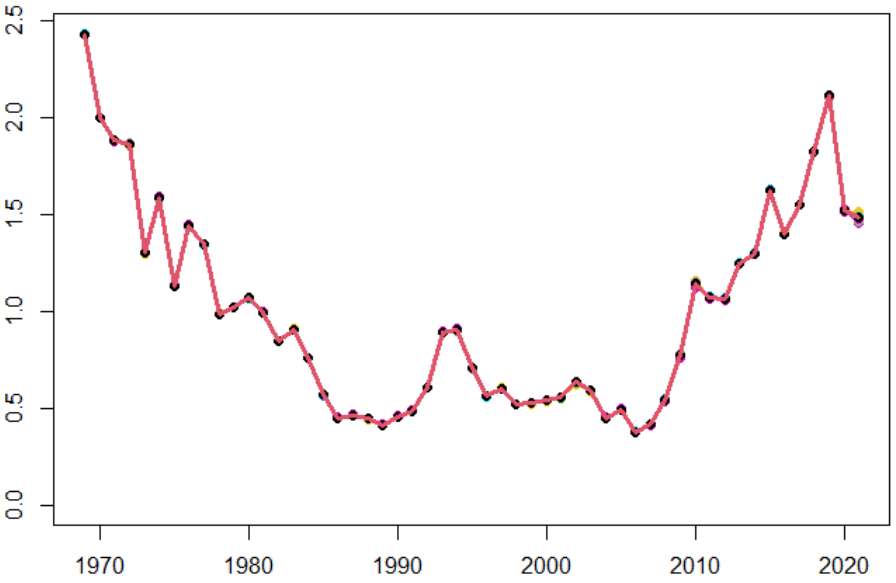

Fig. 28. Sensitivity analysis of k-value in the binomial sub-model for all runs.

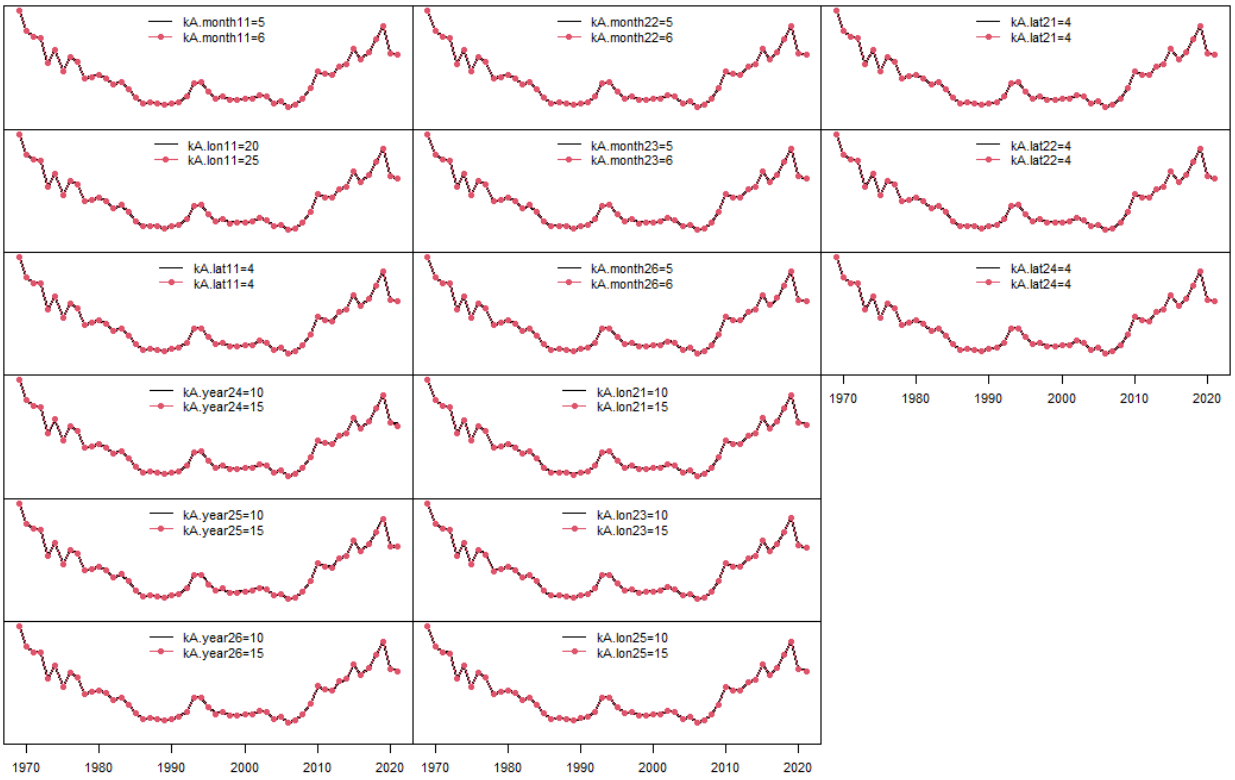

Fig. 29. Sensitivity analysis of k-value in the binomial sub-model for each of run.

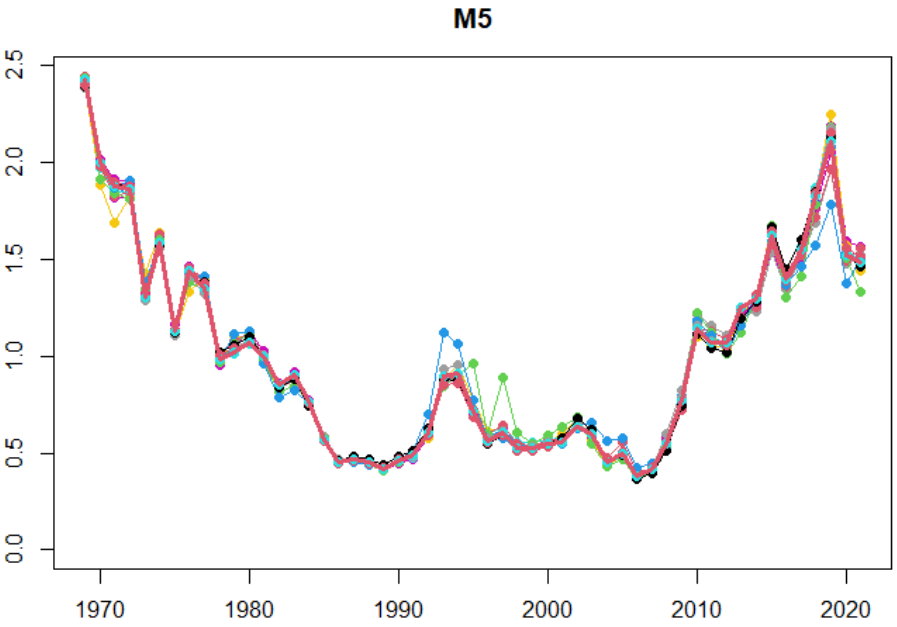

Fig. 30. Sensitivity analysis of k-value in the positive catch sub-model for all runs.

