Some exploratory analyses on age-based longline CPUE of southern bluefin tuna

Norio Takahashi and Tomoyuki Itoh
National Research Institute of Far Seas Fisheries

Summary: Simple exploratory analyses of age-based CPUE were conducted ('age' used as a proxy of 'size'), concerning: 1) validation of the age aggregation and standardization model used for age-based CPUE presented in Itoh and Takahashi (2015); and 2) examination of effect of age on the year*area interaction in CPUE. Results from the analyses supported validation of the age aggregation and standardization model used for the age-based CPUE. Simple GLM analysis suggested that inclusion of the quadric term of average age and its interaction with area into the current Base model for core vessels CPUE index was worth to be considered to cope with impact of age (or size) on CPUE standardization. Results and considerations in this paper are not definitive and further exploration, analysis, and discussion in the CPUE modeling group are necessary for tackling issues of age-based (or size-based) CPUE and a role of age (or size) in CPUE calculation.

要約：以下に関連して、年齢ベース CPUE の簡単な探索的解析を行った（'年齢' を '体長' の代用として用いて）：1）伊藤・高橋（2015）で提示された年齢ベース CPUE に使用された年齢アグリゲーションおよび標準化モデルの妥当性；2）CPUE における年齢の年*海区交互作用への影響の検討。解析結果は年齢ベース CPUE に使用された年齢アグリゲーションおよび標準化モデルの妥当性を支持した。単純な GLM 解析は、年齢（あるいは体長）の CPUE 計算への影響に対処するため、コア船 CPUE 指数の現在の Base モデルへ平均年齢の 2 次項とそれの海区との交互作用を取り込むことは検討に値することを示唆した。この文書で示した結果や検討事項は最終的なものではなく、年齢ベース（あるいは体長ベース）の CPUE の問題、および年齢（あるいは体長）の CPUE 計算における役割に関する問題に取り組むためには、CPUE モデリンググループで今後も探索、解析、議論をしていく必要がある。

Introduction

The CPUE modeling group of the CCSBT Scientific Committee has been discussing about development of size-based (or age-based) CPUE for southern bluefin tuna, *Thunnus maccoyii*, and role of size (or age) in CPUE calculations. This paper provides results and considerations from some simple exploratory analyses of age-based CPUE. In our analyses, 'age' was regarded as a proxy of 'size' because catch at size and effort data are not currently available. The paper consists of two parts: 1) validation of the age aggregation applied and standardization model used for age-based CPUE presented in Itoh and Takahashi (2015); and 2) examination of effect of age on the year*area interaction in CPUE.

Data used and Method

All analyses were based on the same shot-by-shot 'core vessels' data used for the CPUE indices as in Itoh and Takahashi (2016) except catches were disaggregated by age (0, 1, 2, ... 18, 19, 20+).

(1) Validation of the age aggregation and standardization model in Itoh and Takahashi (2015)

To visually examine the similarity of CPUE trends between neighboring ages, the core vessels catch at age (CAA) data were aggregated by year and (CCSBT statistical) area and nominal CPUE trends by area were plotted on graphs. In addition, to identify an age break that separates age groups having resembling CPUE trends, Pearson's correlation coefficients between two CPUE series of adjacent ages by area were calculated. From these two simple analyses, proper age groups that can be aggregated with respect to similarity of CPUE trends were determined.

In order to consider what explanatory variables need to be included in a standardization model for age-based CPUE, the core vessels CAA data were aggregated by year, month, area, latitude (lat5) and longitude (lon5), and nominal CPUE values for the age groups identified above were plotted against possible explanatory variables (e.g., lat5) to examine the relationship between them. Possible explanatory variables to be considered were those included in the Base model in Itoh and Takahashi (2015) as described below:

log(CPUE.age4p+0.2) ~ year + month + area + lat5 + Bigeye.CPUE + Yellowfin.CPUE + month:area + year:lat5 + year:area

(2) Examination of effect of age on the year*area interaction in CPUE

For visually investigating distribution patterns of interaction of CPUE for age 4+, 5x5 degree square, and average age of the catch within a 5x5 square, the core vessels CAA data were aggregated by year, month, area, lat5, lon5, and nominal CPUE for age 4+ and average age within a square were simultaneously plotted on maps by year. The average age within each 5x5 square was calculated as:

$$Average\ \ age\ \ within\ \ a\ \ 5\times 5\ \ square = \sum_{age=0}^{20+} age \times \frac{Catch_{age}}{Total\ \ Catch}, \qquad when\ \ Total\ \ Catch \neq 0$$

Based on consideration of these maps, possible standardization models including the average age term were explored by simple GLM analysis.

Results and Discussion

(1) Validation of the age aggregation and standardization model in Itoh and Takahashi (2015)

<u>Age aggregation</u>

Trends of nominal CPUE for each age (from age 4) by area were shown in Fig. 1. CPUE for age 10 and younger, and age 11 and older were separately plotted on different graphs because of scale differences of CPUE values. CPUE trends of two neighboring ages appear to be similar although there are exceptions for different ages and in different area. Here note that small fish (mostly fish less than 20 kg) have been released/discarded from Japanese vessels in recent years (probably since 2006 when individual quota system was introduced to Japanese SBT longline fishery) and thus CPUE trend of age 4 would be affected by this practice.

Examining individual CPUE trends one by one, partly certain cohorts can be sequentially tracked in trends of different ages. For example, in area 9 in the early 1990s or around 2010, CPUE for ages 4, 5, and 6 sequentially show high values year after year. Such aspect of CPUE trend would be one of reasons why age-based (or size-based) CPUE is worth to pursue.

Fig.2a shows Pearson's correlation coefficients between two CPUE series for adjacent ages. Not surprisingly most of the coefficient values indicate that CPUE trends of two adjacent ages are highly correlated. Correlation of trends between age 4 and 5 is low compared to other pairs. To examine this, correlation coefficients between CPUE for ages 4 and 5 were calculated dividing CPUE before and after 2006 (Fig. 2b). Except areas 5 and 9, the coefficients were lower after 2006, especially noticeable in areas 4, 6, and 7. Therefore, these lower correlations would be partly caused by the release/discard practice of small fish in recent 10 years explained above.

Differences of coefficient values are not large between two adjacent ages (e.g., age 5 and 6 versus age 6 and 7), and the values for correlation of particular adjacent ages are different depending upon areas (e.g., when focusing on age 8 and 9). This makes it difficult to clearly identify age groups which have similar CPUE trends. However, some tendencies can be read off from the graphs having a close look. Although it is different by areas, correlation of CPUE between age 5 and 6 is relatively high. Correlation coefficients from between age 7 and 8 to between age 10 and 11 (or age 11 and 12) are more or less stable at relatively high values, except areas 7 and 8. For age 10 (or age 11) and older, coefficient values vary depending upon areas but CPUE values themselves are small (see the y-axis scale in Fig. 1b). From these considerations, it would be reasonable to treat CPUE aggregating sizes comparable to three age groups of age 5 and 6, from age 7 to 10 (or 11), and age 11 (or age 12) and older. Different patterns of year*area interaction effects on CPUE between age 4 to 10 and age 11 and older shown in Fig. 1 (e.g., compare CPUE trends in area 7) further support a separation of age groups at age 10 (or 11). We omitted age 4 from this age grouping with consideration for a possible impact of the release/discard practice in 10 recent years on its CPUE trend. Trends of nominal CPUE for the three age groups by area were shown in Fig. 3.

Itoh and Takahashi (2015) presented the result from preliminary analysis of standardized CPUE for three age groups, age 5&6, age 7-10 and age 11+ without showing a detailed examination of age aggregation. Our simple analysis supports validity of this age grouping and also suggests that disaggregating age groups further than the three age groups is meaningless because of similarity of CPUE trends among disaggregated groups.

Possible explanatory variables of standardization model

For the three age groups (5&6, 7-10, 11+), nominal CPUE by month and area (month*area), by year and area (year*area), by year and lat5 (year*lat5) were shown in Figs. 4, 5, and 6, respectively. There observed clear interactions between these variables and patterns of the effects vary with age group. For instance, the pattern of CPUE for age group 7-10 by month in area 7 differs from that in area 8 which suggests an interaction (Fig. 4b). This difference seen in age group 7-10 is not same when looking at the pattern of CPUE for age group 11+ by month in area 7 and comparing this to that in area 8 (Figs. 4b and 4c).

For the three age groups, relationships between CPUE of southern bluefin tuna (SBT) and of bigeye tuna (BET), and between CPUE of SBT and of yellowfin tuna (YFT) by lat5 were also shown in Figs. 7 and 8, respectively. Clearly there are particular relationships between CPUE of SBT and of BET, and between CPUE of SBT and of YFT which suggest that CPUEs of BET and of YFT (both species generally inhabit different latitudes or sea temperature ranges from SBT) reflect differences in operational or targeting patterns of Japanese longline vessels in different latitudes.

