Standardisation of size-based indicators for bigeye and yellowfin tuna in the Indian Ocean

Marinelle Basson & Natalie Dowling
CSIRO, Marine Research
Castray Esplanade, Hobart
Australia

Abstract

Information on the size distribution, for example mean length, in the catch can potentially be used as an indicator of stock status. This study looked at the standardisation of such indices particularly for spatial and seasonal effects. Results for bigeye and yellowfin in the longline and purse seine fisheries in the Indian Ocean show that the standardised series are very similar to the 'nominal' indices. This may suggest that, for the current datasets, there is not a strong need to standardise. Standardisation does, however, have other benefits, including the availability of standard errors and insight into the spatial and seasonal patterns from the estimated coefficients.

1. Introduction

Simple indicators from fishery data, e.g. mean length in the catch, are increasingly used as potential indicators of stock status. Such indicators are either used on their own where no assessment is available (e.g. swordfish, skipjack; IOTC WPB and WPTT reports), used instead of an extensive assessment (e.g. southern bluefin tuna during development of management procedure), or used in conjunction with an assessment (bigeye, yellowfin tuna; IOTC WPTT reports). Changes or trends in these indicators could be caused by many different factors and it has been argued that indicators should ideally be standardised, in the same way the commercial longline CPUE is standardised. It is, for example, easy to see that if a stock is spatially disaggregated by size, changes in the timing and location of a fishery could lead to a change in the mean length in the catch which may have nothing to do with a change in the underlying size distribution of the overall exploitable stock.

In this study we consider only the size frequency data and attempt to standardise different measures from these data, including the mean, median and upper 80th percentile, with a view to developing standardised 'stock status indicators'. The idea is simple. At high harvest rates one may see fewer large, old fish. The longline fishery size frequency data could, for example, reflect an absence or decrease in the proportion of old, big fish. The purse seine fishery, on the other hand, could reflect an absence of small fish, which may indicate low recruitment. Given that this gear takes the young faster growing age classes, it could also reveal changes in mean size that may reflect changes in growth rate. Monitoring these indices could therefore provide information on changes in the underlying population structure. It is obvious that changes in the fishery itself could lead to changes in the size frequency, and this should always be borne in mind. This study does not address reasons why there may, or may not, have been changes.

It is important to note that this is only one step in a much wider study that started in September 2003 (SSI-study). The SSI-study aims to explore, via simulation, the robustness and sensitivity of a wide range of indicators, and to develop ways of combining several such indicators into a coherent framework for management decision-making. The study is addressing these issues primarily in the context of the domestic (Australian) fishery, but also in the wider context of the IOTC. The simulation part of the SSI-study may find that mean length is insensitive to stock changes (see e.g. Punt et. al. 2001), and that some other measure, possibly even based on data not currently available for the IOTC fisheries, is more reliable. The work presented here should be seen in this wider context, as an exploration of the effects of, and need for, standardisation of indices based on existing data.

2. Data and Methods

Data

Basic data were taken from the raw size frequency data held by the IOTC. Instead of converting from different measurement types, only data in terms of fork length (FL) were used. The raw data were manipulated to obtain the lower *n*th percentile, median, mean and upper *n*th percentile by gear type, fleet, area, month and quarter (month=quarter for longline).

We considered only two gear types: longline and purse seine. We chose fleets with most samples and longest time series: Japanese longline (LL) data and Spanish, French and NEI-EUR purse seine (PS) data. For the purse seine fleets, temporal trends were almost identical, so all analyses were conducted on data combined across fleets. For the longline data we also looked at the proportion of fish greater than a given length, chosen to reflect approximate size at maturity. For both species this was taken to be 100cm FL, based on the results of Shung (1973) for yellowfin tuna and Whitelaw and Unnithan (1997) for bigeye tuna. For bigeye, this length corresponded to most of the size frequency from longlines (LL), so a larger length threshold of 140cm FL was chosen based on examination of length-frequency data. The indicator is referred to as the 'proportion big'.

For purse seine data we also constructed an indicator based on the proportion of small fish (size corresponding approximately to 1 y.o. fish based on published growth curves: 60cm FL for bigeye and 55cm FL for yellowfin), with the notion that it might reflect recruitment dynamics. Changes in targeting (size) or in fishing techniques could also lead to increases in the proportion small fish in the catch, and the modelling we undertake here does not address the reasons for changes in indicators.

Prior to 1991, sampling of purse seine catch was poorer both in terms of sample size and sampling coverage. For example, there is a lack of data in some months and areas. Sample sizes are also not available in the database. The mean lengths plotted against year showed markedly more variability prior to 1991, and initial modelling attempts showed that it was difficult to deal with the early data. We therefore excluded purse seine data prior to 1991 from the GLM analysis.

Preliminary explorations showed so much similarity between annual values of mean, median, lower 20th and upper 80th percentile that we chose only to apply standardisations to one of these, namely the mean. The mean length and proportion of fish greater than a certain length were also similar. This is not surprising, but both were considered given that the next phase of the study will explore the relative robustness of these different candidate indicators.

Methods

In all cases simple general linear models (in Splus software) were used for standardisation. We identified plausible covariates to include, explored those and interactions between them where sensible, and extracted the appropriate 'standardised' series. Although it is possible to fit a model to size-based indicators from several fleets and gears combined, we decided to separate the gears. We considered that the size frequencies from the two gears are different enough that they may reflect different aspects of the population. Combining the data and 'correcting' for gear type could potentially lead to a loss of signal or information. Given that gear types were treated separately, the main covariates to consider were time and area effects.

There are two reasons for considering spatial factors: 1) to take account of changes in the locations of fishing over time and/or 2) to take account of changes in the size frequency in different areas over time. The second reason could be very important for a species like swordfish which is now thought to have a relatively high expected residence time in an area (Campbell and Hobday, 2003). Different levels of fishing intensity in different areas could therefore have different effects on the mean size in the catch over time. This type of change is more likely to be observed at a relatively coarse spatial scale than a very fine spatial scale. Given our interest in this type of effect for swordfish, we also explored this for bigeye and yellowfin (see below).

We consider that it is inappropriate to incorporate environmental factors into this type of standardisation unless there is clear evidence of a plausible mechanism for catchability changing differentially by size under different environmental conditions. Where environmental conditions in different areas or at different times affect the size distribution, the area or time factors should take care of that aspect. Although it is plausible, and indeed likely, that growth may be affected by environmental conditions, it is unlikely that the size frequency in any given year should be affected by environmental conditions in that year. This is because a size frequency distribution does not reflect the growth in a single year or the growth of a given cohort, but rather the cumulative growth of several cohorts over several years.

Models for longline mean length and proportion big

Multiplicative models were chosen for mean length, with area and time (year and quarter) as the main factors. The log of mean length is therefore an additive model of the relevant factors and we assume that errors are normally distributed. For both species, we chose several different spatial scales. At the fine scale level, we used 10-degree latitude bands and the lat-long grid positions for each quarter. The next level of detail was the "CPUE areas" that had been assigned when standardising catch rates (for bigeye, Okamoto & Miyabe (2003) (Figure 1), for yellowfin, Nishida *et al*. (2003) (Figure 2)), with the notion that the areas have some coherence in terms of fishery and other characteristics. For interest, we also looked at the Longhurst areas (Longhurst 2001), which are meant to represent "habitats" and which may therefore contain fish of certain age or size classes, if distribution is driven by habitat preference, and if the Longhurst area is at a spatial scale capable of distinguishing between those habitats. At the coarsest spatial scale, we considered the eastern and western Indian Ocean with a model that estimates a year effect in each area (covariate 'WestEast').

During the exploratory phase, more complex models with main effects and interactions, for example, between year, quarter and area were fitted and compared to more parsimonious models. Recall that a model with main effects implies that the mean length in any year, area and quarter is the sum of a year effect, an area effect and a quarter effect. A model with an interaction between, say, area and quarter implies that the quarterly effects are different in each of the spatial areas. It is important to bear the meaning of each model in mind, particularly when including interaction terms. Splus software was used and the model definitions for some of the simpler examples we considered are given in terms of S+ notation below:

Model type 1:glm(log(mean length) ~ year + area + quarter)
Model type 2 : glm(log(mean length) ~ year + area*quarter)
Model type 3: glm(log(mean length) ~ year*WestEast + quarter)

The 'proportion big' was fitted with a binomial model and factors were again time and the three different area definitions. The data were weighted by sample size, and the models were fitted to the subset of the data where sample size was greater than 50 (corresponding roughly to the lower quartile of a summary of sample sizes across the whole data set).

Models for Purse Seine mean length and ‘proportion small’

Data based on purse seine fishing contains a further important factor: fishing on free schools (FS data) or fishing on FADs (log school or LS data). The importance of this for the size frequency in the catch has been identified previously in IOTC reports and working papers. Essentially, the FS mean length had a bimodal distribution, whereas the LS mean length had a unimodal distribution (see e.g. Figure 12). Due to these differences, the free school and FAD data were considered using separate GLMs rather than trying to model the data with School Type as a factor. However, no systematic factor could be identified that could remove the bimodality in the free school data. Since this causes serious problems for the analysis and invalidates some of the model assumptions[1], analysis was only undertaken for the associated school data.

The area designations used with the longline data, particularly those based on CPUE standardisation, were not considered appropriate for the purse seine data. This fishery is concentrated in a much smaller area compared to the longline fleets. The lat-long grid points were used as one option, and another was constructed based on the 10-degree latitude bands containing the majority of the data (9N-19S). Data outside the 9N and 19S latitude bands were excluded from the analyses.

Modelling of the proportion small yielded problematic residuals under a binomial assumption. Since the

[1] The problems are similar to a situation where one is trying to fit a regression line through two clusters of points. The clusters will have strong leverage on the resulting regression line, but the model itself may be meaningless.

nominal temporal patterns for the proportion small were largely inversely related to those for the mean length, the standardized indices yielded no new information and, given that they were statistically unreliable, the results are not presented here.