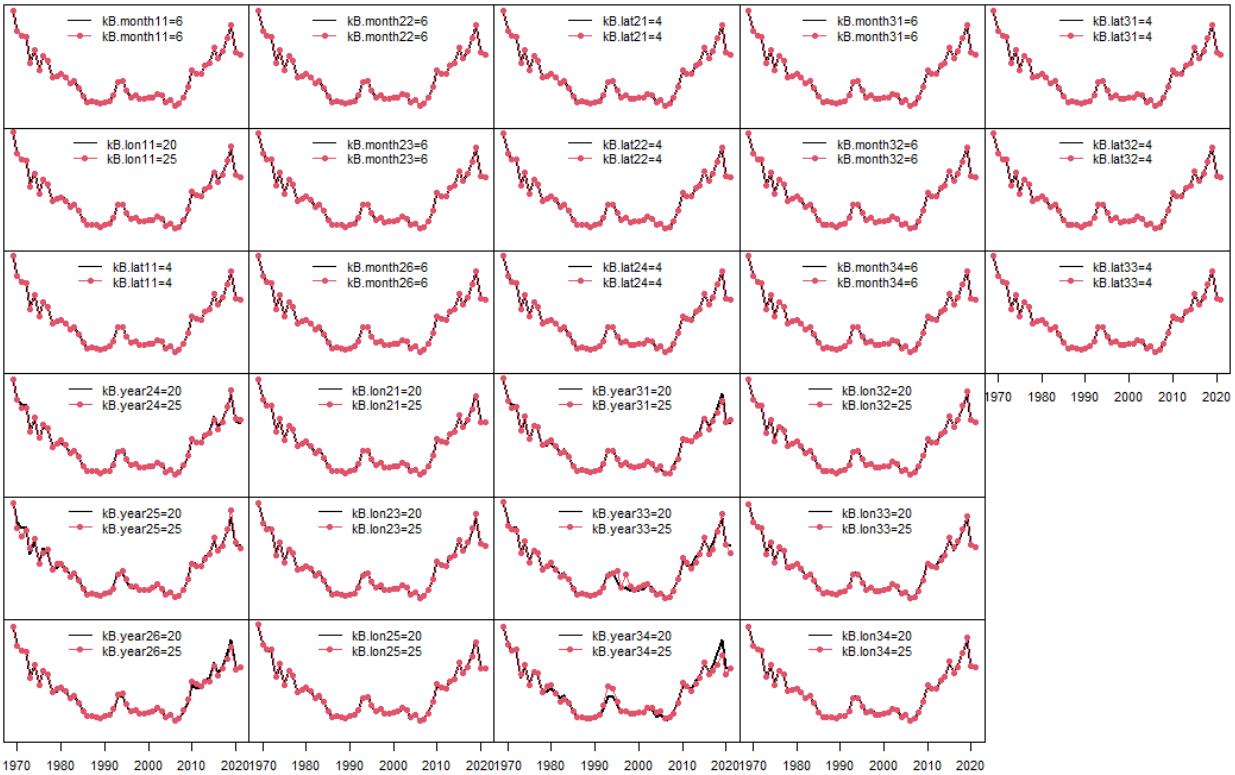

Fig. 31. Sensitivity analysis of k-value in the positive catch sub-model for each of run.

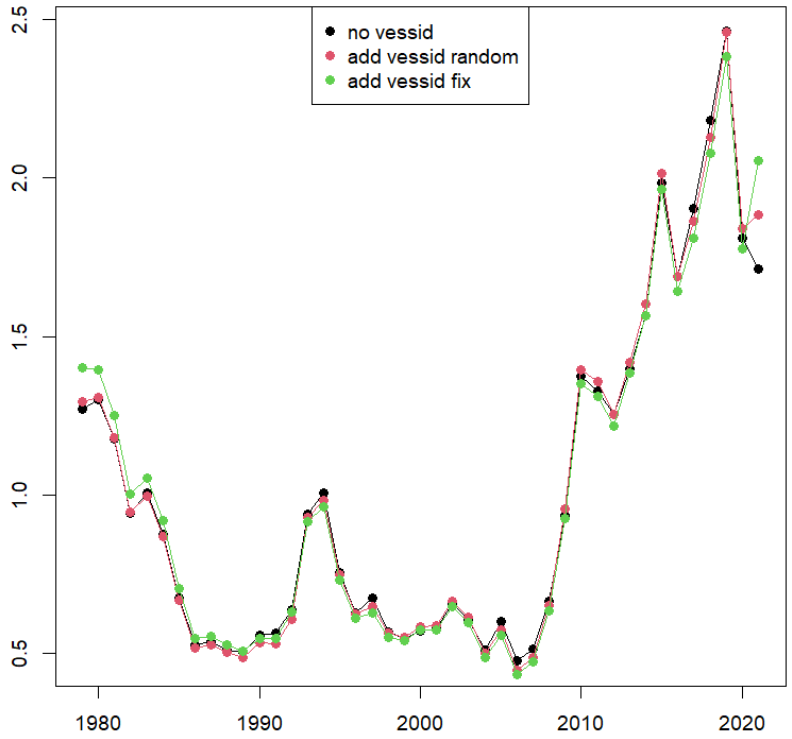

Fig. 32. Sensitivity analysis for the effect of vessel ID.

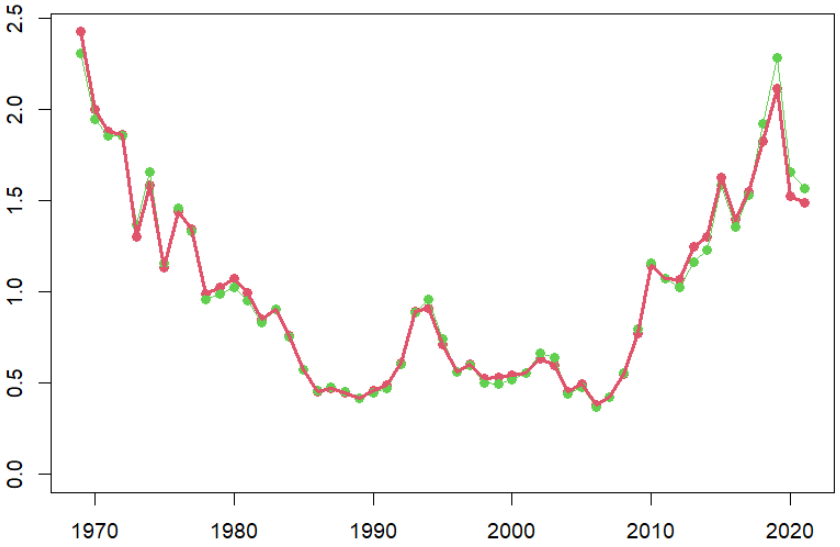

Fig. 33. Sensitivity analysis for the effect of eliminating 30S from the data.
Red is the base case, and green is the sensitivity run (eliminate 30S).