Examination of these figures (Figs. 4 to 8) suggests that year, month, area, lat5, BET CPUE, YFY CPUE and interactions between month and area, year and area, and year and lat5 are all possible terms to be included in standardization models for age(group)-based CPUE. Therefore, the model ('Base' model) used in the preliminary analysis by Itoh and Takahashi (2015) is considered appropriate to standardize CPUE for age groups 5&6, 7-10, and 11+.

(2) Examination of effect of age on the year*area interaction in CPUE

Distributions of nominal CPUE for age 4+ and average age within 5x5 degree square by year were shown in Fig. 9. Here again, recall that small fish mostly less than 20kg have been released/discarded from Japanese vessels in recent years (probably since 2006). This would affect average age calculation (maybe slight overestimation) and thus there needs some caution when interpreting the average age in the recent 10 years.

As having been discussed in the CPUE modeling group, spatiotemporal interactive dynamics of CPUE and average age (seen as a proxy for size) is also evident in Fig. 9. However, it is not easy to specify patterns of how age (or size) affects on year*area interaction from these figures.

As the first attempt, GLM analysis was conducted applying a simple model. Note that, in this GLM analysis, zero catch records were removed from the 5x5 month CAA core vessels data because there were no age frequency data (and thus CAA data) available for zero catch records. We found somewhat a quadric relationship between nominal CPUE for age 4+ and average age (Fig. 10). After exploring several simple models with some interaction terms, we found the model below worth to be considered:

log(CPUE.age4p) ~ year + area + average.age^2 + area:average.age^2

In this model, year and area terms were treated as categorical and average age term was as continuous. Nominal CPUE by average age and area were shown in Fig. 11. This suggests some interactions between average age and area terms because patterns of CPUE vary depending upon areas. For instance, a quadric relationship between CPUE and average age seen in area 4 appears different from that seen in area 9. Attached below is the result summary of 'glm' function of R software. The coefficient table shows that the quadric term of average age and its interaction with area are all significant.

---------- glm summary start ----------

```
Call:
glm(formula = log(cp4p) ~ fyear + farea + AvAge2 + farea:AvAge2)

Deviance Residuals:
    Min        1Q    Median       3Q      Max
-7.4576   -0.3459   0.1629   0.5509   3.2816

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.285825   0.209255   1.366 0.172182
fyear1987       0.150172   0.207522   0.724 0.469403
fyear1988      -0.154725   0.212536  -0.728 0.466735
fyear1989      -0.112303   0.207323  -0.542 0.588125
fyear1990       0.185582   0.203234   0.913 0.361320
fyear1991      -0.216033   0.199483  -1.083 0.279008
fyear1992       0.062184   0.204614   0.304 0.761242
fyear1993       0.156873   0.208267   0.753 0.451438
fyear1994       0.432808   0.218515   1.981 0.047820 *
fyear1995       0.403128   0.218881   1.842 0.065718 .
fyear1996       0.167073   0.218410   0.765 0.444429
fyear1997      -0.035003   0.206535  -0.169 0.865444
fyear1998       0.014408   0.209275   0.069 0.945120
fyear1999       0.100442   0.206383   0.487 0.626561
fyear2000      -0.004197   0.212759  -0.020 0.984263
fyear2001       0.339511   0.211821   1.603 0.109197
fyear2002       0.442603   0.226175   1.957 0.050554 .
fyear2003       0.091388   0.224599   0.407 0.684146
fyear2004       0.249151   0.213345   1.168 0.243071
fyear2005      -0.123464   0.216313  -0.571 0.568248
fyear2006      -0.152696   0.221887  -0.688 0.491456
fyear2007      -0.779248   0.232001  -3.359 0.000804 ***
fyear2008      -0.078398   0.223540  -0.351 0.725858
fyear2009       0.185066   0.231479   0.799 0.424138
fyear2010       0.178476   0.234602   0.761 0.446926
fyear2011       0.003237   0.228986   0.014 0.988723
fyear2012       0.474005   0.226397   2.094 0.036465 *
fyear2013       0.270966   0.235480   1.151 0.250051
fyear2014       0.777489   0.239520   3.246 0.001198 **
fyear2015       0.803994   0.228034   3.526 0.000436 ***
farea7          0.913342   0.232588   3.927 9.02e-05 ***
farea8          0.371697   0.185374   2.005 0.045139 *
farea9          0.532124   0.177824   2.992 0.002816 **
farea56        -0.275924   0.239089  -1.154 0.248668
AvAge2         -0.021579   0.002256  -9.565  < 2e-16 ***
farea7:AvAge2   0.011915   0.003689   3.229 0.001269 **
```

```
farea8:AvAge2    0.015780   0.002791   5.654 1.89e-08 ***
farea9:AvAge2    0.017066   0.002756   6.191 7.79e-10 ***
farea56:AvAge2   0.020176   0.002821   7.152 1.37e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 1.192928)

    Null deviance: 2250.9  on 1457  degrees of freedom
Residual deviance: 1692.8  on 1419  degrees of freedom
AIC: 4435.3

Number of Fisher Scoring iterations: 2
---------- glm summary end ----------
```

One immediate idea come to mind from this simple GLM analysis is inclusion of the quadric term of average age (or average size?) within 5x5 square and its interaction with area term into the current Base model for the core CPUE index. A problem here is that there are no age frequency data available for squares without catch (0 catch). To standardize core CPUE by a model with these terms, some assumption needs to be made on average age for squares with 0 catch.

Results and considerations presented in this paper are not definitive and further exploration, analysis, and discussion in the CPUE modeling group are necessary for tackling issues of age-based (or size-based) CPUE and a role of age (or size) in CPUE calculation.

References

Itoh T, Takahashi N (2015) Standardized CPUE by age group in SBT. CCSBT/CPUE2015/03.

Itoh T, Takahashi N (2016) Update of the core vessel data and CPUE for southern bluefin tuna in 2016. CCSBT-ESC/1609/**.

(a) From age 4 to age 10

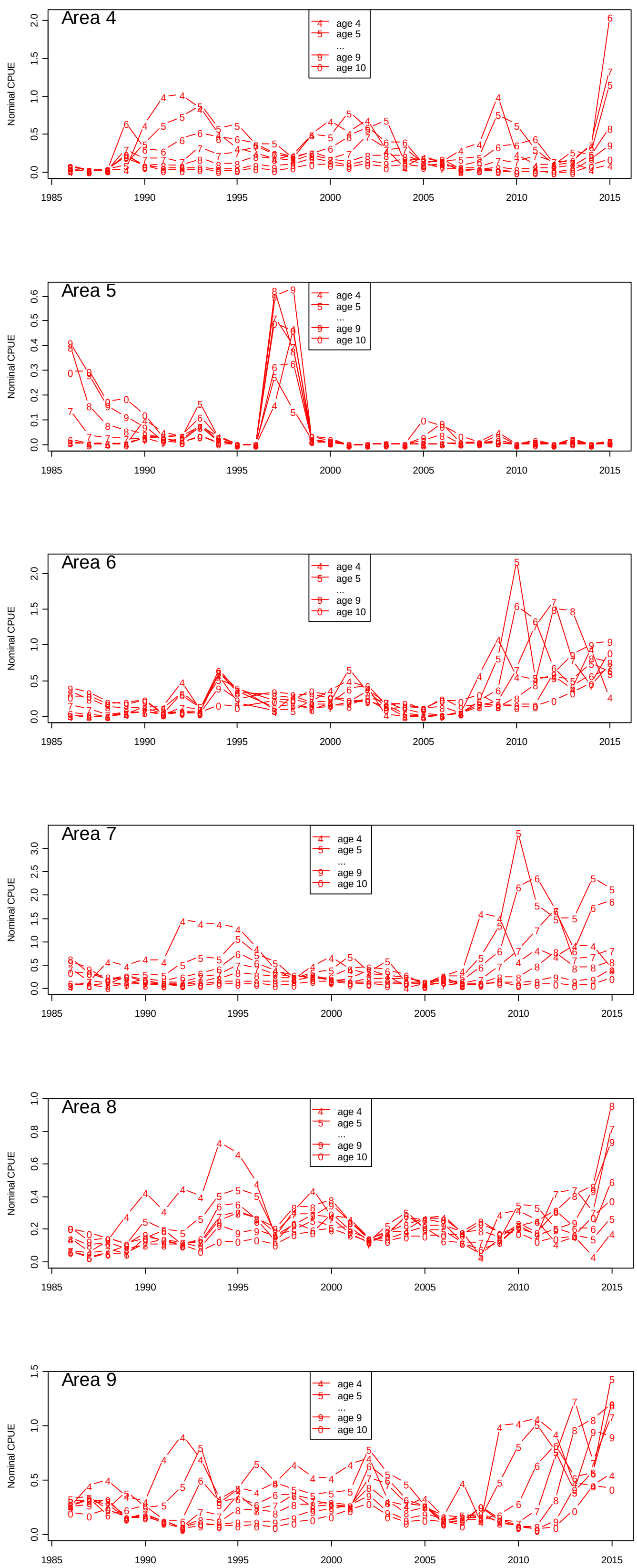


Fig. 1. Nominal CPUE for each age (from age 4) by area. (a) From age 4 to age 10 (b) From age 11 to age 20+).