Model selection

The key aim of fitting GLMs to the two indicators is to remove changes due to different timing (within the year) or different locations of fishing, and to extract the so-called year-effects from the models. The aim is not to determine statistically which factors are significant or not. Nonetheless, it is good practice to compare the goodness of fit of different models to avoid under- or over-fitting. The deviance residuals and q-q plots associated with each model were used to evaluate goodness of fit, homogeneity of variance and extent of conformance to a normal distribution. Where models were nested, appropriate significance tests were used to compare them, but in the case of the binomial models, these are unlikely to be reliable because of over-dispersion. Where models were not nested, particularly where the different area-designations were used, Akaike's information criterion (AIC) was used. Our main concern was, however, with the year-effects and whether they were sensitive to the different model formulations or variables included.

3. Results

3.1 Longline: Bigeye Tuna

Nominal trends and exploratory analyses

The mean lengths by area, quarter and year were highly variable. The annual averages of mean length were also very variable, ranging between about 117 and 134cm. Although there appear to be increases and decreases in the annual average series, these should be seen within the context of the range of the size frequency. Figure 3a shows the annual averages (over location and quarter) of the mean and of the 20^{th} and 80^{th} percentiles. Changes in the mean (and median) time series were small relative to the inter-percentile range. Similarly, the decrease in bigeye mean size over the last 5 years appears relatively insignificant relative to the long-term pattern, and the size range. The inter-quantile range has not changed substantially over the whole period, varying between 22 and 35cm with no systematic temporal pattern.

The pattern of the nominal time series plot for the proportion of fish >140cm FL was similar to that of mean length (Figure 3b), as expected. There was an increase in the proportion of large fish in the early 1990s to historically high levels (approx. 40%), followed by a decrease to the pre-1990 levels.

Exploratory analysis was undertaken to investigate the effect of alternative temporal scales and area designations on the inter-annual trends. This included comparing nominal time series broken down by quarter and area to find any obvious patterns, and comparing standardised indices across alternative models. The following points summarise the minor findings from the exploratory analyses:

- Plots of mean length (and the other candidate indicators) over time and by the different area designations did not show any clear patterns to inform model choice. Plots were generally noisy, with missing data and much inter-annual variability.

- For both mean length and "proportion big", the relative pattern of the year coefficients was very similar between model types and choice of area designation. Although models with interaction terms generally gave a better fit to the data, the standardised indices were essentially unchanged. Figure 4 shows the similarity in mean length GLM year coefficients for 5 alternative model types/area designations.

- Models using the CPUE standardisation areas generally gave lower AIC values compared to those using other area designations. Therefore, only the main effects models using the CPUE standardisation areas are subsequently considered in detail.

GLM standardisation of mean length and proportion >140cm FL

Results are presented for mean length modelled as:

glm(log(mean length) ~ year + CPUEarea +quarter).

This model has a common spatial pattern for each year and a common quarterly pattern in each year and area. The standardized indices showed a similar temporal trend to the nominal pattern (Figure 5a). However, the standardised indices for the proportion of bigeye > 140cm FL, using the same factors as the mean length model, showed a clear decline from the mid 1970s to 1990, which was not as apparent in the nominal trend (Figure 5b).

Figure 6 shows diagnostics and results for the mean length model. Residuals did not show any systematic patterns, either overall or when plotted against each set of predictor variables. The q-q plot showed that the data were almost normally distributed, with deviation mainly at the lower tail. These large negative residuals persisted in models with interactions. It is interesting to note that areas 1, 2, 3 and 6 had higher coefficients. These areas corresponded to the western Indian Ocean (Figure 1). Areas 4, 5 and 7, corresponding to the eastern Indian Ocean, had lower coefficients. The lowest-value quarter coefficient was in quarter 3. The results imply a 5cm difference between mean length in quarter 1 versus quarter 3, and a 14cm difference between mean length in area 1 versus area 7.

Residuals for the model of proportion of bigeye >140cm FL were also reasonably homoscedastic, close to normal and showed no pattern when plotted against each set of predictor variables. The area and quarter coefficients followed a similar pattern to those from the mean length model.

The standard errors of the standardised indices tended to be inversely related to sample size (Figure 7). This effect was more marked when weighting by sample size was used in the models. There is therefore less relative certainty associated with the standardised trends from more recent years, and this will have an effect on how such a series is implemented as a "stock status indicator". This comment also applies to the other standardised indices presented below.

3.2 Longline: Yellowfin Tuna

Nominal trends and exploratory analyses

Mean fork length of Japanese yellowfin catch showed the most dramatic change in the first 5 or 6 years of the fishery with a drop from ~145cm to ~115cm (Figure 8a). As with bigeye, plots of the annual average 20^{th} percentile, mean and upper 80^{th} percentile showed that, apart from the first 5-6 years in the time series, changes in the mean were small relative to the size range of the length frequency. The inter-percentile range varied between 12 and 29cm, with the narrowest range at the start of the time series, followed by an increase and then a decrease in the final few years. We did not have time to enquire whether there are any problems with, or less reliability in the data, prior to, say, 1965 (the first year for which bigeye data are available). Since catches in these early years were low relative to catches in later years, it seems unlikely that this would have been the effect of fishing, but this cannot be ruled out.

It is interesting that time series of minimum, median and maximum length values generally also showed parallel trends, with the notable exception of a decrease in minimum size in the late 1960’s and early 1970’s. This could possibly indicate one or more strong recruitments. One should also remember that the minimum and maximum would be far more sensitive to sample size and sampling noise than, say, the mean or median.

There was much inter-annual variation but no consistent temporal trend in the proportion of fish >100cm (Figure 8b). As expected, this series showed the same features as the mean length series.

The following points summarise the minor findings from the exploratory analyses:

- Plots of candidate indicators over time and by the different area designations did not show any clear patterns to inform model choice. Plots were generally noisy, with missing data and much inter-annual variability.

- A decrease in mean length since 1952 (see below) was seen in both the western and eastern Indian Ocean, but the increase in the mid 1980s until 1994 was mainly in the western Indian Ocean, and most prominent in the northern hemisphere.

- For mean length, Model 3, which incorporated latitude, longitude and quarter in a single factor, gave the lowest AIC and residual deviance of all the models. For the proportion of fish >100cm FL, models including the CPUE standardisation areas gave the lowest AIC among all model types. However, for both indicator variables, there was very little difference between model types and choice of area designation, in terms of the relative pattern of the time series of year coefficients, with the extent of similarity consistent with that shown for bigeye (Figure 4). As such, only the main effects models using the CPUE standardisation areas are subsequently considered in detail.

GLM standardisation of mean length and proportion >100cm FL

Results are presented for mean length modelled as:

glm(log(mean length) ~ year + CPUEarea +quarter).

Standardised indices showed a similar trend relative to the nominal plot, with the most recent value similar to the long-term mean (Figure 9a). Figure 10 shows diagnostics and the other estimated coefficients. Residuals showed some heteroscedasticity and many large negative values. Apart from the tails, the distribution was reasonably normal. The area coefficients were highest for area 4 (south-east Indian Ocean, Figure 2), and lowest for area 5 (north-east Indian Ocean south of the Bay of Bengal). The quarter coefficients varied depending on whether an interaction term was included in the model, but were consistently low for quarter 2 and high for quarters 1 and 4.

Fitted and observed values of the 'proportion big' (>100cm FL) were all clustered close to 1.0, indicating that for the majority of samples, all or almost all fish are “mature”. The standardized indices showed a consistent pattern with the nominal trend (Figure 9b). Figure 11 shows that residuals were reasonably homoscedastic but there is again a tail of large negative residuals. This feature was present irrespective of the model choice. Area coefficients showed somewhat different relative patterns than in the mean length model. Quarter coefficients were similar, except for quarter 1, relative to those from the mean length model. As for bigeye, standard errors were, as expected, larger in years where sampling was low.

3.3. Purse seine: Bigeye Tuna

Nominal trends and exploratory analyses

We have already commented on the bimodal nature of the mean length for free school data. This bimodality is illustrated in a histogram of mean length (with means being calculated for each month/year/5-degree square/fleet combination, Figure 12b). The annual average (over month, location, fleet) of mean length and of 20th and 80th percentiles generally resulted in parallel trends but, unlike the longline data, changes in the mean were large relative to the inter-percentile range (Figure 12a). This suggests the possibility that more than one age class is being exploited in very different proportions from year to year. It is interesting to note that results in Stequert and Conand (2003) imply that the mode at 60cm coincides with 1-year old fish and the mode at 110cm (or above) coincides with 2-and-a-half year or older fish. As stated in the methods, a systematic factor responsible for the bimodality could not be identified and thus no GLM standardisations were undertaken on the free school data.

Mean lengths from associated schools (LS data) were generally smaller than those from free schools and the distribution of mean length was unimodal (Figure 12). Changes in the mean length were small relative to the inter-percentile range, which varied between 12 and 22cm.

The proportion of 'small' fish (<60cm FL) for free and associated schools showed an inverse temporal patterns to that for mean length. Since these temporal patterns gave no new information to that provided by the mean length, and due to persistent problems with the residuals under a binomial assumption, no GLM analyses are presented for the proportion of small fish.

The exploratory analyses revealed the following minor points:

- Temporal plots of free and associated mean size and by 10-degree latitude bands and Longhurst areas did not show any clear patterns to inform model choice. Data were generally concentrated within one area, with missing data and much inter-annual variability characterizing the other areas.

- The substitution of month for quarter in the GLMs, and/or the inclusion of fleet or area (Longhurst, latitude band) factors in the GLMs made no discernable difference to the standardised index. The similarity between models was even greater than for the longline analyses shown in Figure 4.

GLM standardisation of mean length

There appears to be a strongly significant interaction between year and quarter, and including this interaction reduced the deviance considerably[2]. This implies that the quarterly pattern in mean length varied according to year. Results for the following simple models are presented:

Model 1 (main effects): glm(log(mean length) ~ year + quarter)

Model 2 (interaction effects): glm(log(mean length) ~ year + year:quarter)

The standardized indices for the main effects model were almost identical to the nominal series (Figure 13), but those for the interaction effects model were slightly different. Note that the indices in Figure 13 are relative to quarter 1. For the main effects model, the pattern over time would be identical for each quarter, but the level would be different. For the interaction model, however, the pattern over time would differ slightly depending on which quarter one is looking at.