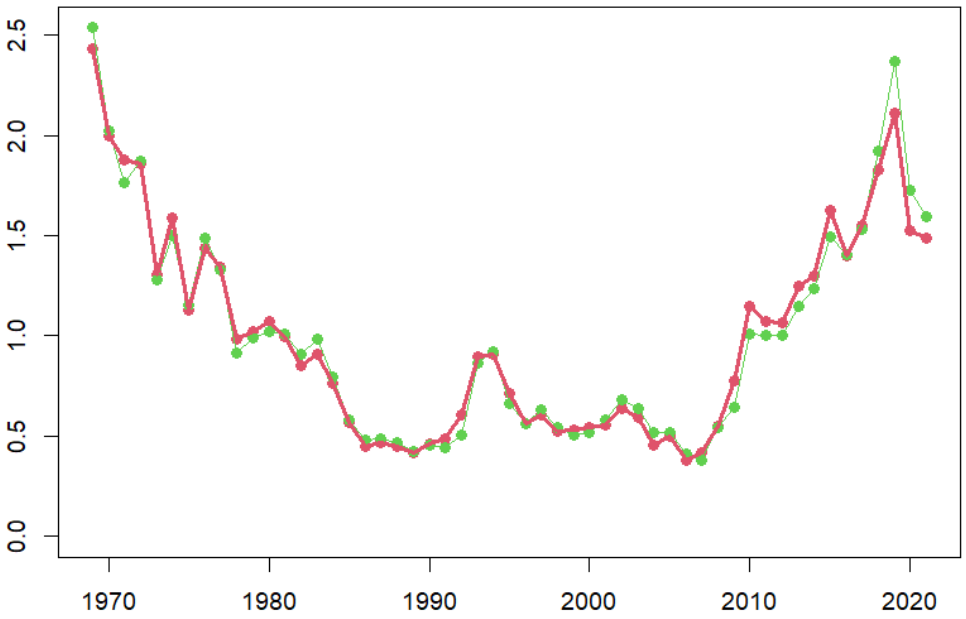

Fig. 34. Sensitivity analysis for the effect of age-5 plus instead of age-4 plus.
Red is the base case, and green is the sensitivity run (age-5 plus).

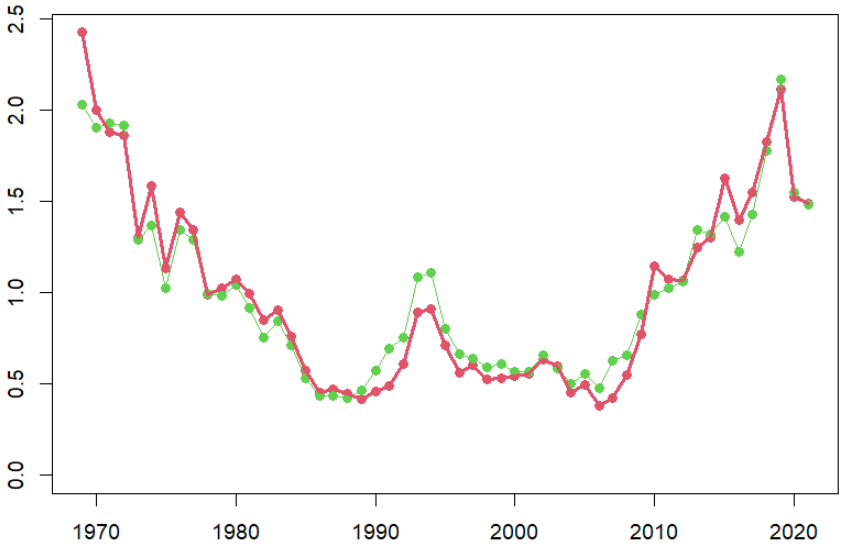

Fig. 35. Sensitivity analysis for the effect of of all ages instead of age-4 plus.
Red is the base case, and green is the sensitivity run (all ages).

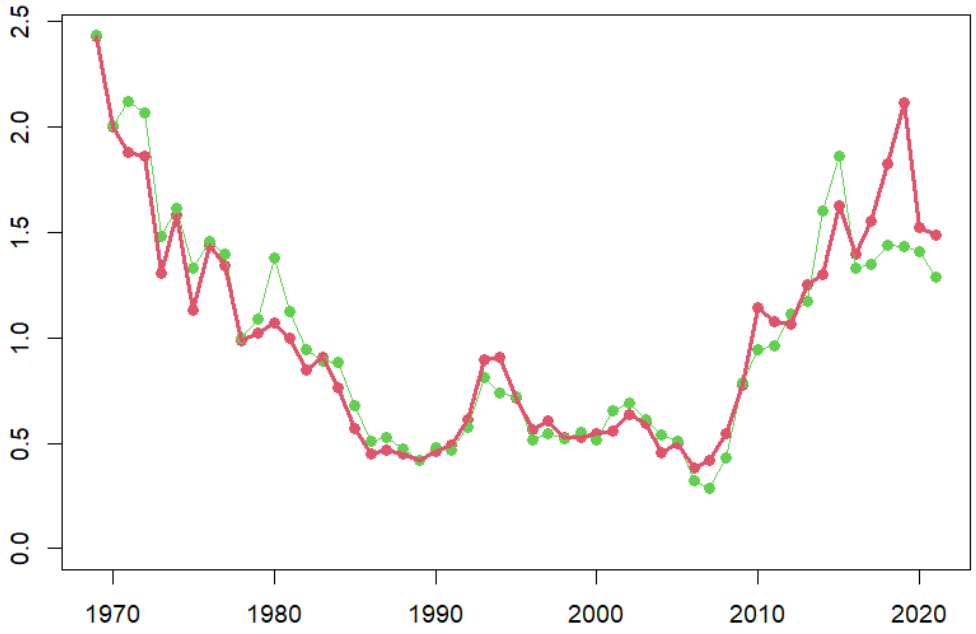

Fig. 36. Sensitivity analysis for the effect of data resolution in 5 degree.
Red is the base case, and green is the sensitivity run (aggregated by month and 5 degree latitude and 5 degree longitude).

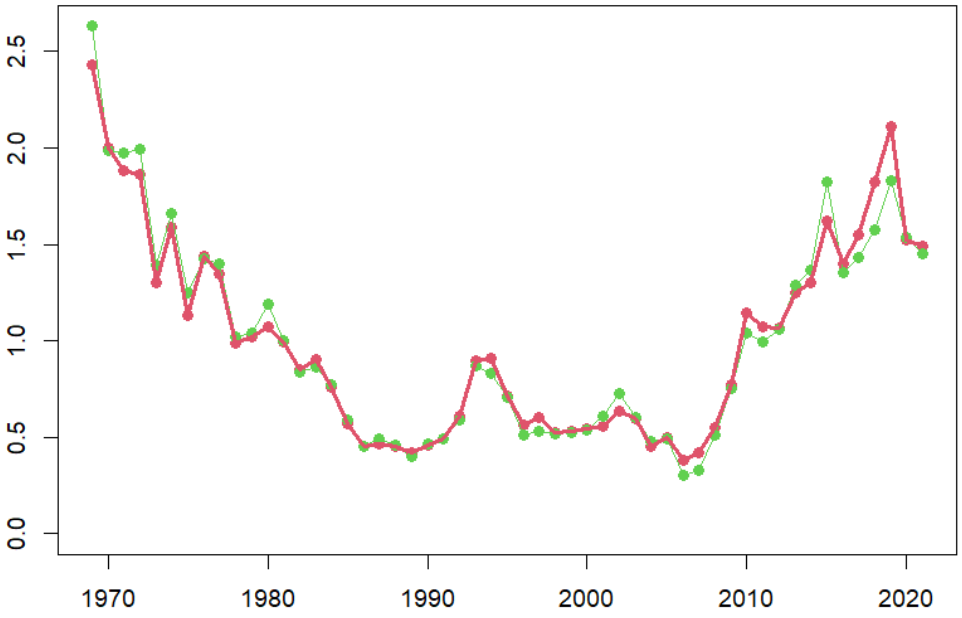

Fig. 37. Sensitivity analysis for the effect of data resolution in 1 degree.
Red is the base case, and green is the sensitivity run (aggregated by month and 1 degree latitude and 1 degree longitude).