(b) From age 11 to age 20+

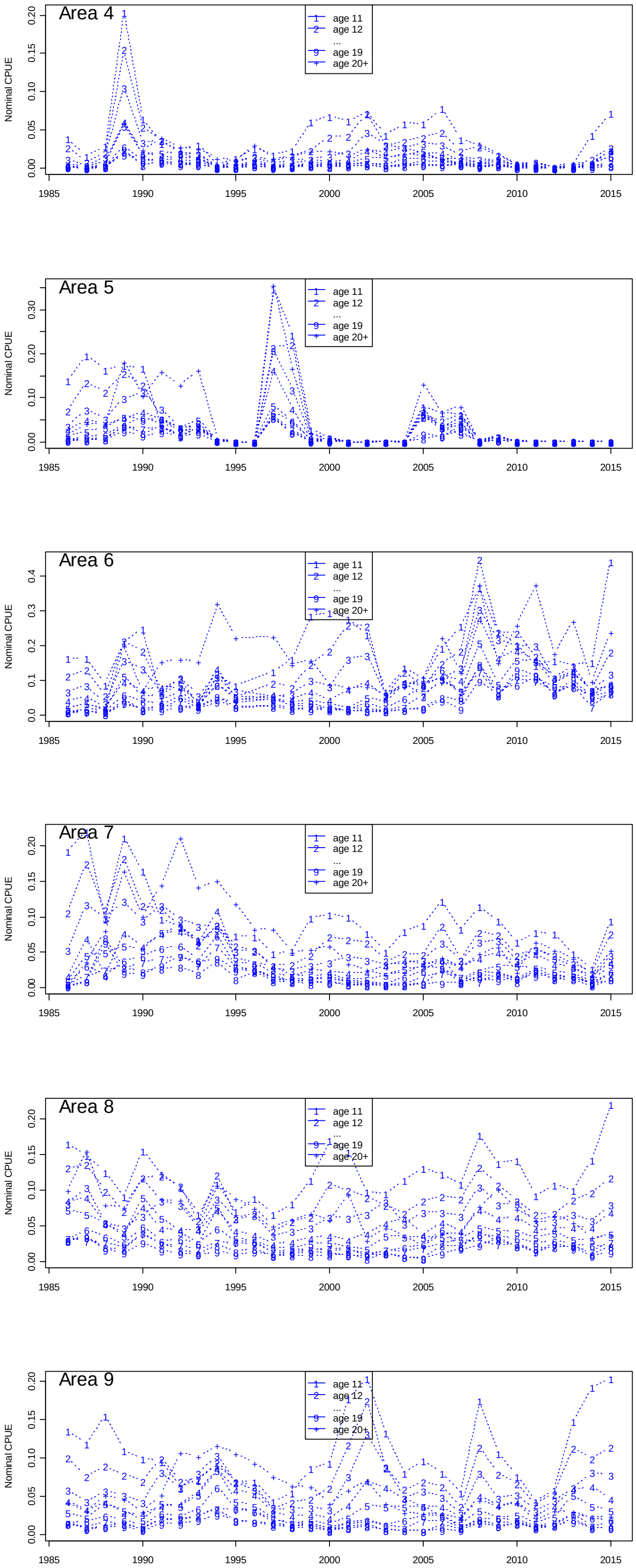


Fig. 1 (cont'd). Nominal CPUE for each age (from age 4) by area. (a) From age 4 to age 10 (b) From age 11 to age 20+).

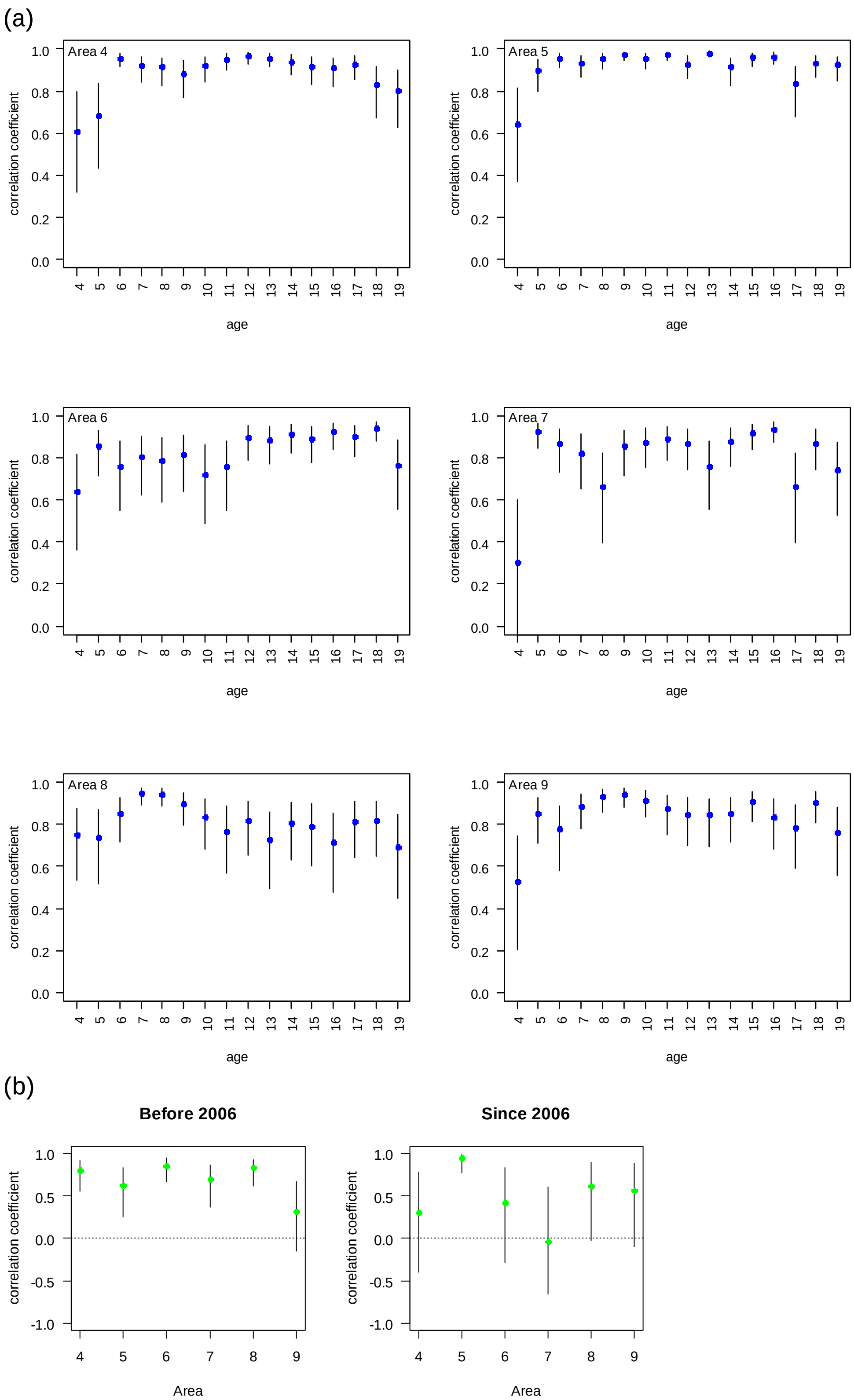


Fig. 2. (a) Pearson's correlation coefficients between two CPUE series for adjacent ages by area. Vertical lines indicate confidential intervals. For example, '8' on the x-axis indicates a correlation coefficient between CPUE series for ages 8 and age 9. (b) The correlation coefficients between CPUE series for ages 4 and 5 only. Correlations were examined dividing CPUE series before and after 2006. Vertical lines indicate confidential intervals. Note that the x-axis indicates area in this case.

(a) age groups 5&6, and 7-10

Fig. 3. Nominal CPUE for age groups 5&6, 7-10 and 11+ by area. (a) age groups 5&6 and 7-10 (b) age group 11+).