Diagnostics for the two models are shown in Figures 14 and 15. There are some positive outliers but no systematic patterns or heteroscedasticity. Outliers were not removed as doing so did not change the overall time series pattern of the standardized indices. The q-q plots for both models showed some deviation from normality, particularly at the right tail. However, as the emphasis of the GLMs is not on optimal model selection, some deviance is acceptable. The comparison shows that, although the residuals are slightly smaller for the more complicated model, the large positive values, and the departure from normality persist.

The quarterly coefficients in the main effects model were highest for quarters 1 and 2, suggesting that the mean size of associated school fish is larger over the summer months (Figure 14). In the interaction model, the quarterly effects are estimated relative to quarter 1 in each year. Figure 15 shows that the relative quarterly patterns were similar for years i) 1991-1994, and 2000, ii) 1996 and 2001, and c) 1999 and 2002. There is no pattern apparent in this, and it is probably simply due to 'natural' inter-annual variability. It does, however, show that quarter 2 has the highest coefficients of the three in most of the years, as suggested by model 1.

3.4 Purse seine: Yellowfin Tuna

Nominal trends and exploratory analyses

Mean lengths from associated schools were again generally smaller than those from free schools. The mean and 20th percentile trends were mostly parallel, but the 80th percentile was very variable and appears to have dropped from ~100cm in 1992 to ~70cm in the most recent years (Figure 16). This could be worrying, particularly when seen together with the slight decrease in the longline mean length since 1994.

Free school data and the proportion small yellowfin (<55cm FL) in the associated school data were not used in GLM analyses for the same reasons as given for bigeye.

Exploratory analyses for the yellowfin data lead to similar conclusions as those summarised for bigeye above.

GLM standardisation of mean length

Results are presented for the two models used with the LS data for bigeye. An interaction between year and quarter was again considered to be important. The standardized indices for the main effects model (Model 1) were almost identical to the nominal mean length (Figure 17), and those from the interaction model were

[2] Model 1 reduced the null deviance (19.83) by 43% (residual deviance 11.36), while model 2 reduced the null deviance by 59% (residual deviance 8.07) (ANOVA of model 1 vs. model 2: df=33, F=51.88, $P < 0.000001$).

also quite similar. Recall, however, that for the interaction model, the annual time trends would be slightly different depending on which quarter one considers.

Residual and q-q plots for the two models were very similar and are only shown for model 2. There was a cluster of very large positive residuals for the lower fitted values (Figure 18), due to positive outliers in 2001 and 2002, from all quarters. Apart from this cluster of residuals, there are no other systematic patterns or heteroscedasticity. Outliers were not removed as doing so did not change the overall time series pattern of the standardized indices. More complex models did not 'eradicate' the large positive residuals.

As with bigeye tuna, the quarterly coefficients of the main effects model were highest for quarters 1 and 2, suggesting that the mean size of associated school fish is larger over the summer months. The temporal pattern for the interaction effects was reasonably similar for quarters 3 and 4, particularly prior to 1998 (Figure 18). Relative quarterly patterns are similar for years i) 1993-1994, 1998 and 2000 and ii) 1996-1997, 2001 and 2002.

4. Discussion

The results presented here show that it is possible to standardise the mean length and other similar indices reasonably successfully. Admittedly, there are cases where the model fit is rather poor with undesirable features in the residuals. In our experience, there is very little that can be done (given the set of available covariates) to improve this. The standardisation does not, however, seem to have much of an effect on the indicators considered here. The pattern of an annual 'nominal' index was found to be very similar to its standardised version. This suggests that, apart from occasional checks to see that this still holds, it may not be necessary to standardise these particular indices for yellowfin and bigeye.

Although the standardised and 'nominal' series have almost identical patterns over time, there are a few useful by-products of the standardisation process. Most importantly, there are estimates of standard error that are likely to be more robust than direct estimates from the raw data. When implementing such an index in a monitoring framework the standard errors would play a strong role in assessing whether there has been enough of a change in the index to warrant management action. The increase in standard errors when sample sizes decrease was noted, and this would have to be taken into account when implementing such indices into a monitoring or management framework.

The standardisation also allows for the integration of data in a statistical framework and for 'like with like' comparisons in different areas or in the same area in different quarters. The standardisation can reduce the effects of 'noise' in the data. This does not mean that careful scrutiny of the detailed data are not necessary. Particularly when changes are observed, there would be a need to try to understand why and where those changes are occurring.

Although one would hope that standardisation could remove the effects of spatial or temporal changes in the fishery, we have already noted that there may be other potential factors that affect catch at size but which cannot be incorporated due to an absence of data or the aggregated nature of the data. Therefore, this type of analysis does not, and cannot, take into account changes due to factors that could not be included in the model.

We have also noted that mean or median length may not be a sensitive or reliable indicator. It is easy to show, using simple yield per recruit analysis, that the largest change in mean length as fishing mortality increases occurs for slow growing species. This suggests that mean length may be particularly poor for yellowfin which is relatively fast growing, but it could be reasonable for a species like swordfish. This will be investigated under the SSI-study, and results will be presented to the IOTC WPB in the future.

Acknowledgements

This work was supported by funding from the Australian Fisheries Research Development Corporation (project 2003/042) and CSIRO. We thank the IOTC Secretariat, Miguel Herrera in particular, for assistance with the data and for prompt responses to our questions. Any misinterpretations of the data are entirely our fault.

References

Campbell, R.A. and A. Hobday 2003. Swordfish-environment-seamount-fishery interactions off eastern Australia. Report to the Australian Fisheries Management Authority, Canberra, Australia.

Longhurst, A.R. 2001. Ecological geography of the sea. Academic Press, San Diego, California.

Nishida, T, Bigelow, K., Mohri, M. and F. Marsac 2003. Comparative study on Japanese tuna longline CPUE standardisation of yellowfin tuna (*Thunnus albacares*) in the Indian Ocean based on two methods: general linear model (GLM) and habitat-based model (HBM)/GLM combined – (1958-2001). IOTC WPTT-03-05.

Okamoto, H. and N. Miyabe 2003. Standardised Japanese longline CPUE for bigeye tuna in the Indian Ocean up to 2001. IOTC WPTT-03-11.

Punt, A.E.; Campbell, R.A.; Smith, A.D.M. 2001. Evaluating empirical indicators and reference points for fisheries management: application to the broadbill swordfish fishery off eastern Australia. Marine & Freshwater Research Vol. 52, no. 6, pp. 819-832.

Shung, S.H. 1973. The sexual activity of yellowfin tuna caught by the longline fishery in the Indian Ocean, based on the examination of ovaries. Bulletin of the Far Seas Fisheries Laboratory Vol. 9, pp. 123-142.

Stequert, B. and F. Conand. 2003. Age and Growth of bigeye tuna (Thunnus obesus) in the western Indian Ocean. IOTC WPTT-03-Inf2.

Whitelaw, A.W. and Unnithan, V.K. (1997). Synopsis of the distribution, biology and fisheries of the bigeye tuna (*Thunnus obesus*, Lowe) with a bibliography. Australia, CSIRO Marine Laboratories Report 228, 62pp.

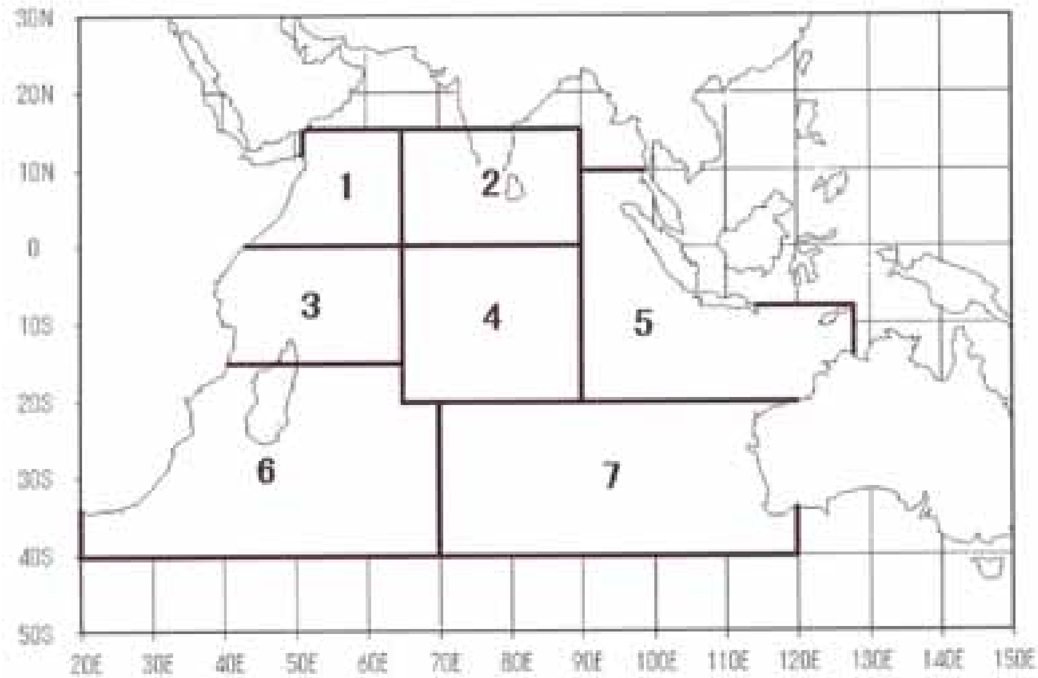


Figure 1: The areas used for standardisation of bigeye tuna longline CPUE (Okamoto and Miyabe, 2003).