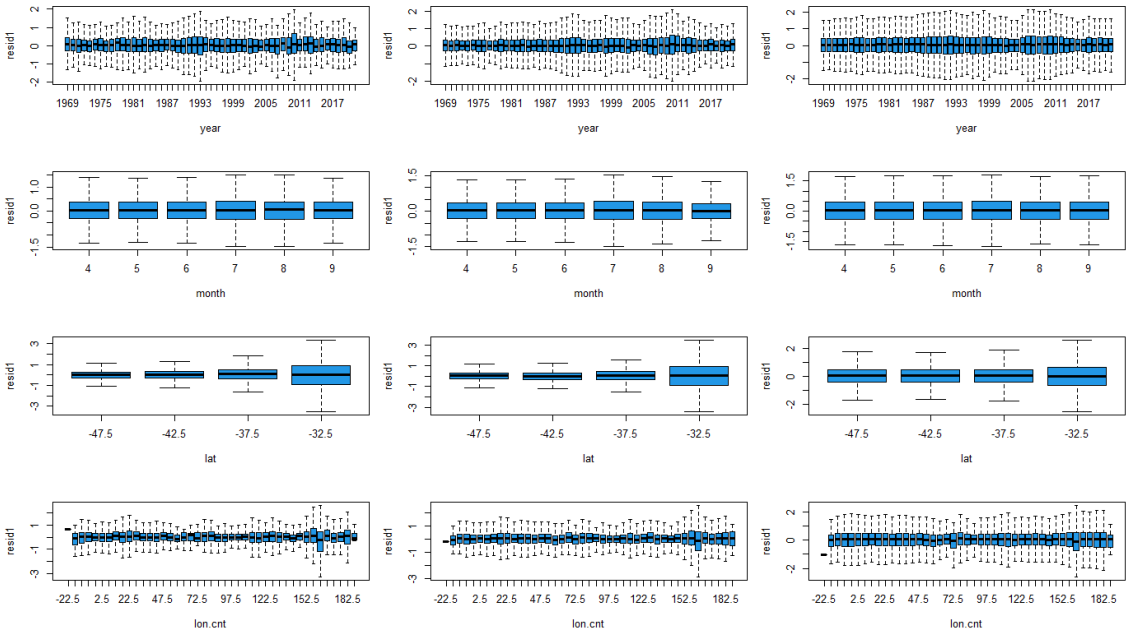

Fig. 38. Sensitivity analysis for the effect of data aggregation by residuals.
Left is 5 degree, middle is 1 degree and right is shot-by-shot data.

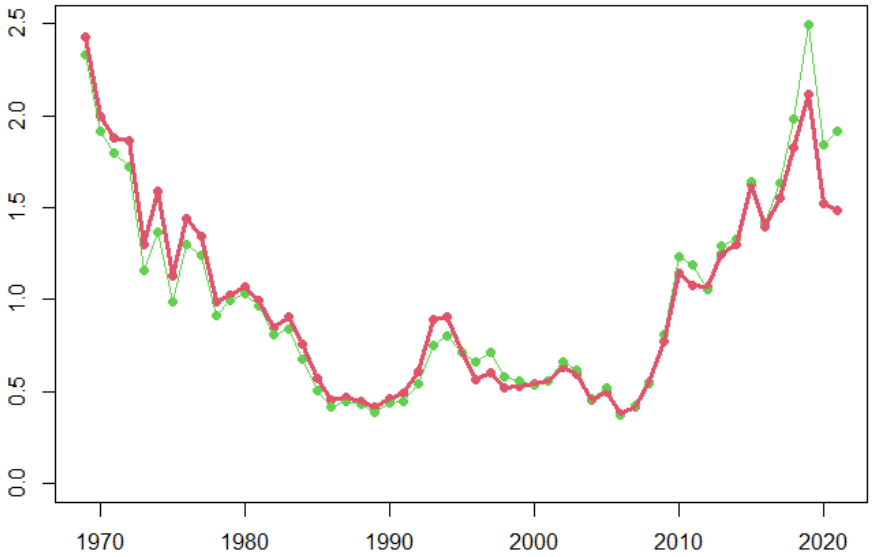

Fig. 39. Sensitivity analysis for the effect of resolution in model.
Red is the base case, where 5 degree for latitude and longitude, and green is the sensitivity run (1 degree in the model).

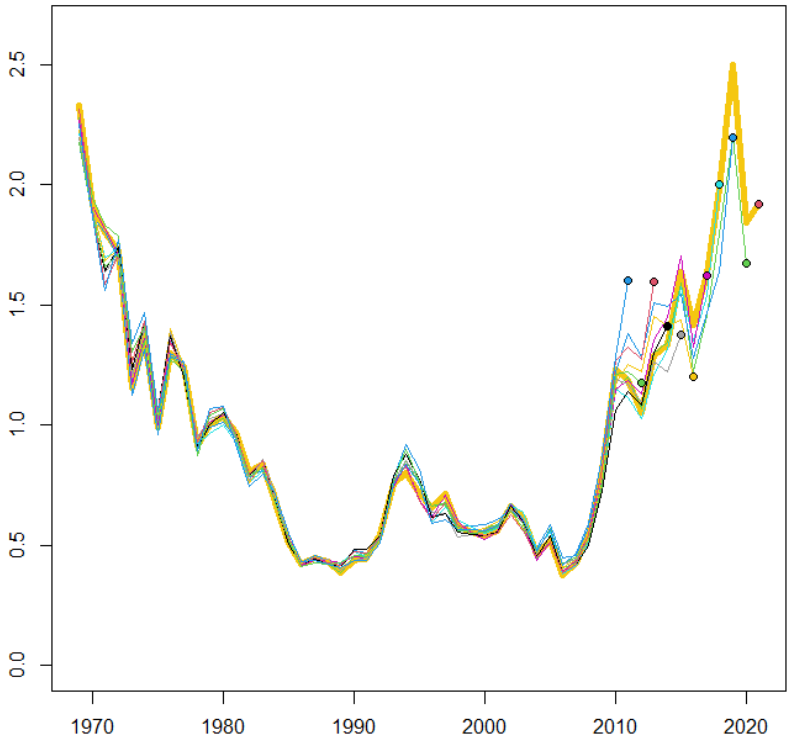

Fig. 40. Retrospective analysis in the sensitivity analysis of 1 degree model.

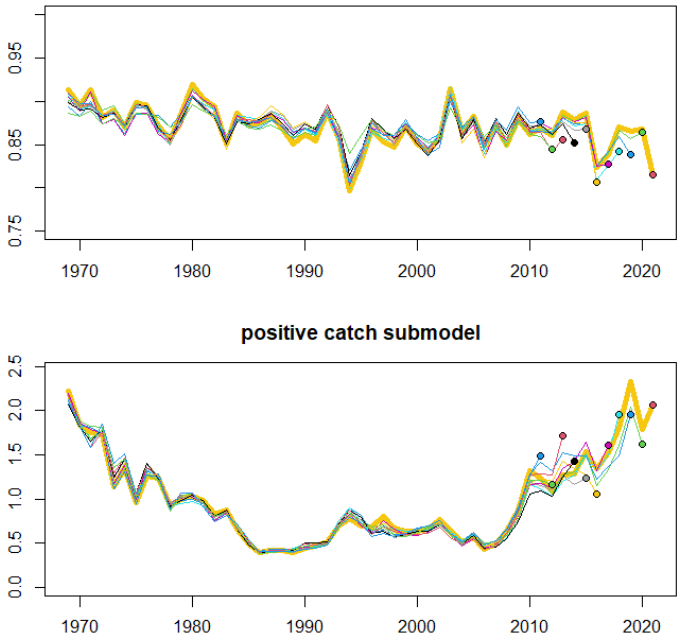

Fig. 41. Retrospective analysis in the sensitivity analysis of 1 degree model by sub-model.
Upper panel is by binomial submodel and lower panels is by positive catch submodel.