(b) age group 11+

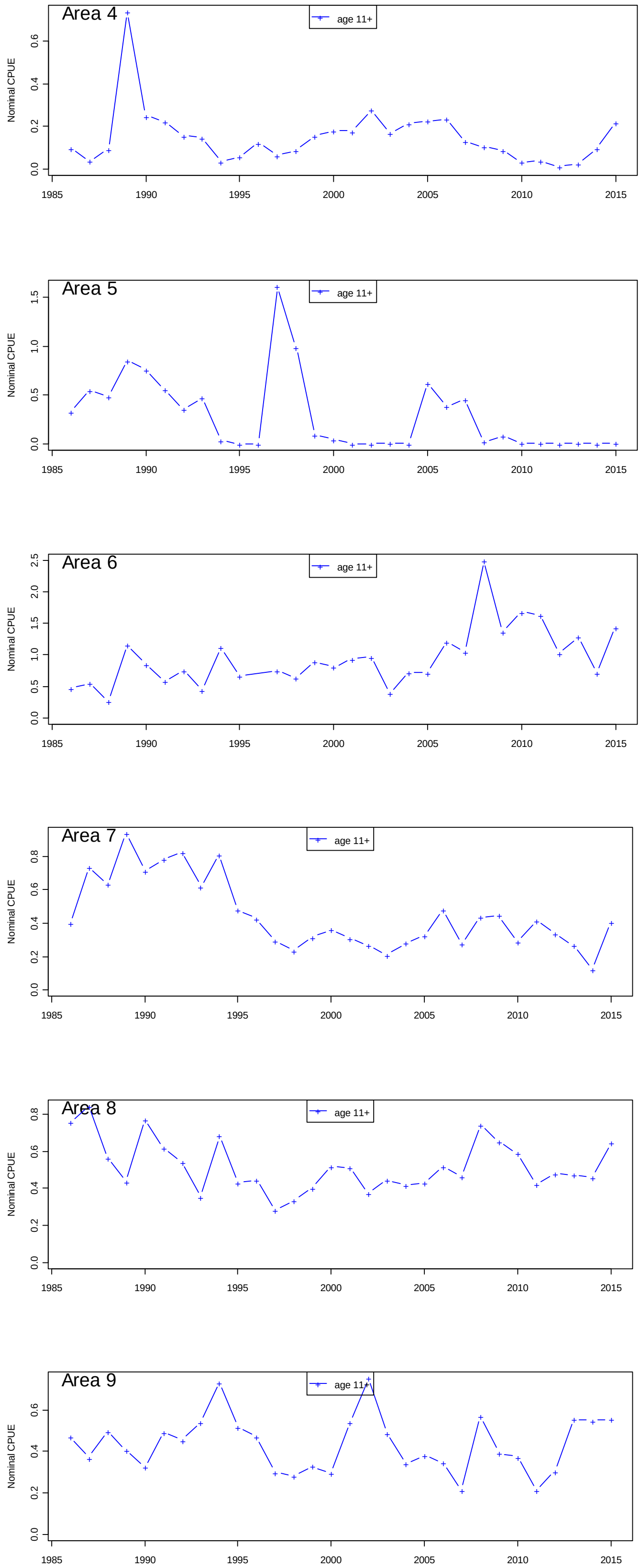
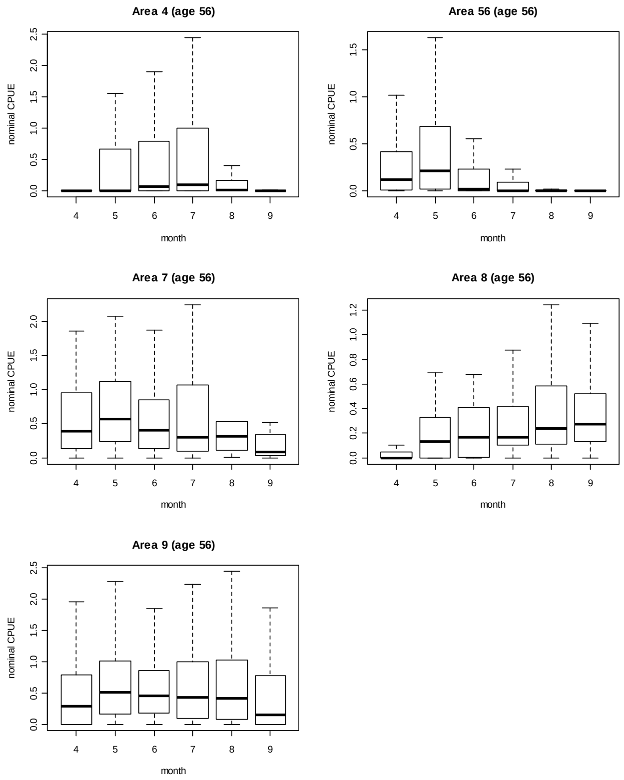
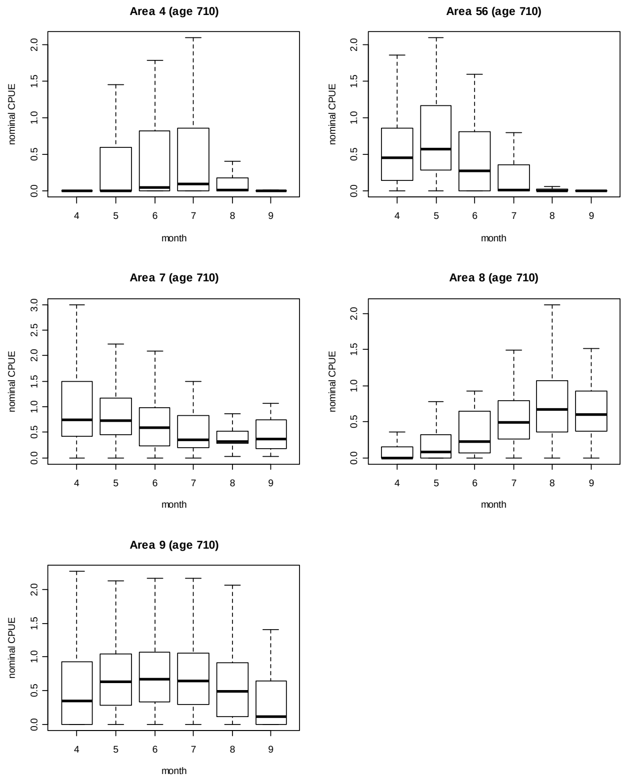


Fig. 3 (cont'd). Nominal CPUE for age groups 5&6, 7-10 and 11+ by area. (a) age groups 5&6 and 7-10 (b) age group 11+).

(a) age group 5&6



(b) age group 7-10



(c) age group 11+

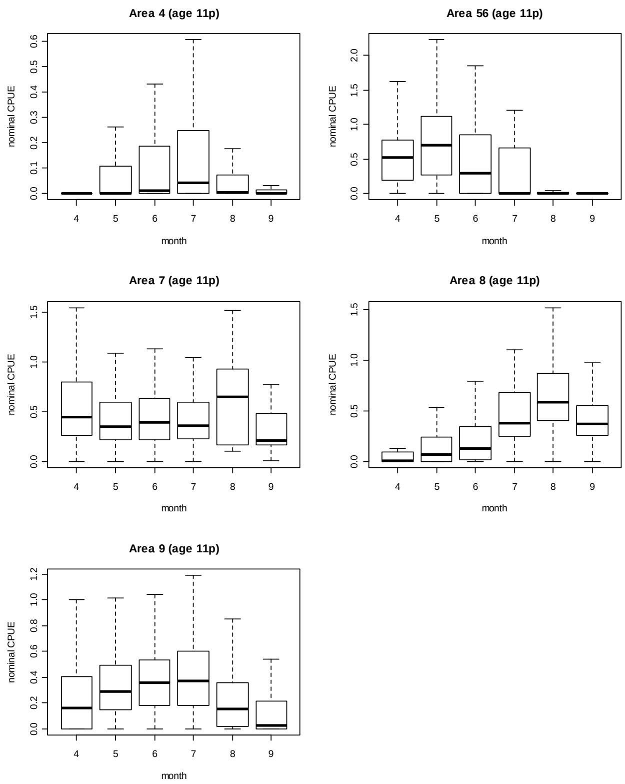
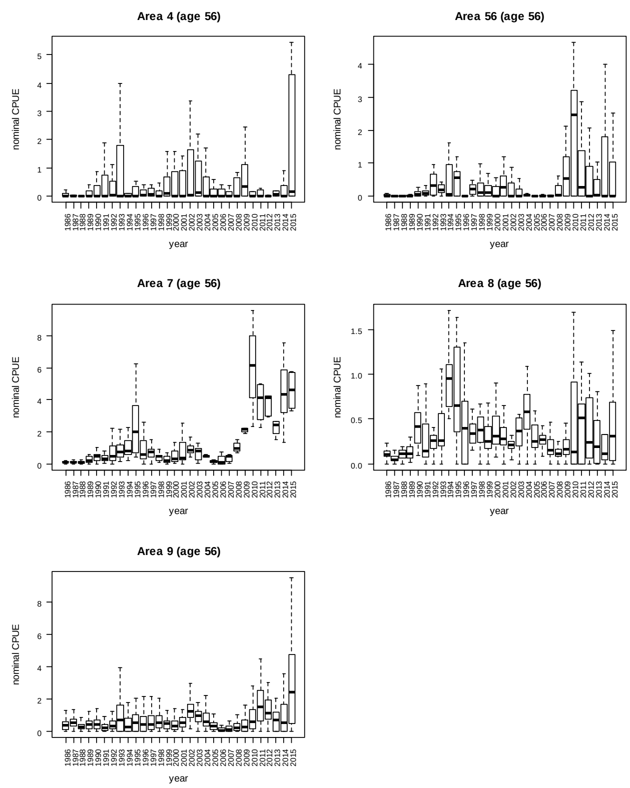
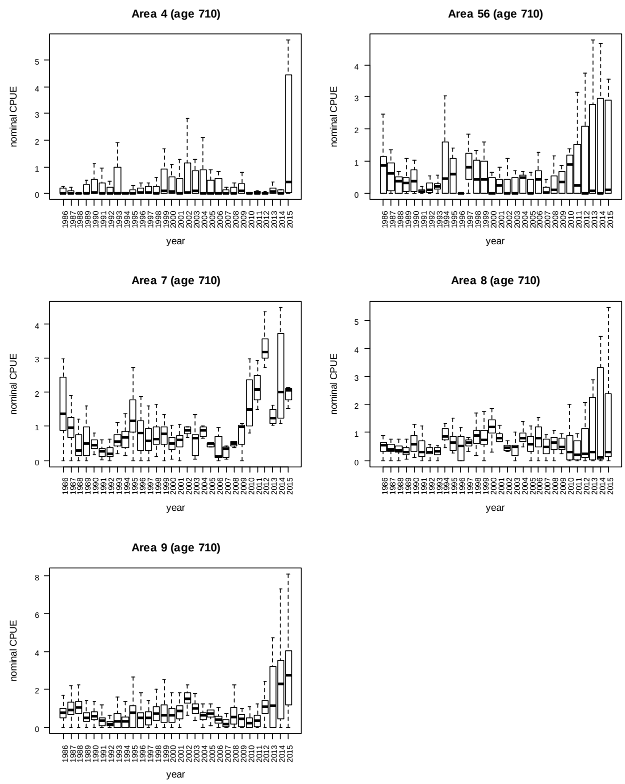