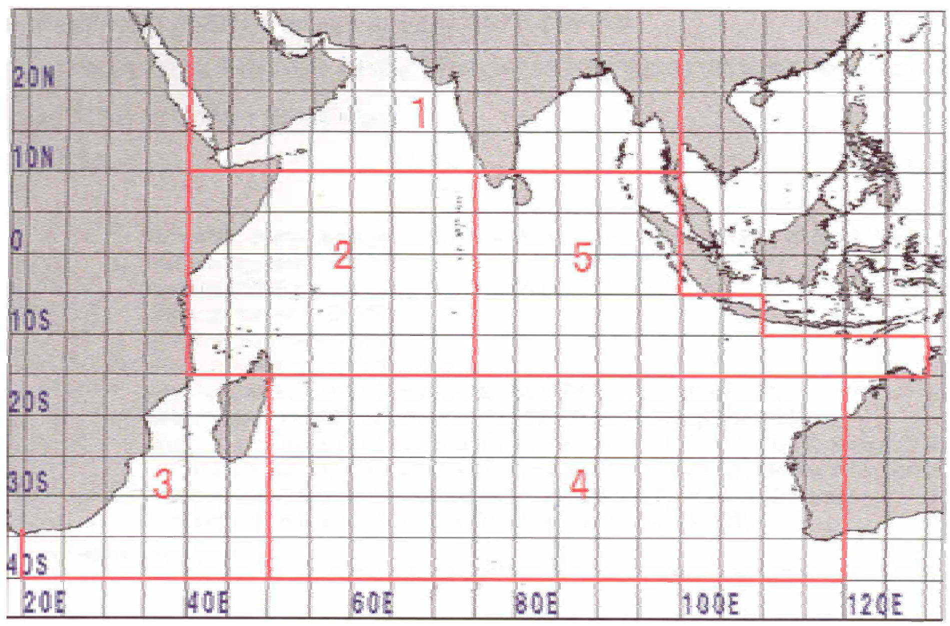


Figure 2: The areas used for standardisation of yellowfin tuna longline CPUE (Nishida *et al*., 2003).

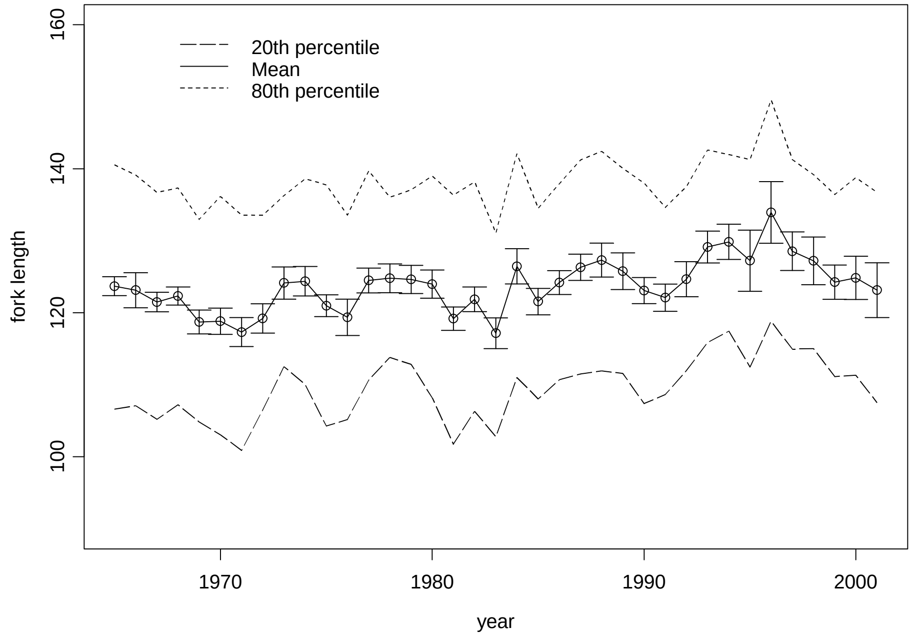


Figure 3a: Bigeye, Japanese Longline. Annual averages of mean length and average 20th & 80th percentiles (FL in cm) for the whole of the IO. Error bars are ±1 standard error (s.e.).

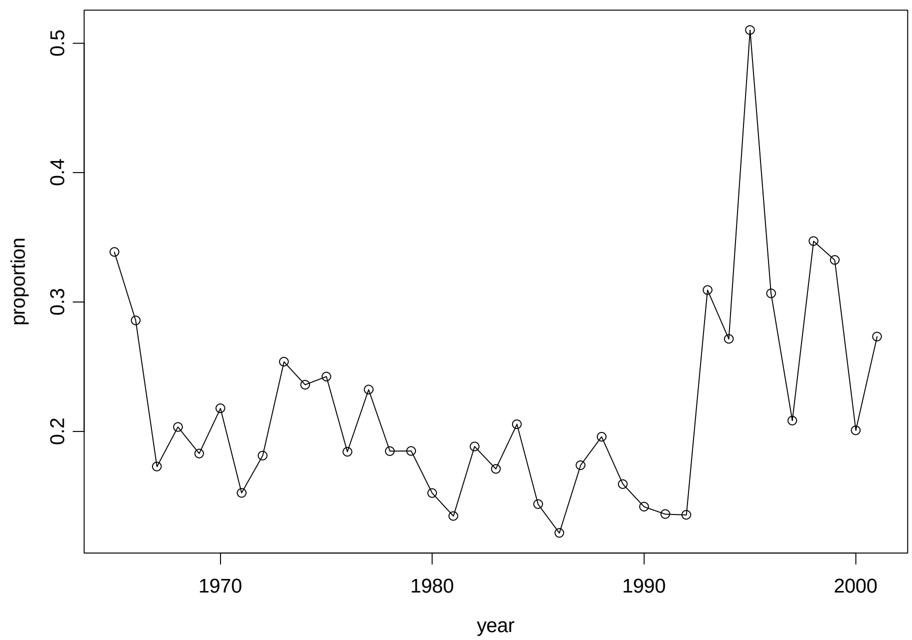


Figure 3b: Bigeye, Japanese Longline. Annual proportion >140cm FL (FL in cm) for the whole of the IO.

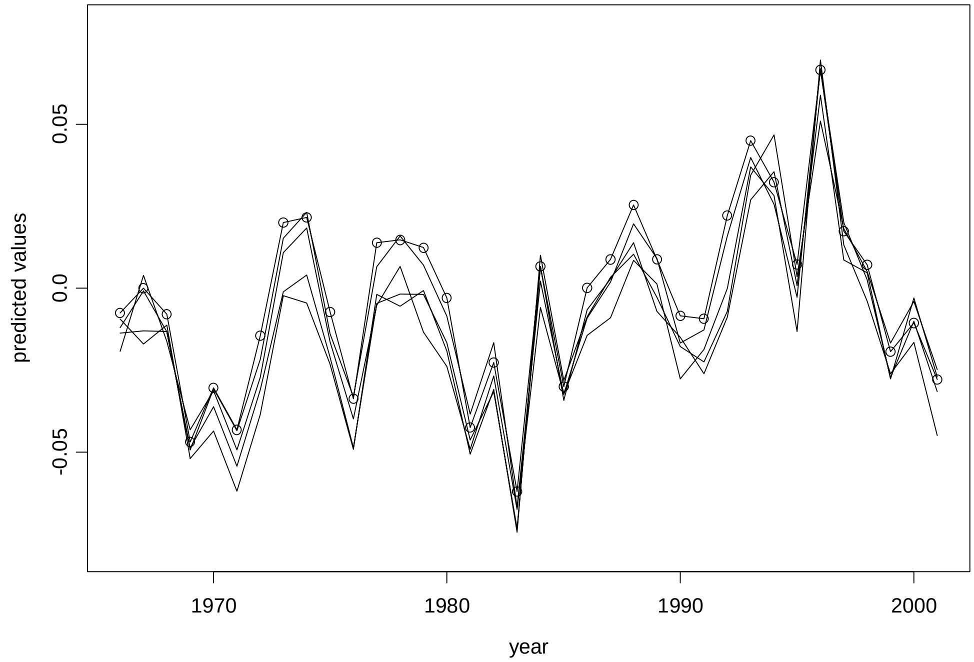


Figure 4: Bigeye, Japanese Longline. Year coefficients overlaid for GLMs of log(mean length) as a function of i) Year + CPUEarea + Quarter, ii) Year + Longhurst + Quarter, iii) Year + Latitude band + Quarter, iv) Year + LongLatQtr, v) Year + CPUEarea*Quarter. (Individual models have not been identified; the figure is meant to illustrate the similarity between models.)