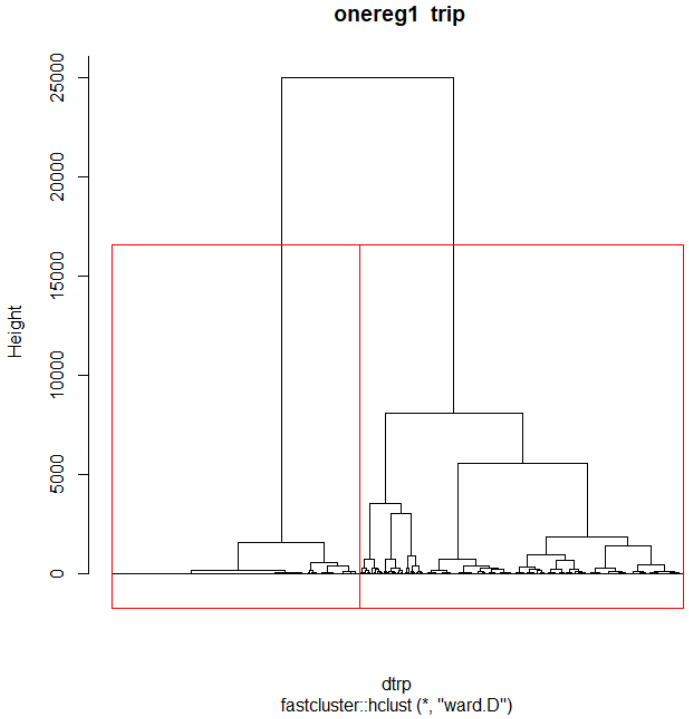

Fig. 42. Dendrogram in the cluster analysis of sensitivity analysis for the effect of 2 clusters.

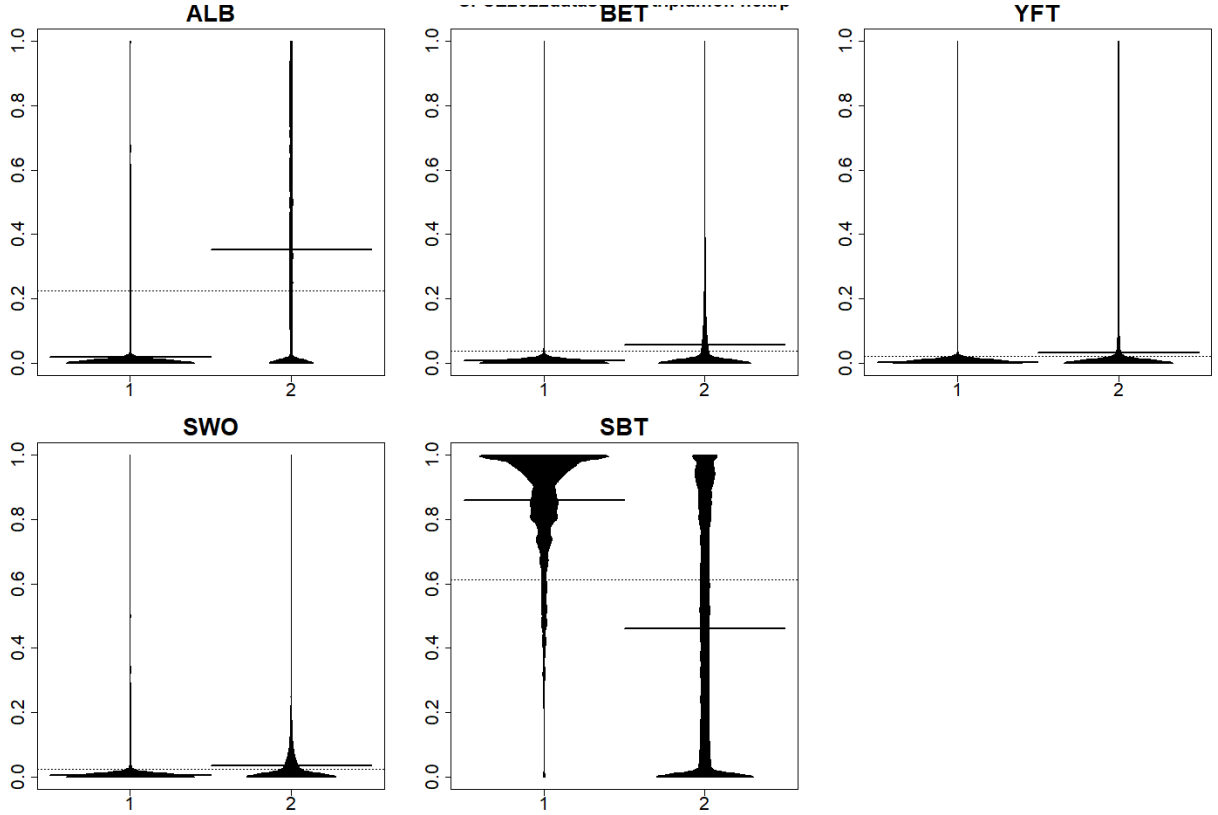

Fig. 43. Occurrence by species in the group in the 2 cluster analysis as a sensitivity analysis.

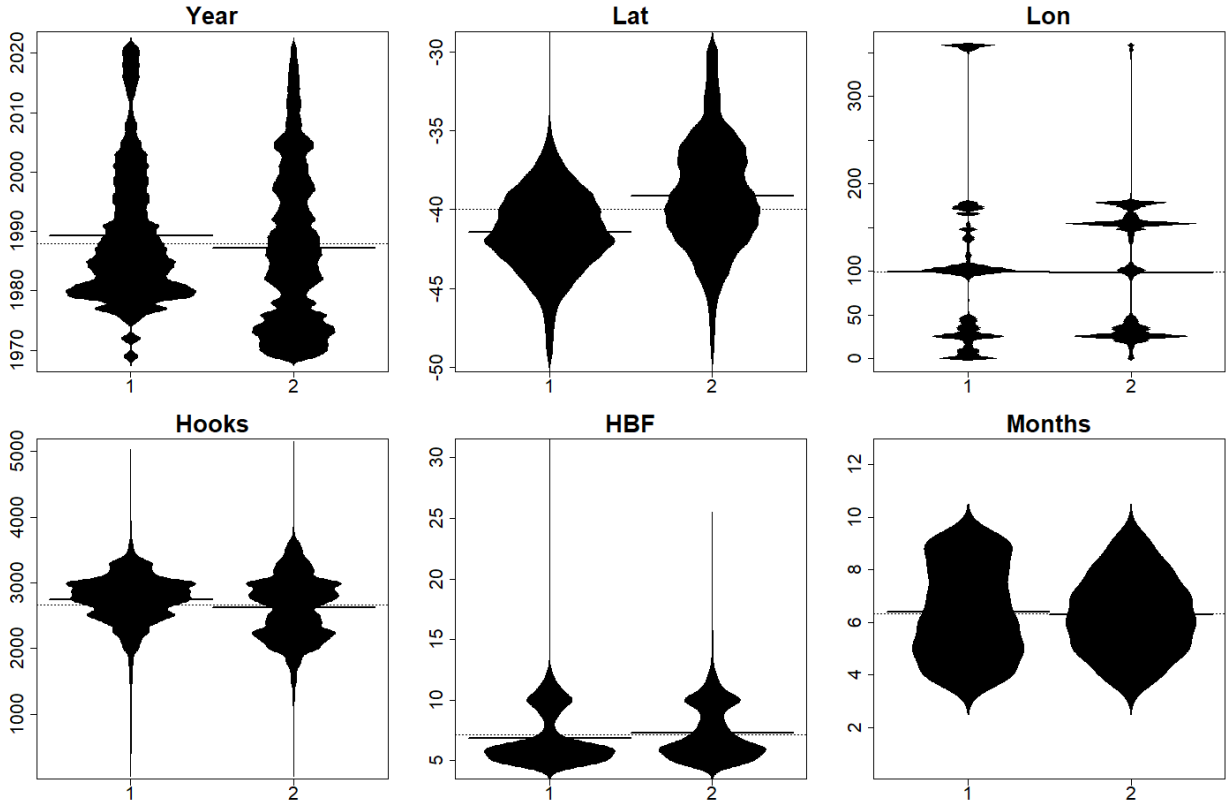

Fig. 44. Occurrence by variables in the group in the 2 cluster analysis as a sensitivity analysis.