Fig. 4. Nominal CPUE by month and area (month*area interaction). (a) age group 5&6 (b) age group 7-10 (c) age group 11+.

(a) age group 5&6



(b) age group 7-10



(c) age group 11+

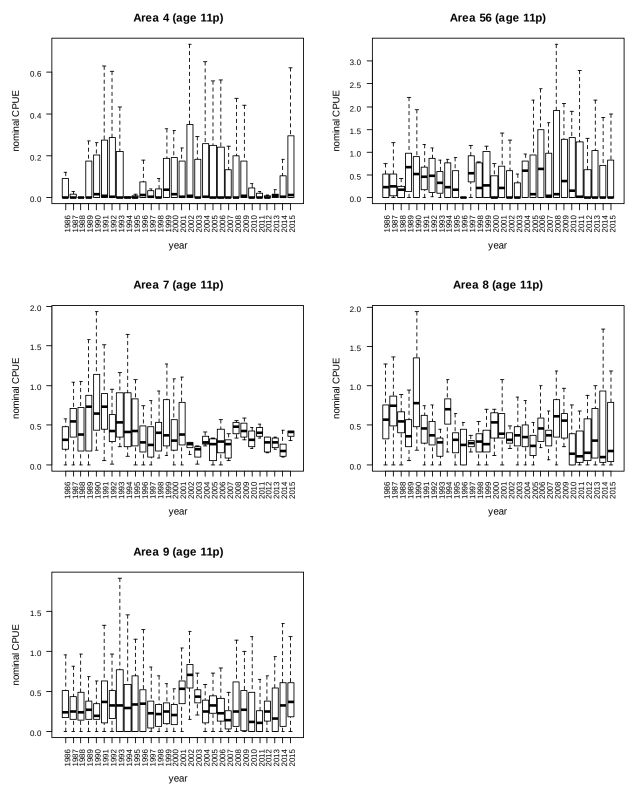
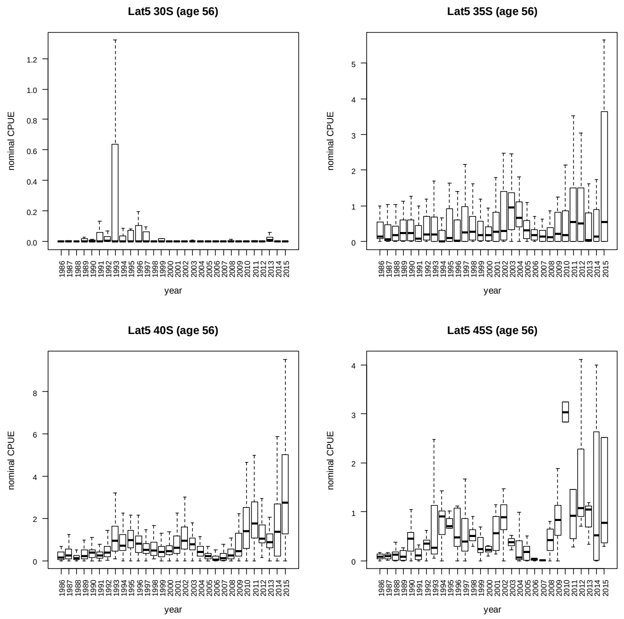
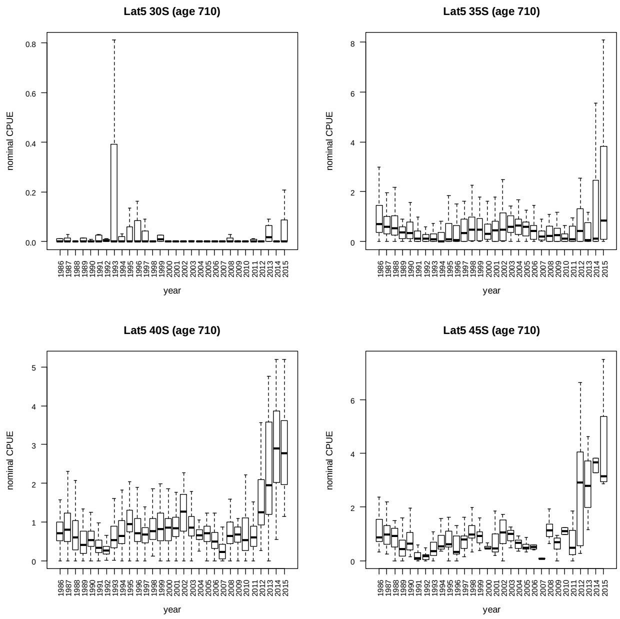


Fig. 5. Nominal CPUE by year and area (year*area interaction). (a) age group 5&6 (b) age group 7-10 (c) age group 11+.

(a) age group 5&6



(b) age group 7-10



(c) age group 11+

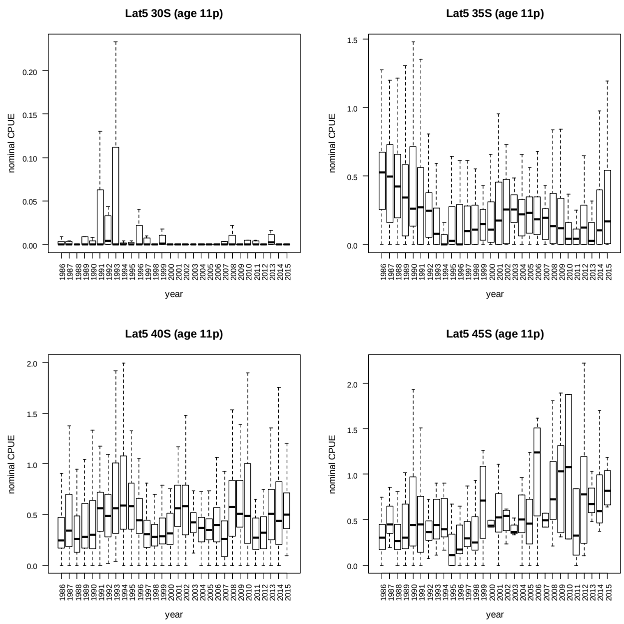


Fig. 6. Nominal CPUE by year and lat5 (year*lat5 interaction). (a) age group 5&6 (b) age group 7-10 (c) age group 11+.