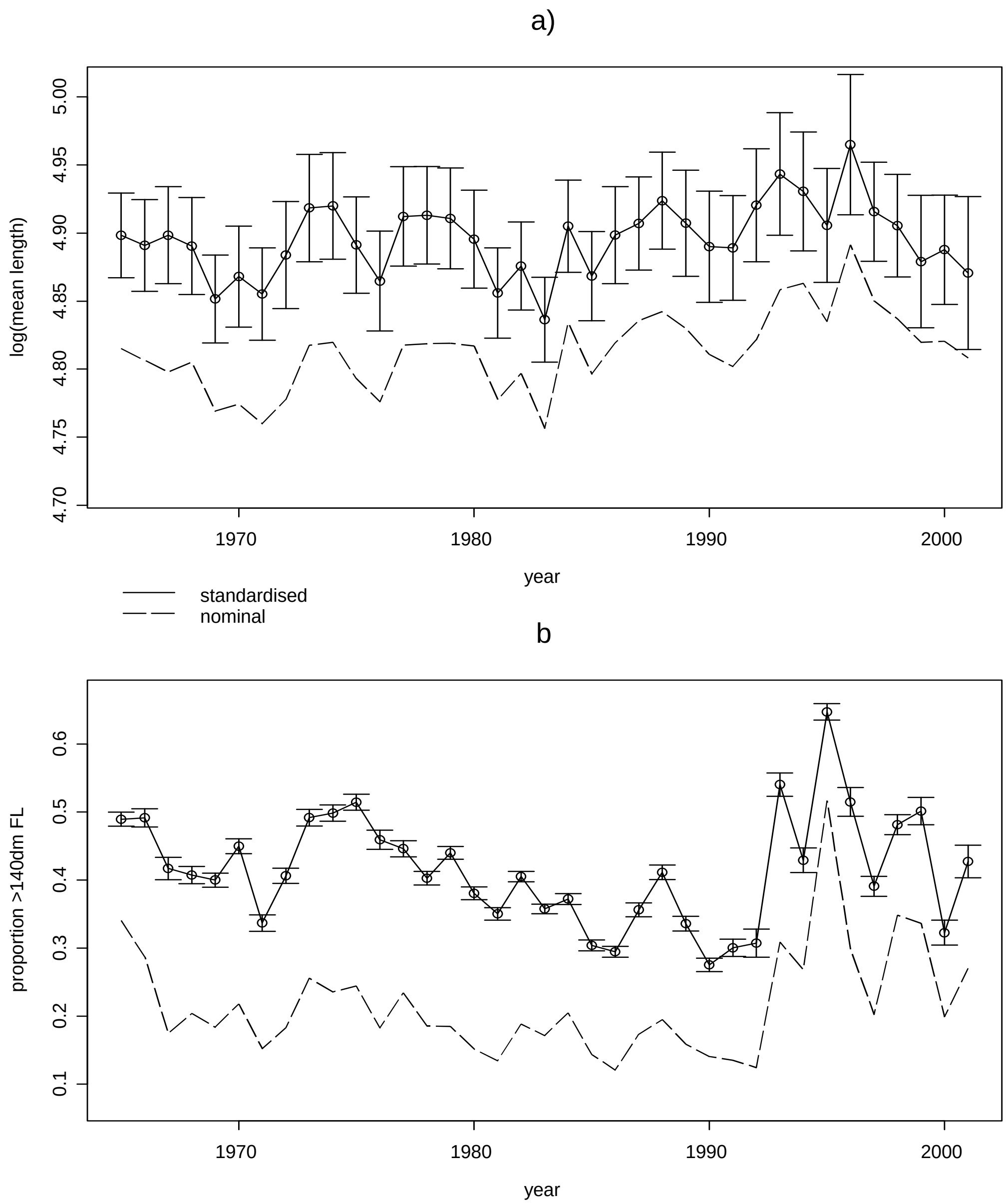


Figure 5: Bigeye, Japanese Longline. Time series of a) standardized and nominal bigeye mean length indices on the log scale, and b) standardized and nominal indices for proportion of bigeye >140cm FL. Error bars are ±1 s.e. The differences in magnitude are because the standardized indices are the predicted values for CPUE area 1 in quarter 1, while the nominal indices are the annual averages across all areas and quarters.

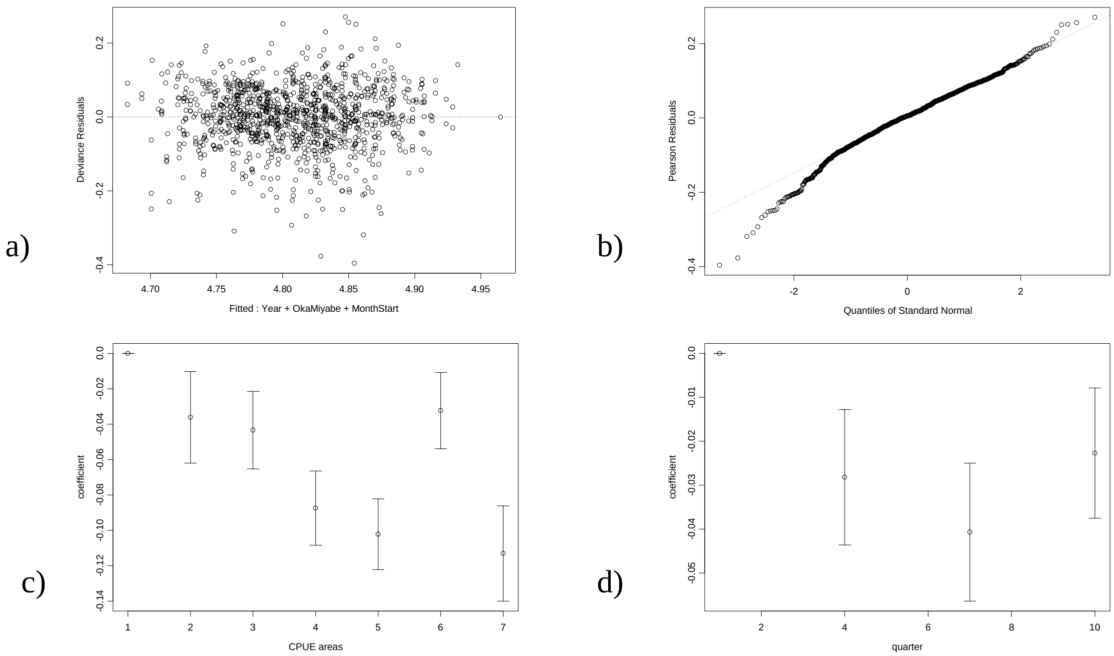


Figure 6: Bigeye, Japanese Longline. Plots of a) deviance residuals vs. fitted values, b) Pearson residuals vs. quantiles of standard normal, c) CPUE area coefficients and d) quarter coefficients, for the bigeye mean length main effects GLM. Error bars are ±1 s.e.

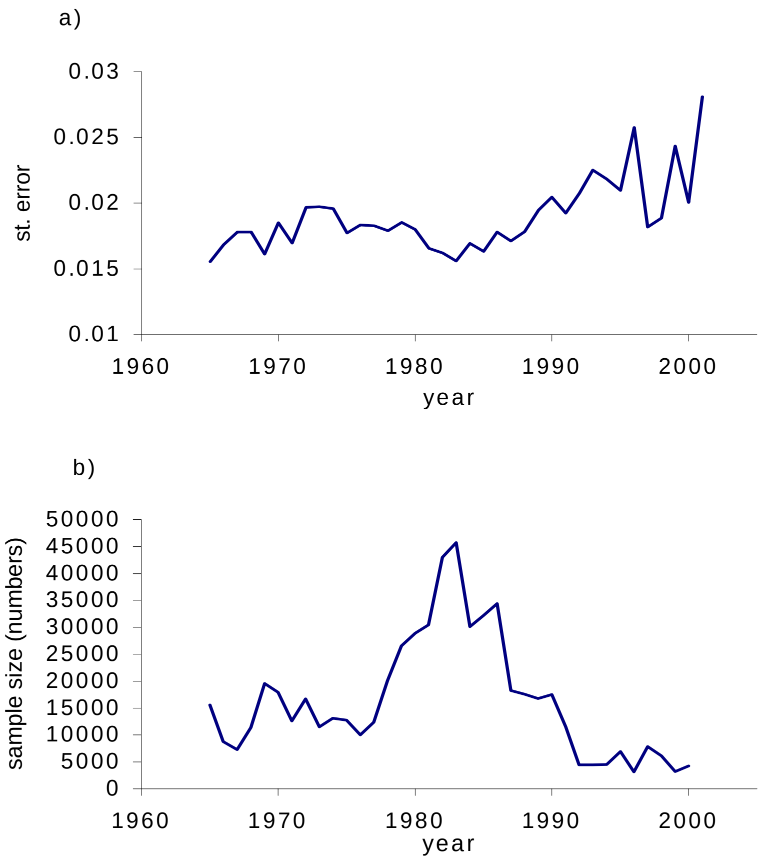


Figure 7: Bigeye, Japanese Longline. Standard errors associated with standardized indices for a) bigeye mean length and b) annual total sample sizes for bigeye mean length.

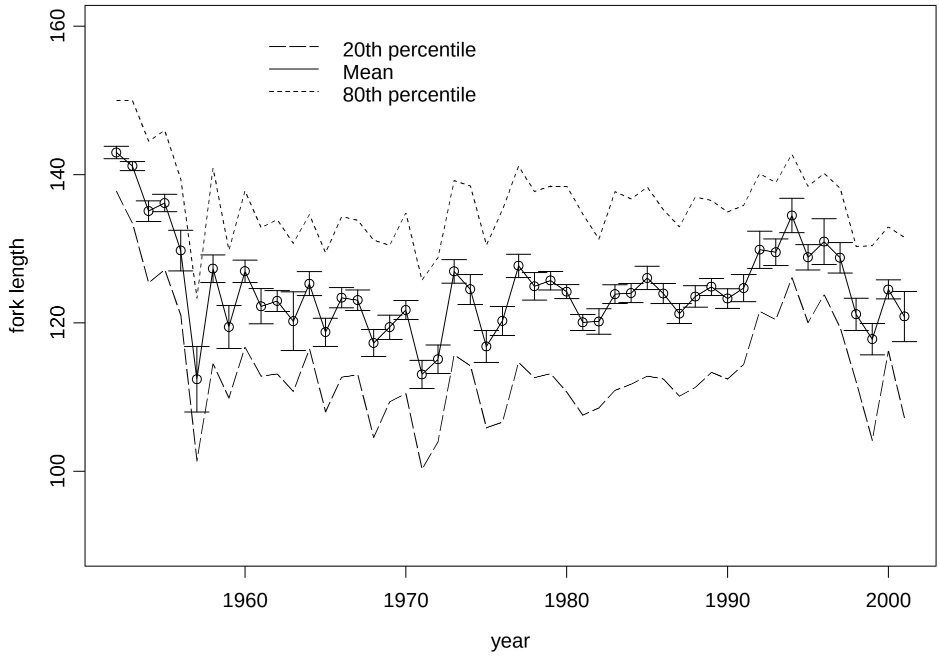
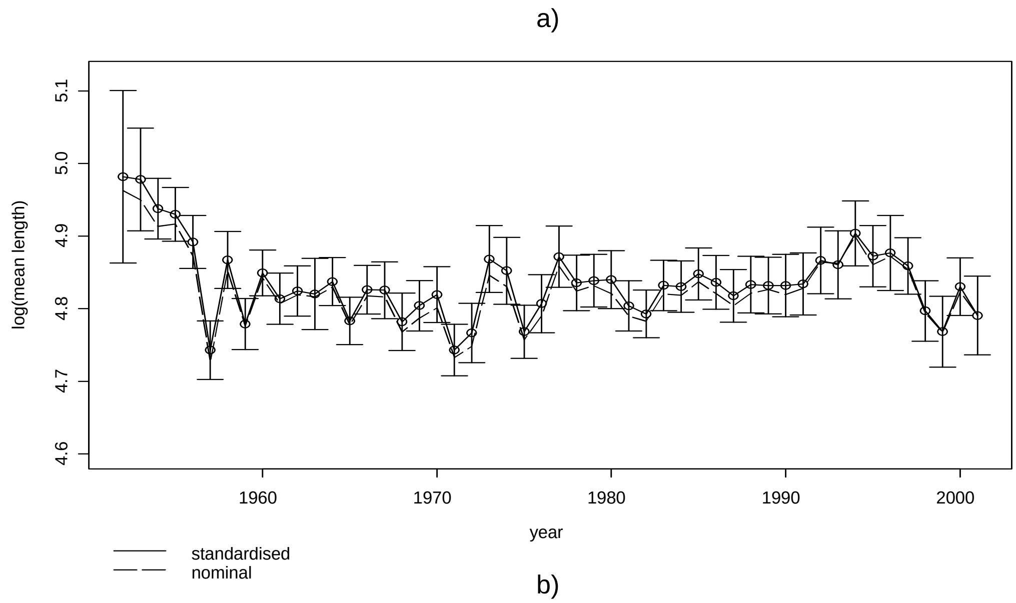


Figure 8a: Yellowfin, Japanese Longline. Annual averages of mean length and 20th & 80th percentiles (FL in cm) for the whole of the IO. Error bars are ±1 s.e.