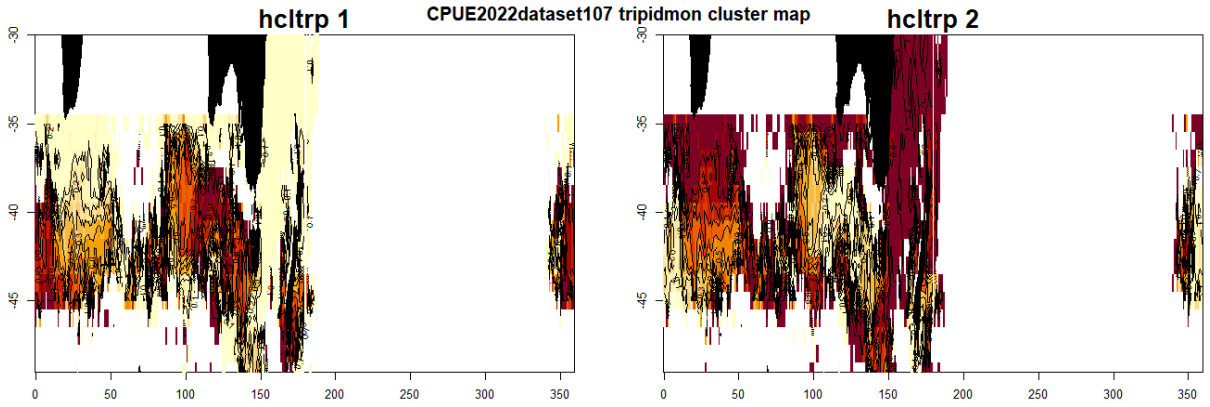

Fig. 45. Occurrence on map in the group in the 2 cluster analysis as a sensitivity analysis.

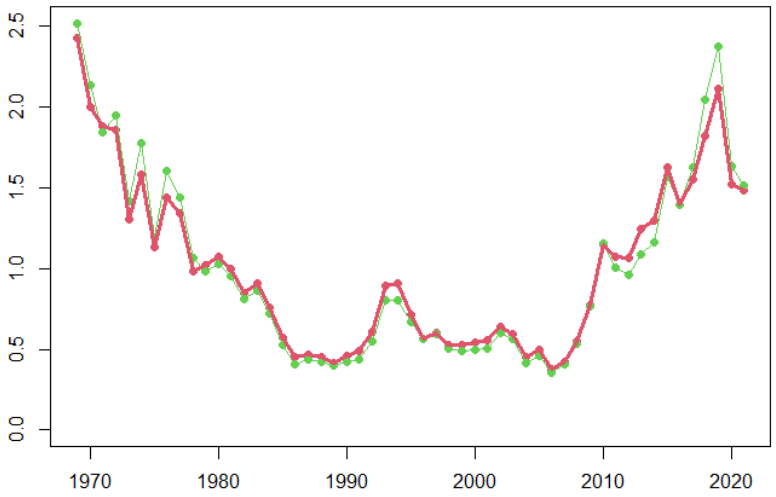

Fig. 46. Sensitivity analysis for the effect of 2 clusters instead of 4 clusters in the abundance index.
Red is the base case, and green is the sensitivity run (2 clusters).

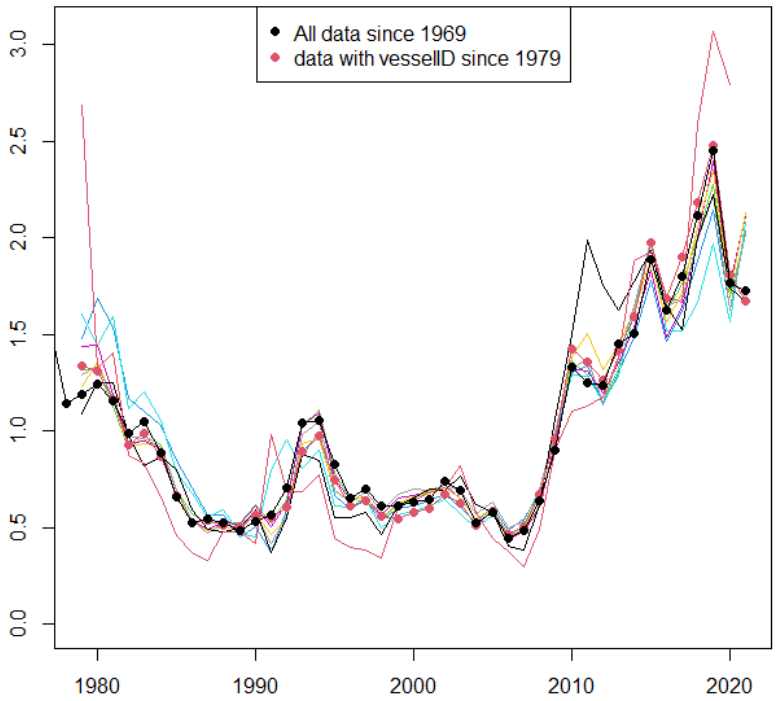

Fig. 47. Sensitivity analysis of core vessels selection.
Black and red lines with dots are the index from all vessels since 1969 and 1979, respectively. Black line is adjusted to the mean between 1979 and 2021. Indices from different core vessels data are shown in different colors.

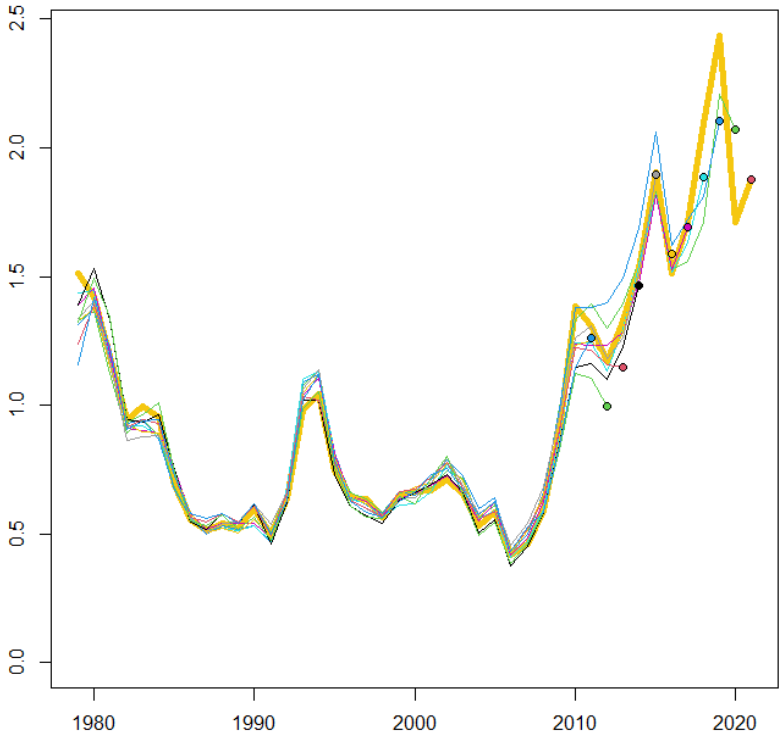

Fig. 48. Retrospective analysis in the sensitivity analysis of the core vessel.