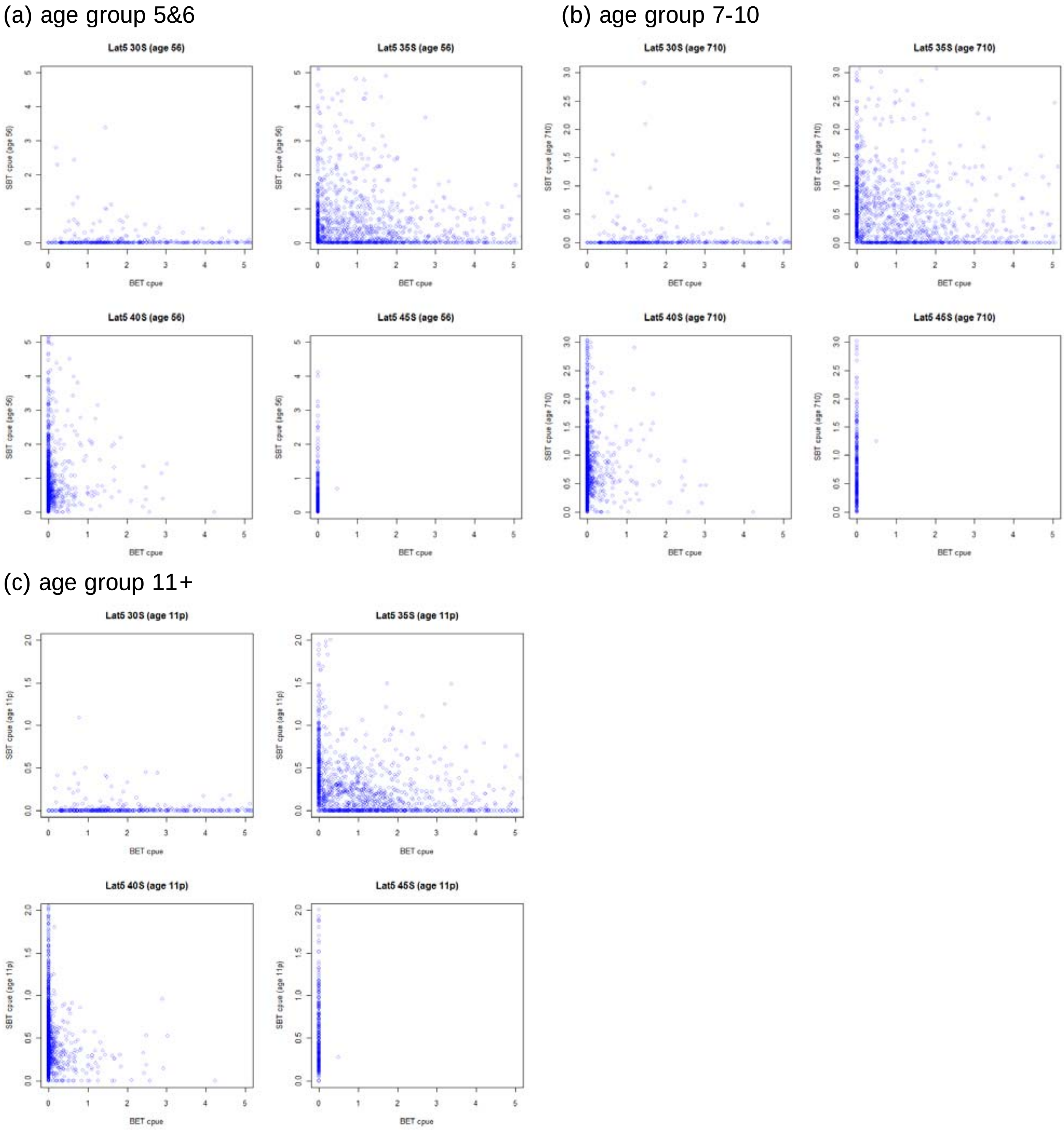


Fig. 7. Relationships between nominal CPUE of southern bluefin tuna (SBT) and of bigeye tuna (BET) by 5 degree latitudinal band (lat5). (a) age group 5&6 (b) age group 7-10 (c) age group 11+.

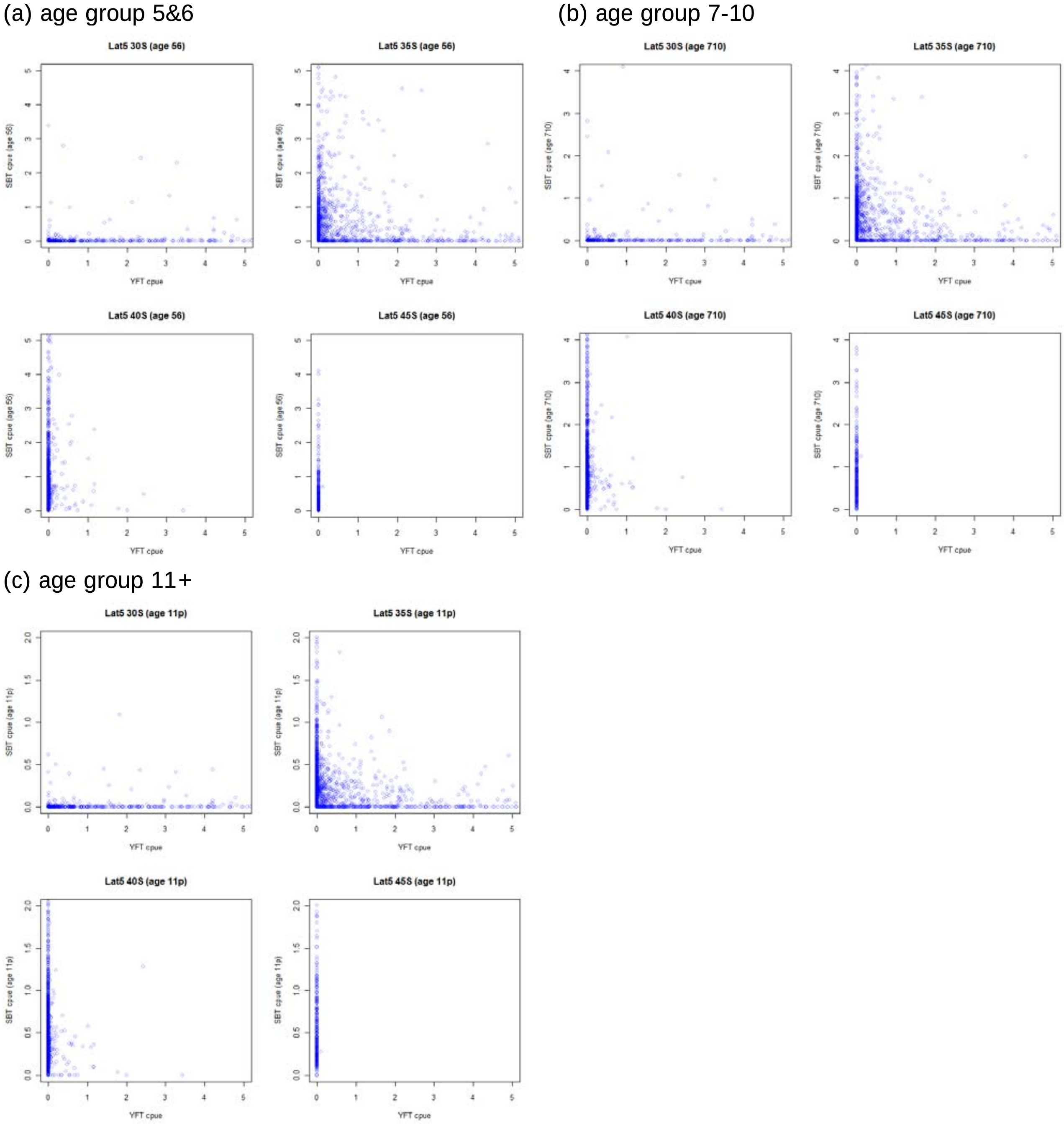


Fig. 8. Relationships between nominal CPUE of southern bluefin tuna (SBT) and of yellowfin tuna (YFT) by 5 degree latitudinal band (lat5). (a) age group 5&6 (b) age group 7-10 (c) age group 11+.