Figure 8b: Yellowfin, Japanese Longline. Annual proportion >100cm FL (FL in cm) for the whole of the IO.



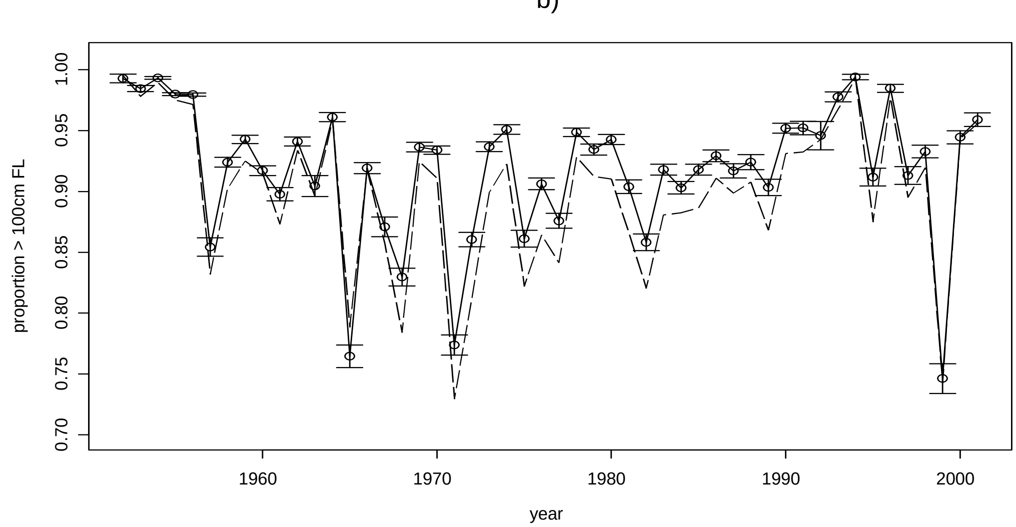


Figure 9: Yellowfin, Japanese Longline. Time series of a) standardized and nominal yellowfin mean length indices on the log scale, and b) standardized and nominal indices for proportion of yellowfin >100cm FL. Error bars are ±1 s.e. Note that the standardized indices are the predicted values in CPUE area 1 in quarter 1.

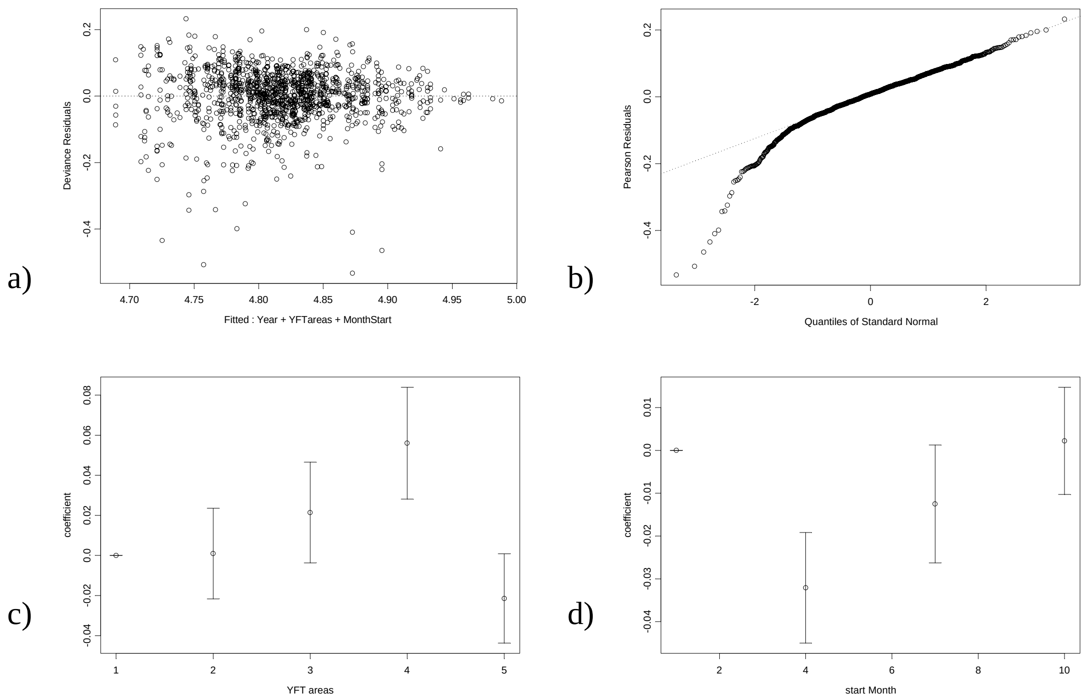


Figure 10: Yellowfin, Japanese Longline. Plots of a) deviance residuals vs. fitted values, b) Pearson residuals vs. quantiles of standard normal, c) CPUE area coefficients and d) quarter coefficients, for the yellowfin mean length main effects GLM. Error bars are ±1 s.e.

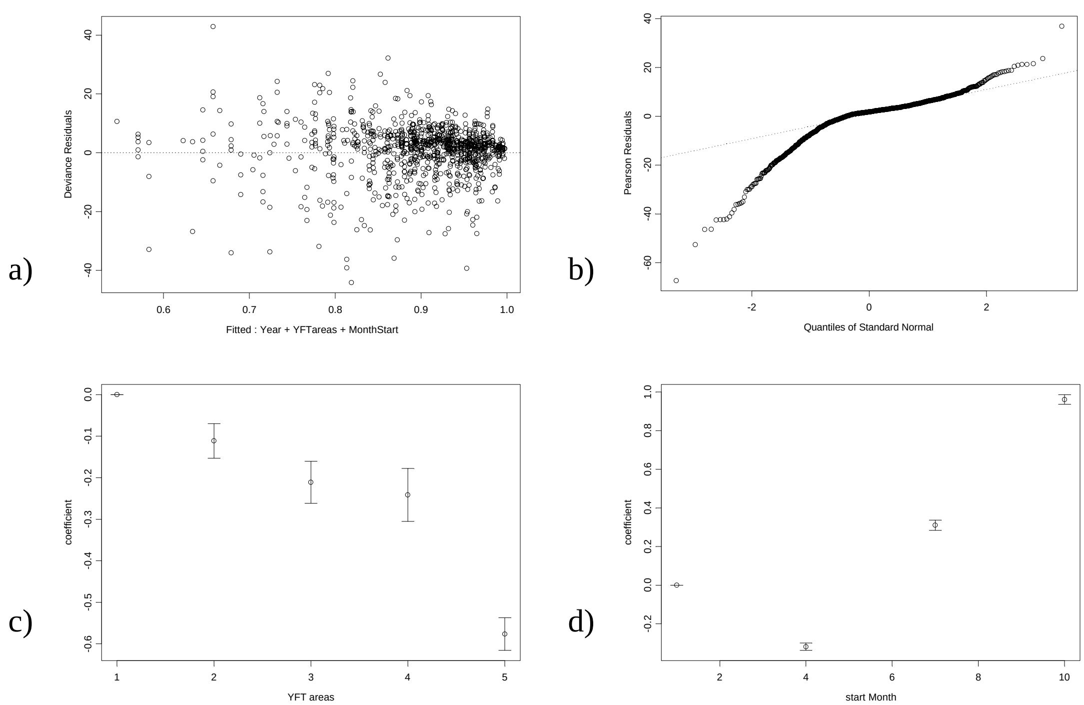


Figure 11: Yellowfin, Japanese Longline. Plots of a) deviance residuals vs. fitted values, b) Pearson residuals vs. quantiles of standard normal, c) CPUE area coefficients and d) quarter coefficients, for the GLM of proportion of yellowfin >100cm FL. Error bars are ±1 s.e.