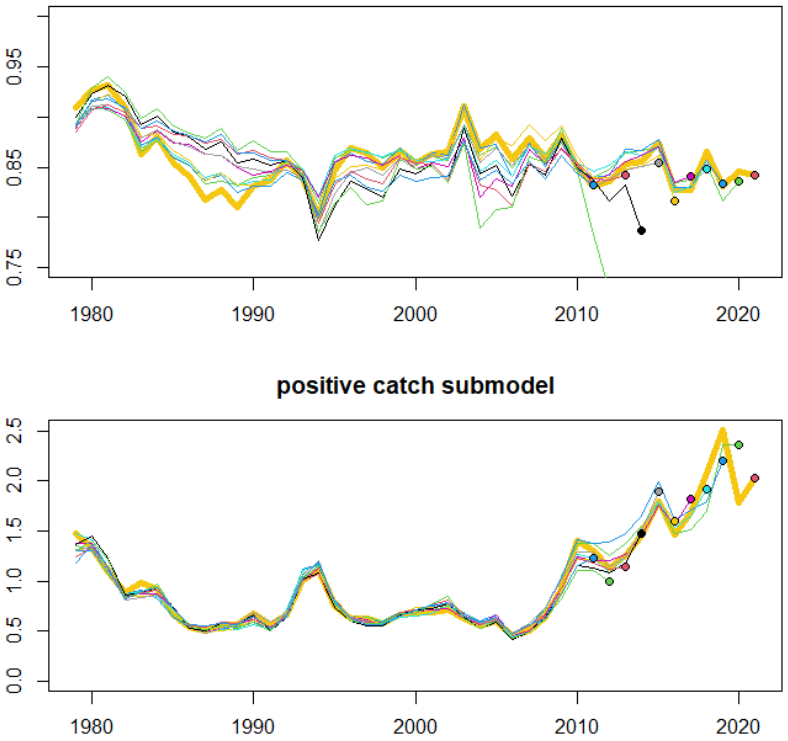

Fig. 49. Retrospective analysis in the sensitivity analysis of the core vessel by sub-model.
Upper panel is by binomial submodel and lower panels is by positive catch submodel.

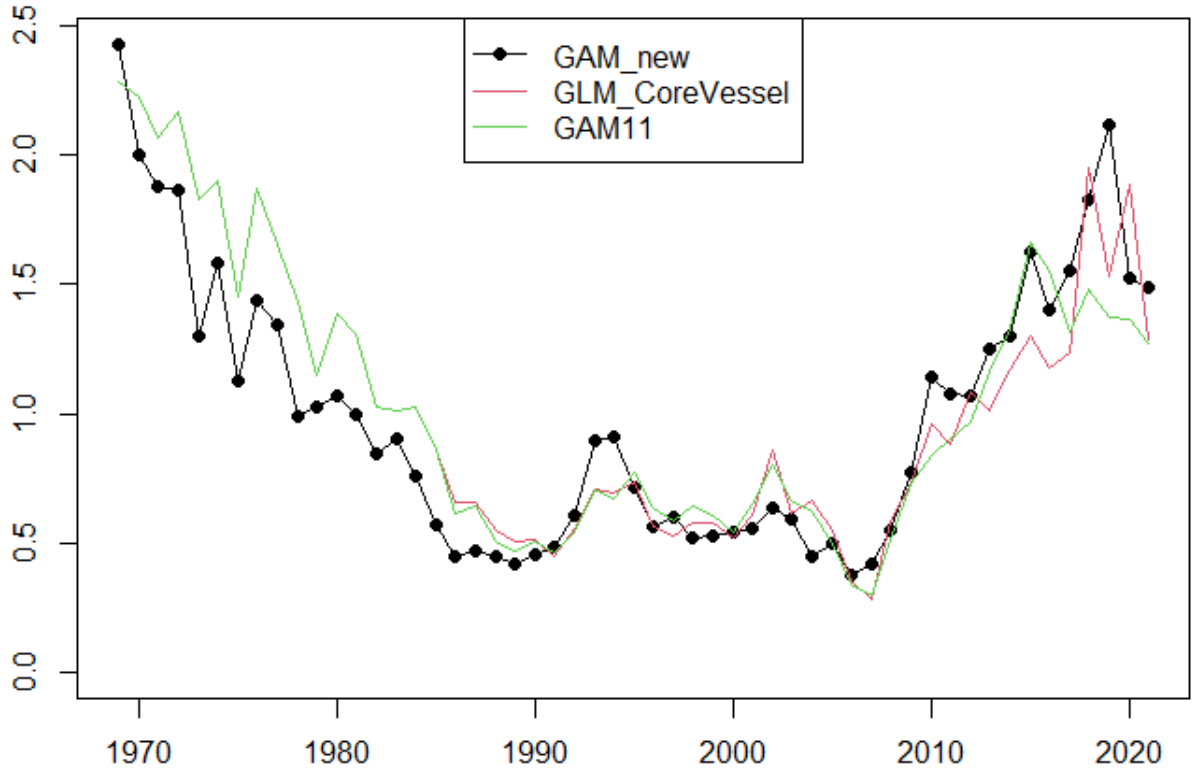

Fig. 50. Comparison of three abundance indices.
GLM_CoreVessel is the index by core vessel data with GLM base model in W0.8. GAM11 is the GAM model W0.8 used for the stock assessment in 2020.

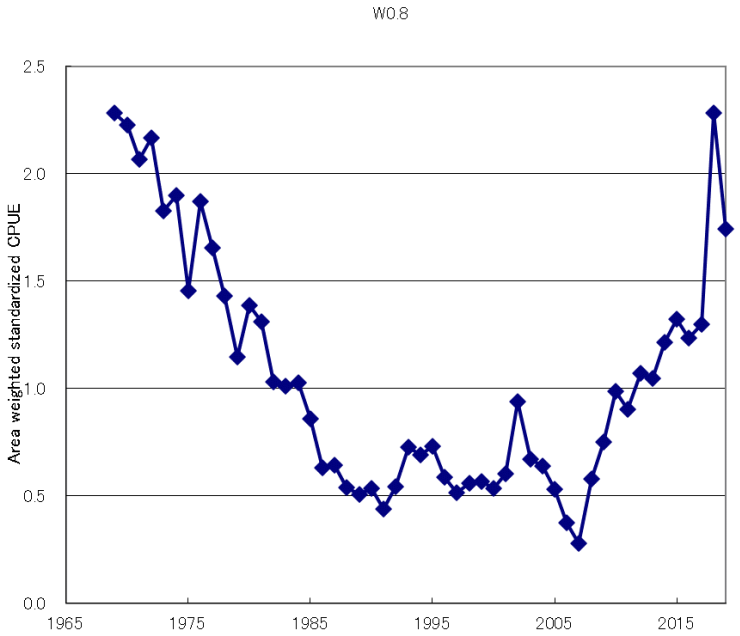

Fig. 51. Core vessel CPUE in W0.8 at ESC25 held in 2020.