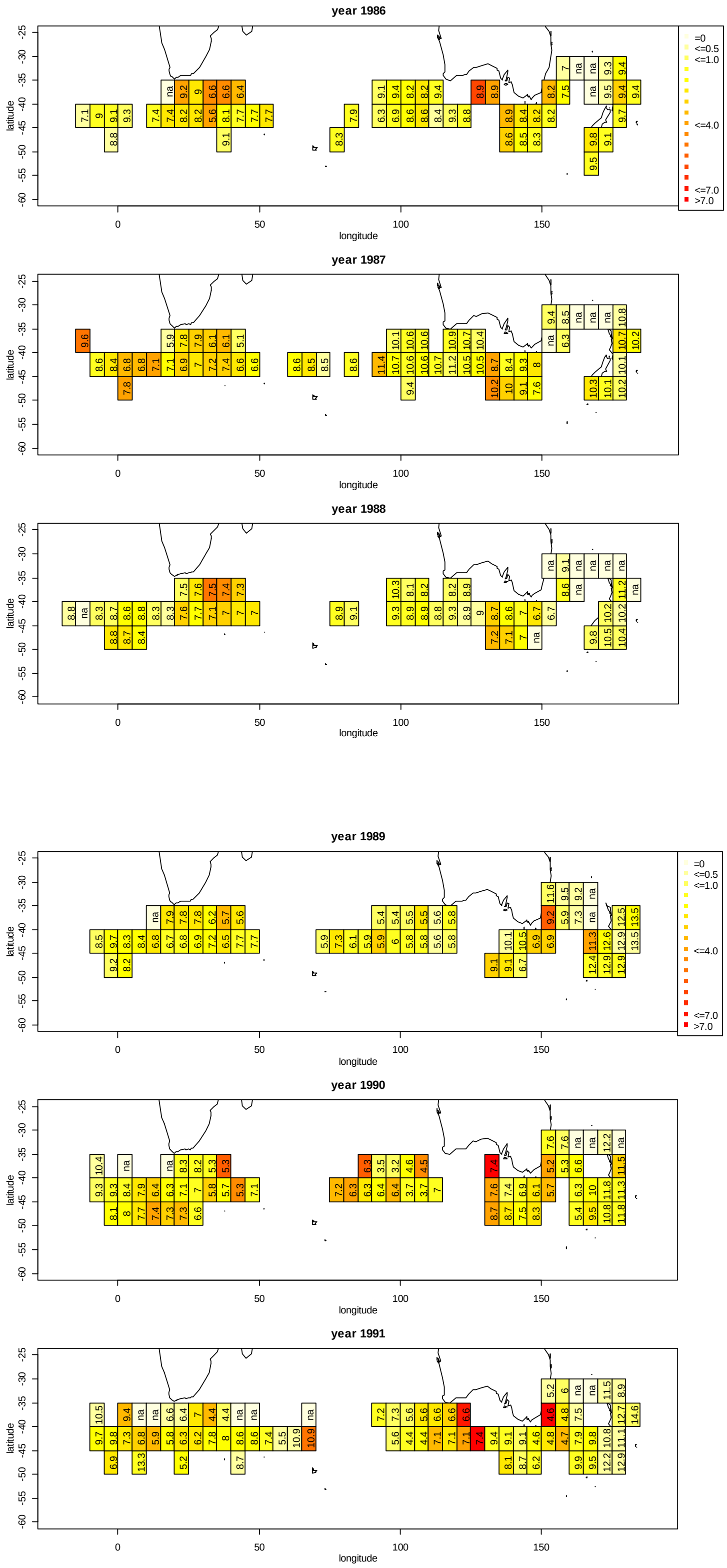


Fig. 9. Distributions of nominal CPUE for age 4+ and average age within 5x5 square by year. For each square, CPUE is represented by color spectrum and a number indicates the average age. Note that ‘na’ means that catch within a square was 0 and no age frequency data available.

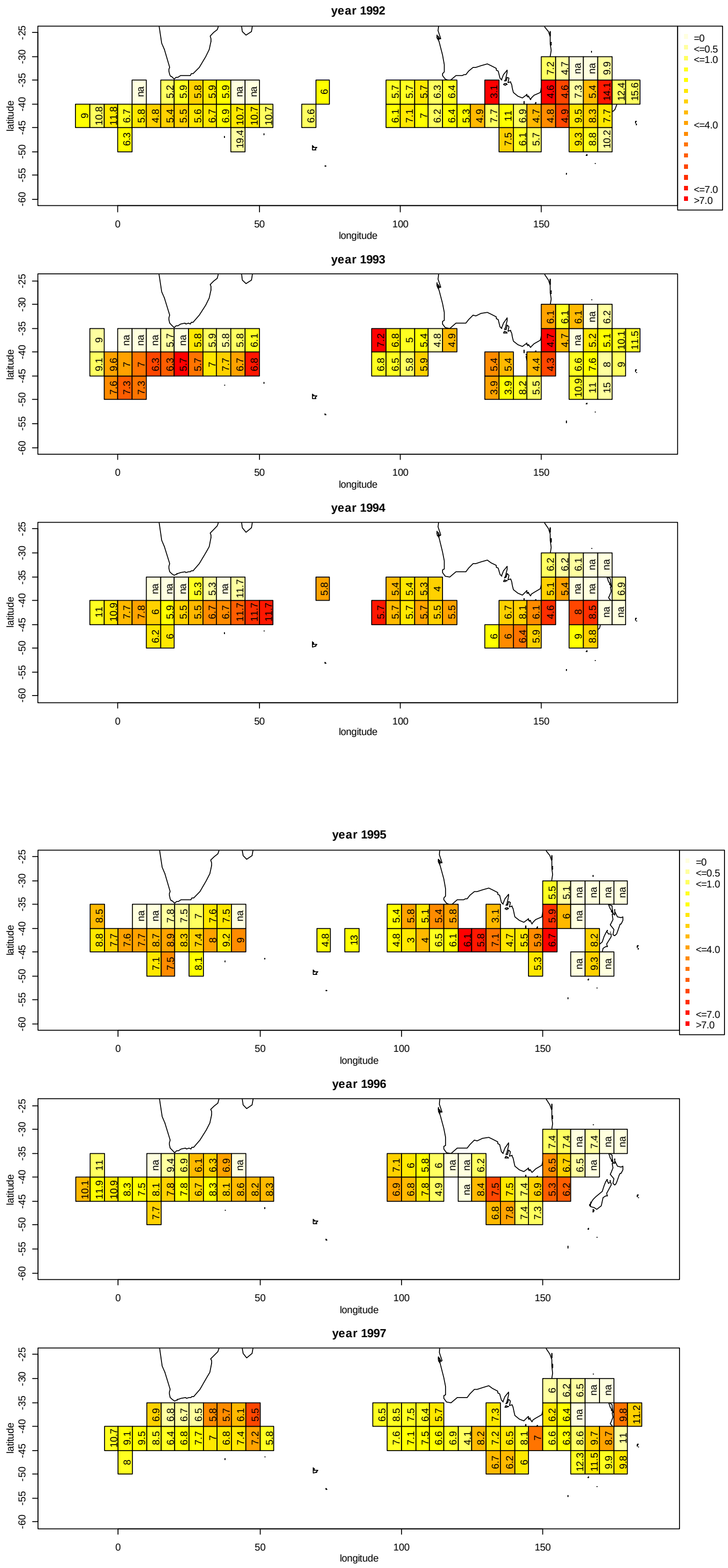
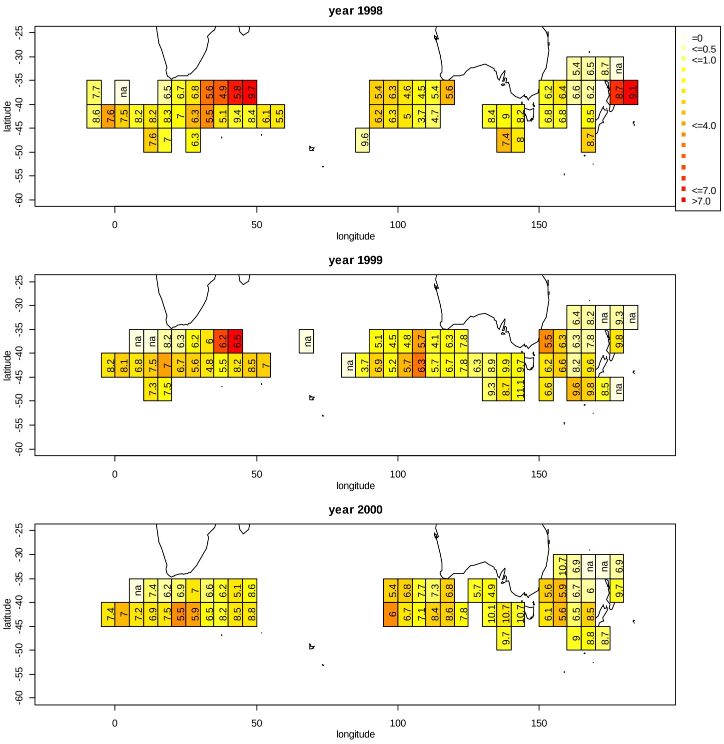


Fig. 9 (cont'd). Distributions of nominal CPUE for age 4+ and average age within 5x5 square by year. For each square, CPUE is represented by color spectrum and a number indicates the average age. Note that 'na' means that catch within a square was 0 and no age frequency data available.