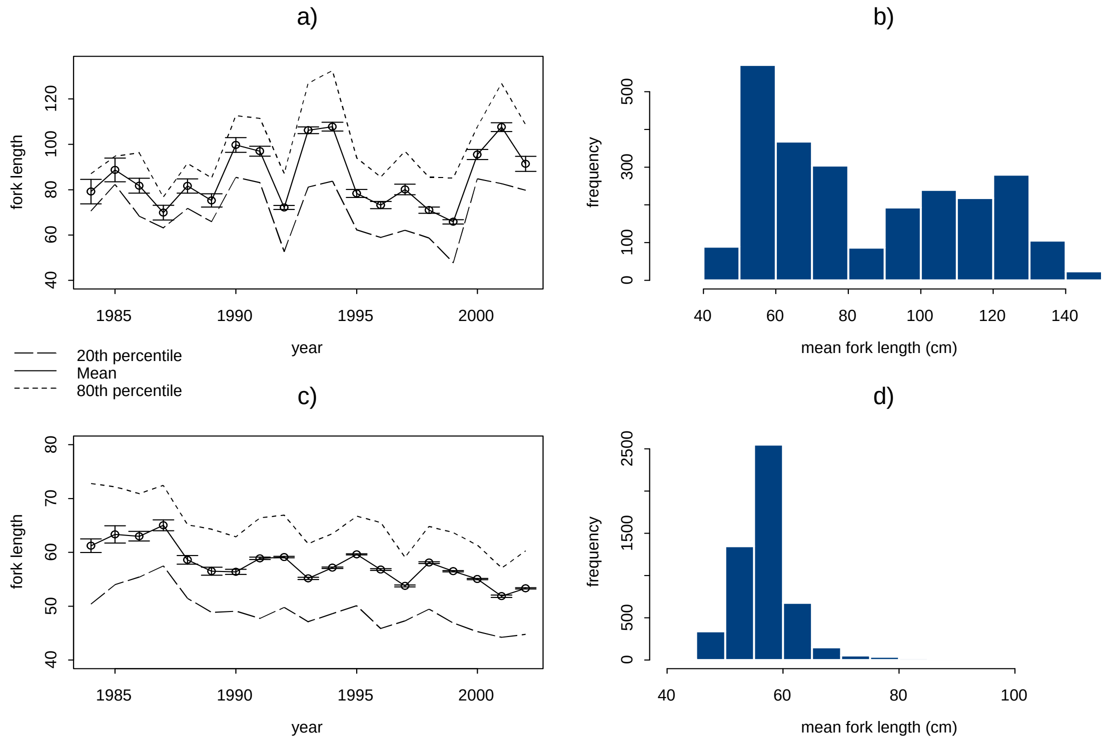


Figure 12: Bigeye, Purse Seine. Free school a) average annual mean, 20th and 80th percentiles of length, and b) histogram of mean lengths. Associated school c) average annual mean, 20th and 80th percentiles of length, and d) histogram of mean lengths. Mean lengths are calculated for each available month/year/5-degree square/fleet. combination for Spanish, French and NEI-EUR fleets, across all areas fished. Error bars are ±1 s.e.

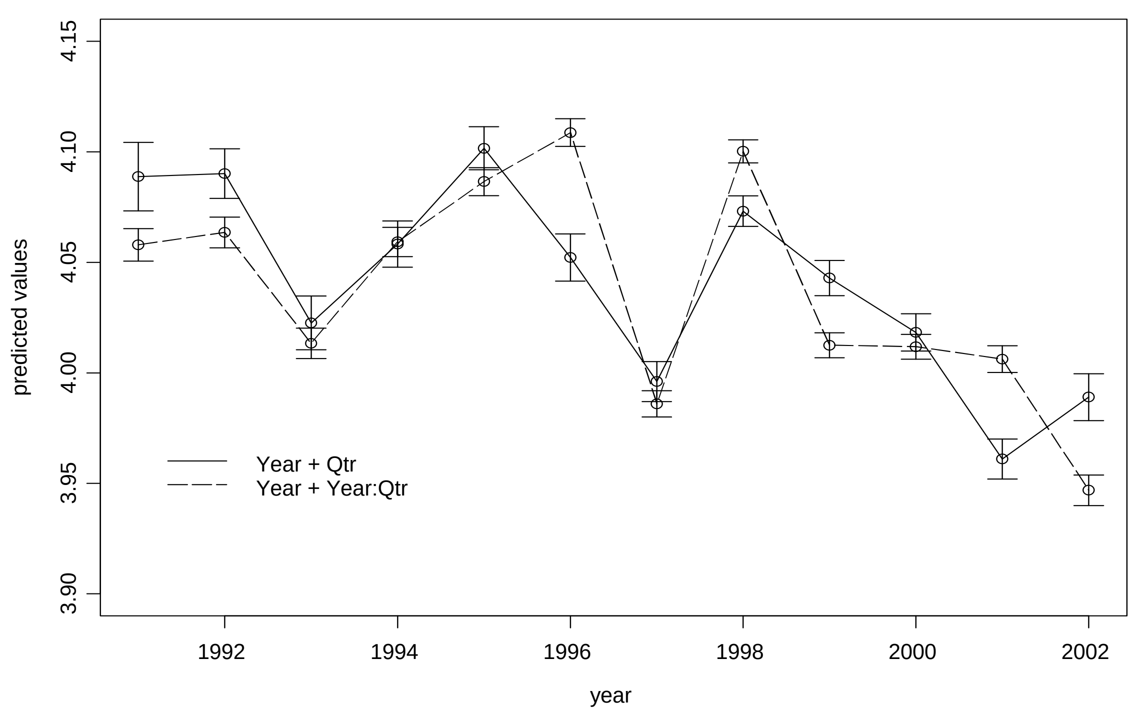


Figure 13: Bigeye, Purse Seine. Time series of standardized mean length indices, on the log scale, for each of the two models (see text). Error bars are ±1 s.e.

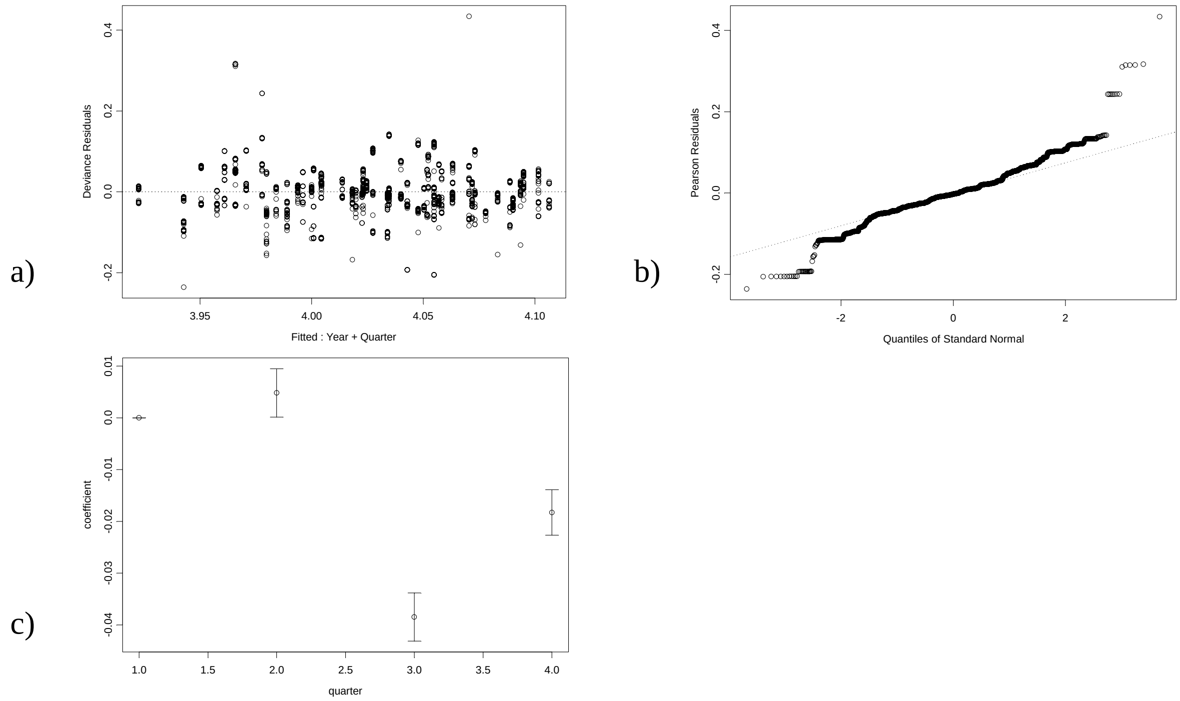


Figure 14: Bigeye, Purse Seine. Plots of a) deviance residuals vs. fitted values, b) Pearson residuals vs. quantiles of standard normal, c) quarter coefficients, for the bigeye GLM model 1. Error bars are ±1 s.e.

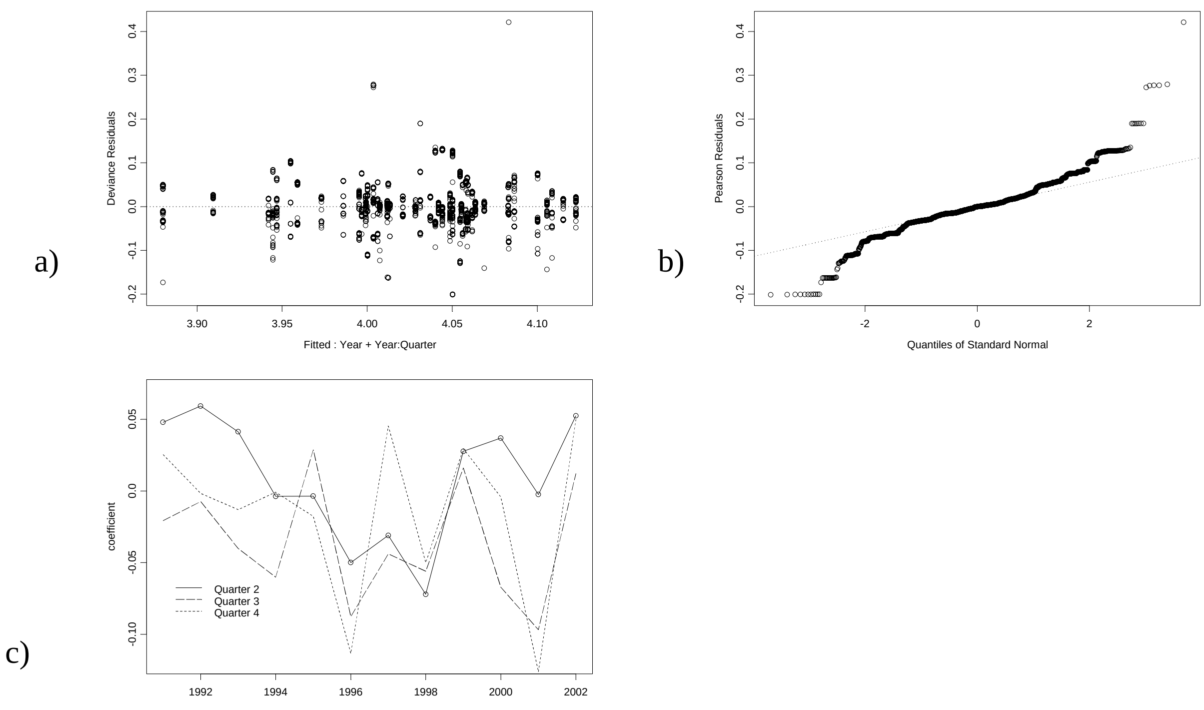


Figure 15: Bigeye, Purse Seine. Plots of a) deviance residuals vs. fitted values, b) Pearson residuals vs. quantiles of standard normal, c) quarter coefficients, for the bigeye GLM model 2. Error bars are ±1 s.e.