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	3.80	0.9984	0.55
ti(lon.cnt)	19	18.67	0.9965	0.54
ti(lat)	3	2.30	0.9943	0.40
ti(lon.cnt,lat)	27	18.34	0.9409	0.02
ti(month,lat)	12	9.16	1.0036	0.63
ti(lon.cnt,month)	36	31.83	0.9961	0.55
ti(year,lat)	27	22.38	1.0055	0.68
ti(year,lon.cnt)	81	73.07	0.9307	0.00
ti(year,month)	36	34.35	0.9746	0.22
s(log(hook))	9	8.25	0.9860	0.23

PCSM

Term	k'	edf	k-index	p-value
ti(month)	5	4.67	1.0091	0.74
ti(lon.cnt)	19	17.65	0.9975	0.41
ti(lat)	3	2.96	0.9927	0.33
ti(lon.cnt,lat)	43	35.38	0.9973	0.41
ti(month,lat)	14	11.76	1.0021	0.56
ti(lon.cnt,month)	94	72.70	1.0075	0.70
ti(year,lat)	57	47.19	1.0036	0.60
ti(year,lon.cnt)	334	246.15	0.9742	0.03
ti(year,month)	95	80.61	1.0112	0.79
ti(lat,month,year)	151	113.13	0.9802	0.12
ti(lat,lon.cnt,month)	83	60.51	1.0092	0.63
ti(lat,lon.cnt,year)	295	238.55	0.9490	0.00
ti(year,lon.cnt,month)	782	580.12	0.9495	0.00
s(log(hook))	9	7.57	1.0217	0.92

Table 3. Diagnostic statistics of GAM.

Sub-model	BSM	PCSM
n.data	794,481	702,481
dev.expl	73.23%	49.13%
AIC	153,078	1,491,407
residual.df	794,203	700,906
REMLscore	1,705,734	375,849

Table 4. Abundance index as the base case.

Year	Index
1969	2.42910
1970	1.99828
1971	1.88008
1972	1.86216
1973	1.30350
1974	1.58577
1975	1.13166
1976	1.43957
1977	1.34506
1978	0.98698
1979	1.02370
1980	1.07063
1981	0.99663
1982	0.85022
1983	0.90602
1984	0.76263
1985	0.57080
1986	0.45156
1987	0.46857
1988	0.45037
1989	0.41795
1990	0.45992
1991	0.48896
1992	0.60896
1993	0.89510
1994	0.91043
1995	0.71296
1996	0.56431
1997	0.60376
1998	0.52519
1999	0.52944
2000	0.54541

Year	Index
2001	0.55715
2002	0.63665
2003	0.59650
2004	0.45361
2005	0.49771
2006	0.38058
2007	0.42048
2008	0.54826
2009	0.77517
2010	1.14512
2011	1.07434
2012	1.06752
2013	1.24914
2014	1.29931
2015	1.62459
2016	1.39826
2017	1.55107
2018	1.82492
2019	2.11360
2020	1.52267
2021	1.48767

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	73.23%	153,078	794,203	1,705,734	2,767
modA2.no5	-ti(lon, lat)	72.76%	155,682	794,220	4,007,886	5,372
modA2.no6	-ti(month, lat)	72.97%	154,509	794,211	4,184,870	4,199
modA2.no7	-ti(lon, month)	72.89%	154,911	794,234	2,030,894	4,600
modA2.no8	-ti(year, lat)	72.56%	156,822	794,223	7,193,970	6,512
modA2.no9	-ti(year, lon)	71.81%	160,987	794,276	2,173,631	10,676
modA2.no10	-ti(year, month)	72.27%	158,426	794,237	1,278,770	8,116
modA2.no15	-cl	70.80%	166,882	794,203	2,260,677	16,571
modA2.no16	-s(log(hook))	72.82%	155,388	794,211	1,812,437	5,077
modA1	Main effects	66.77%	189,473	794,391	1,245,736	39,163
modA2.p11	+ti(lat, month, year)	73.64%	150,772	794,177	12,211,104	462
modA2.p12	+ti(lat, lon, month)	73.51%	151,518	794,174	565,423,241	1,208
modA2.p13	+ti(lat, lon, year)	73.48%	151,662	794,174	4,177,097	1,352
modA2.p14	+ti(year, lon, month)	73.73%	150,311	794,151	1,298,659	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	49.13%	1,491,407	700,906	375,849	0
modB3.no5	-ti(lon, lat)	49.09%	1,491,992	700,895	376,060	584
modB3.no6	-ti(month, lat)	49.10%	1,491,798	700,902	375,974	391
modB3.no7	-ti(lon, month)	49.07%	1,492,363	700,862	376,290	956
modB3.no8	-ti(year, lat)	49.09%	1,491,982	700,906	375,999	575
modB3.no9	-ti(year, lon)	48.91%	1,494,576	700,866	376,796	3,169
modB3.no10	-ti(year, month)	49.09%	1,492,003	700,899	376,029	596
modB3.no11	-ti(lat, month, year)	49.08%	1,491,984	700,965	375,933	577
modB3.no12	-ti(lat, lon, month)	49.08%	1,492,141	700,919	376,033	734
modB3.no13	-ti(lat, lon, year)	48.79%	1,495,728	701,071	376,676	4,320
modB3.no14	-ti(year, lon, month)	48.15%	1,503,822	701,437	378,176	12,415
modB3.no15	-cl	48.81%	1,495,847	700,902	376,974	4,440
modB3.no16	-s(log(hook))	49.09%	1,492,023	700,911	376,004	616
modB1	Main effects	40.85%	1,594,318	702,390	398,925	102,911
modB2	Main and 2 way interactions	47.09%	1,517,112	701,842	380,842	25,705

Table 7. 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	598,042	598,042	100%	1,672	1,672	100%
2	56	3	598,042	231,455	39%	1,672	283	17%
3	56	5	598,042	164,768	28%	1,672	165	10%
4	56	7	598,042	120,900	20%	1,672	107	6%
5	50	3	598,042	214,842	36%	1,672	260	16%
6	40	3	598,042	182,569	31%	1,672	215	13%
7	30	3	598,042	143,390	24%	1,672	158	9%
8	20	3	598,042	91,185	15%	1,672	93	6%
9	20	5	598,042	57,220	10%	1,672	49	3%

			SBT catch			Run time in minutes			
Run	xx	yy	Original	core vessel data	%	BSM	PCSM	Total	%
1	9999	1	16,428,540	16,428,540	100%	6.78	14.79	21.57	100%
2	56	3	16,428,540	9,717,227	59%	3.05	6.42	9.47	44%
3	56	5	16,428,540	7,396,109	45%	2.23	4.96	7.19	33%
4	56	7	16,428,540	5,814,778	35%	1.86	3.03	4.89	23%
5	50	3	16,428,540	9,316,260	57%	2.42	5.17	7.59	35%
6	40	3	16,428,540	8,356,128	51%	2.18	4.25	6.44	30%
7	30	3	16,428,540	7,046,070	43%	1.73	3.22	4.95	23%
8	20	3	16,428,540	5,174,534	31%	1.17	1.77	2.94	14%
9	20	5	16,428,540	3,542,045	22%	0.75	1.34	2.09	10%