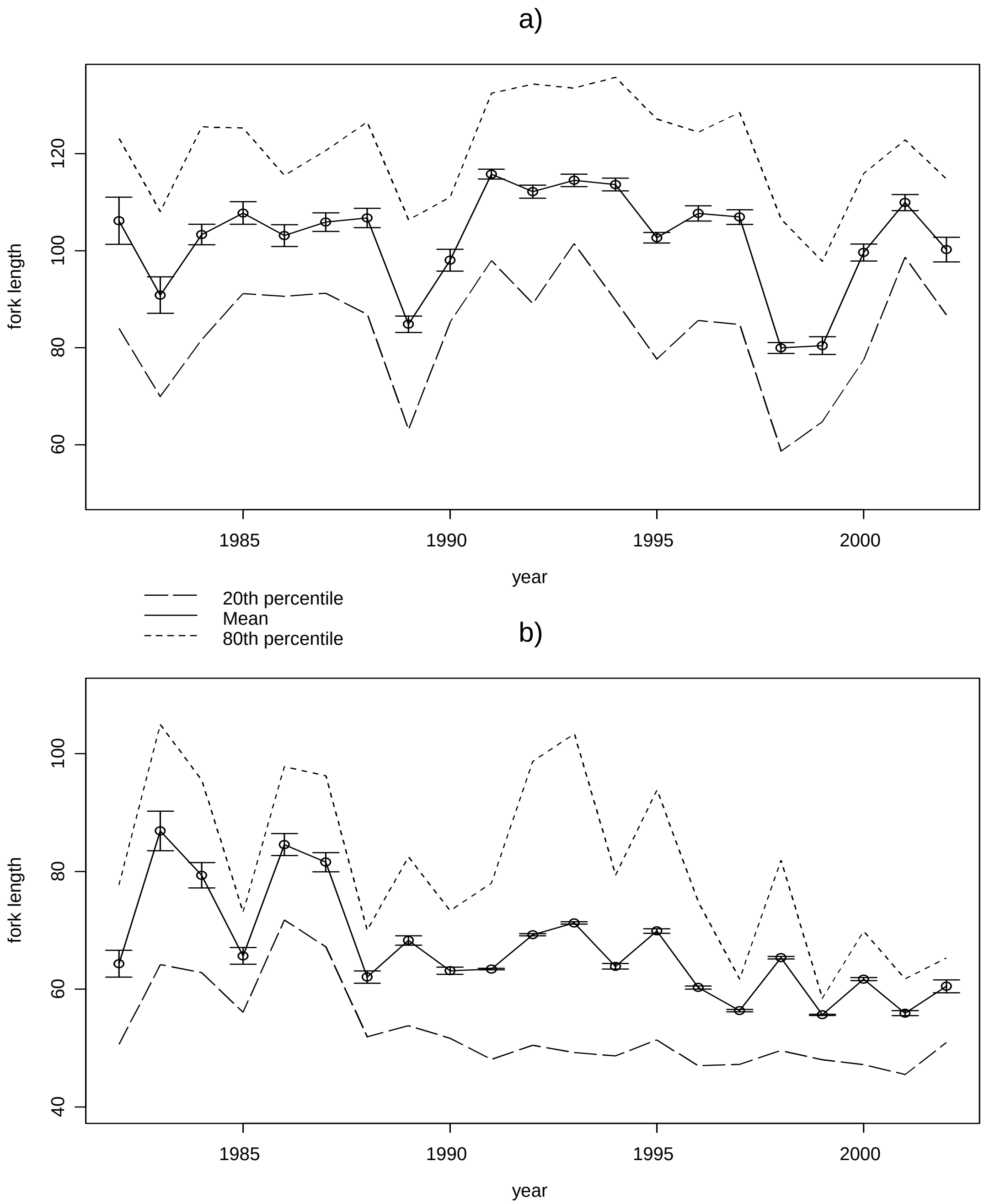


Figure 16: Yellowfin, Purse Seine. Annual averages of mean length and 20th & 80th percentiles (FL in cm) combined Spanish, French and NEI-EUR fleets for a) free schools and b) log schools. Error bars are ±1 s.e.

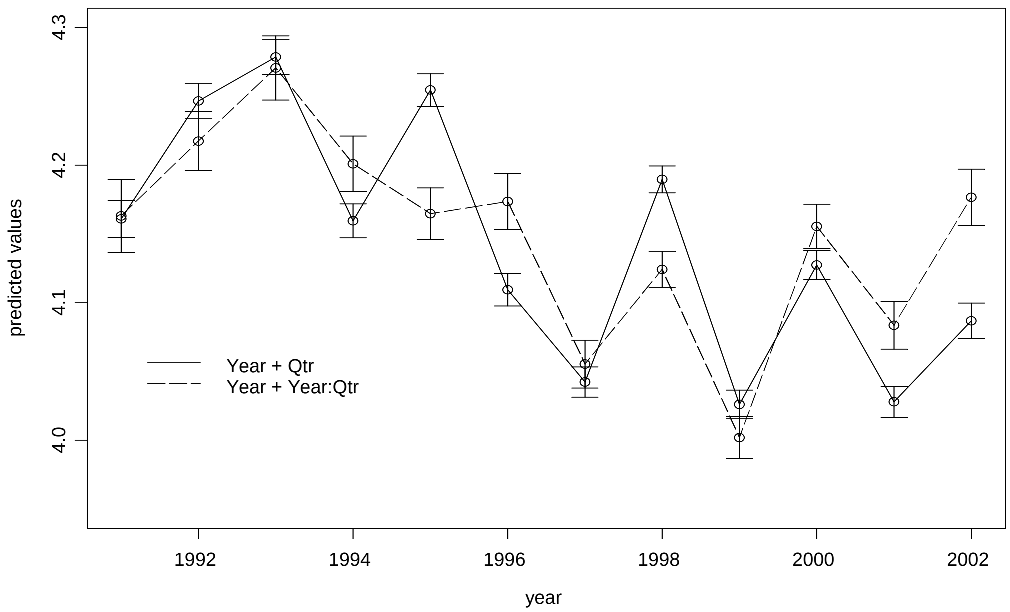


Figure 17: Yellowfin, Purse Seine. Time series of standardized LS mean length indices (on the log scale) for each of the two models (see text). Error bars are ±1 s.e.

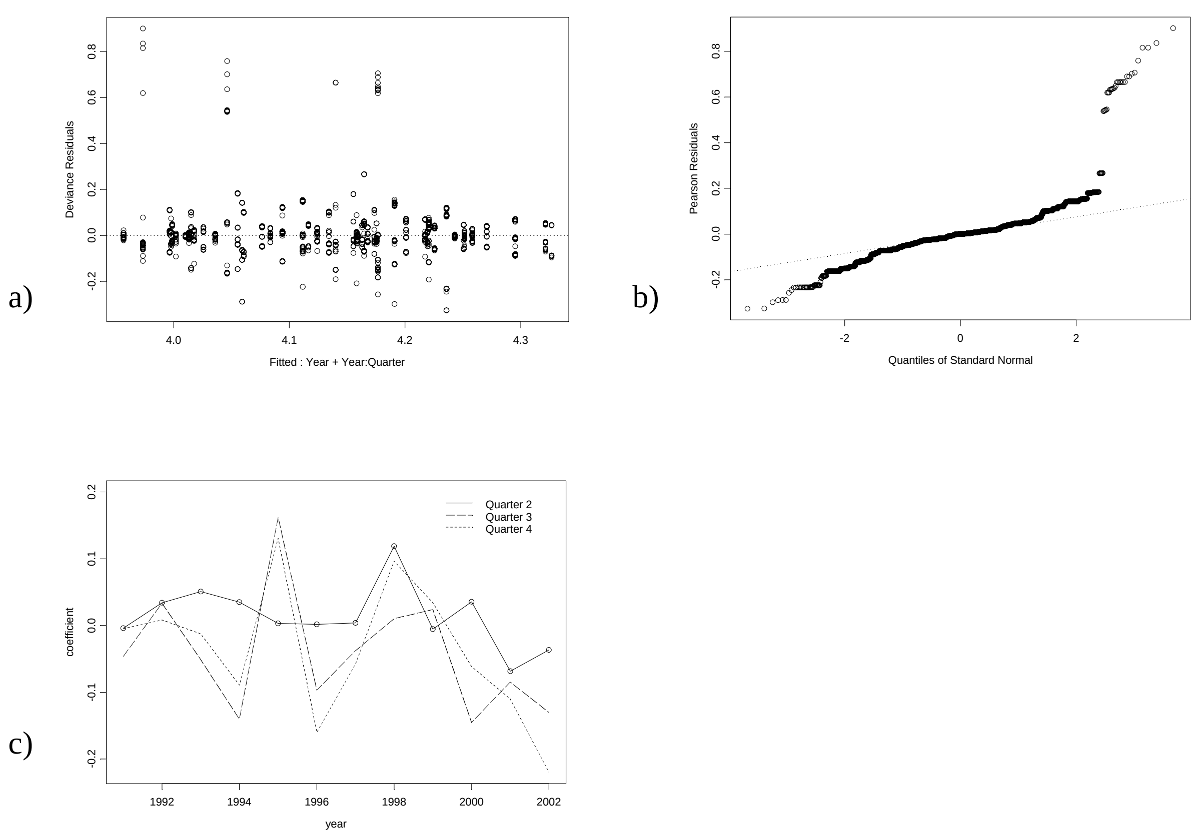


Figure 18: Yellowfin, Purse Seine. LS data. Plots of a) deviance residuals vs. fitted values, b) Pearson residuals vs. quantiles of standard normal, c) quarter coefficients, for the yellowfin GLM model 2. Error bars are ±1 s.e.