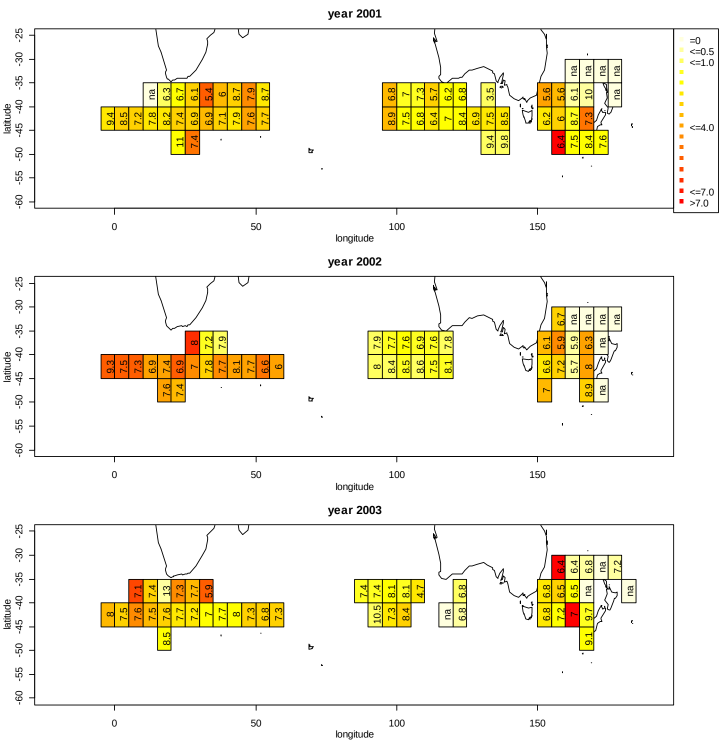


Fig. 9 (cont'd). Distributions of nominal CPUE for age 4+ and average age within 5x5 square by year. For each square, CPUE is represented by color spectrum and a number indicates the average age. Note that 'na' means that catch within a square was 0 and no age frequency data available.

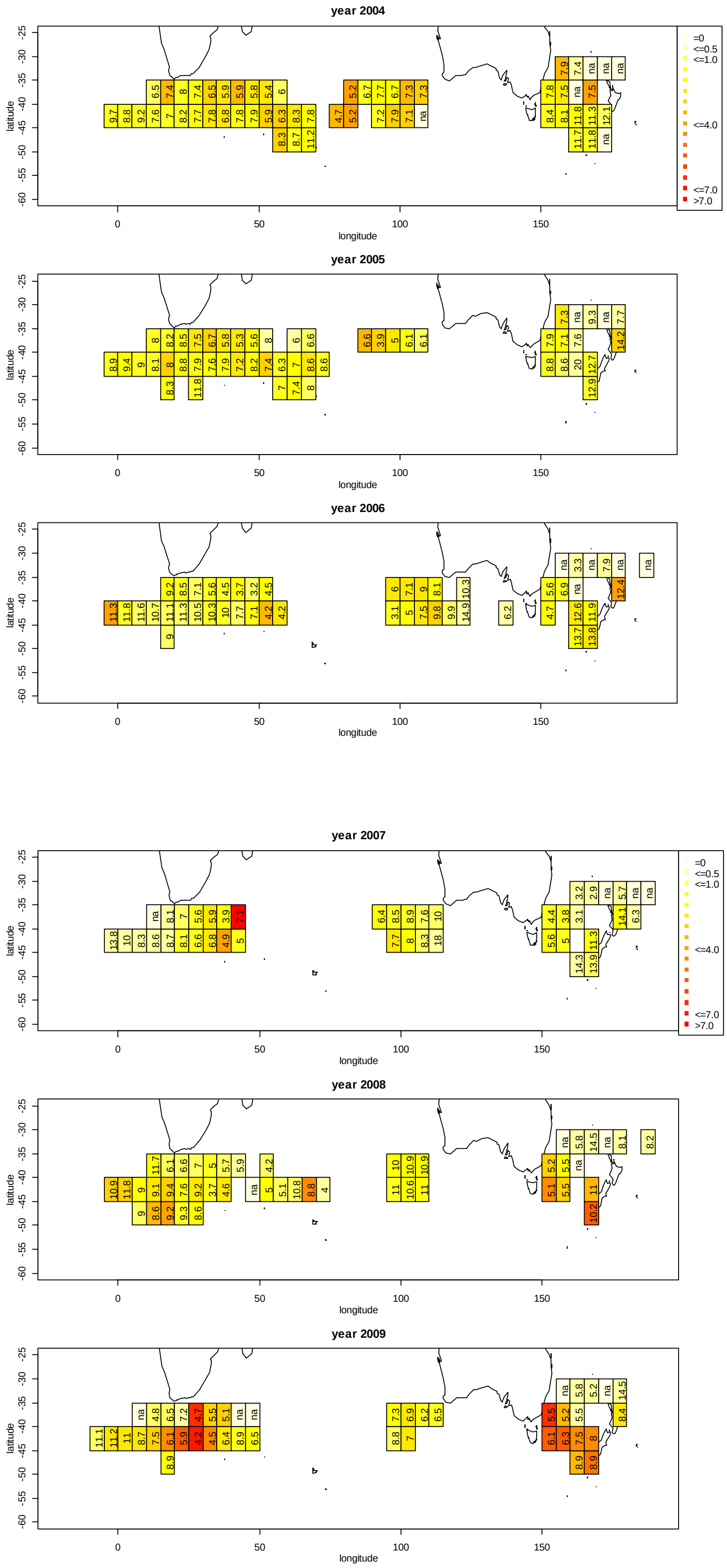


Fig. 9 (cont'd). Distributions of nominal CPUE for age 4+ and average age within 5x5 square by year. For each square, CPUE is represented by color spectrum and a number indicates the average age. Note that 'na' means that catch within a square was 0 and no age frequency data available.

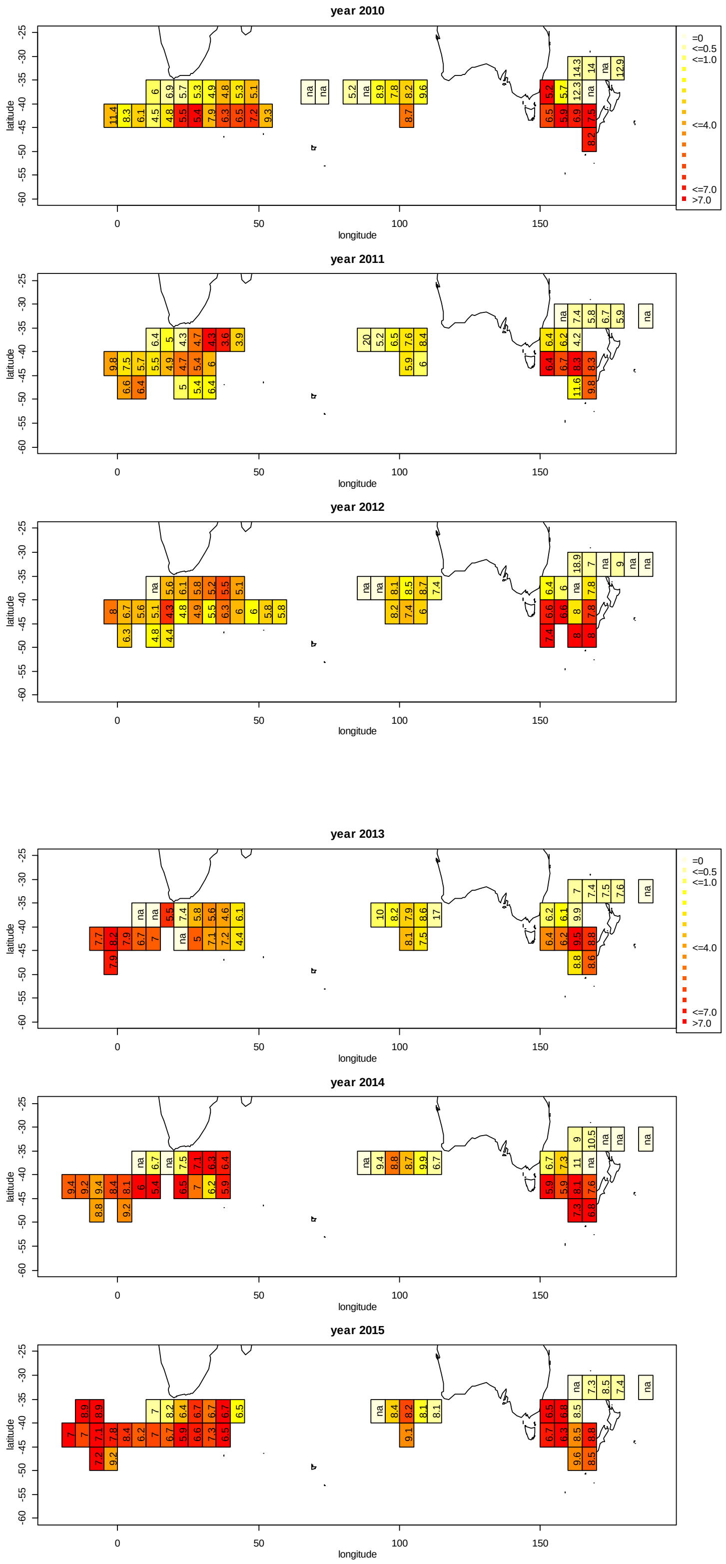


Fig. 9 (cont'd). Distributions of nominal CPUE for age 4+ and average age within 5x5 square by year. For each square, CPUE is represented by color spectrum and a number indicates the average age. Note that 'na' means that catch within a square was 0 and no age frequency data available.

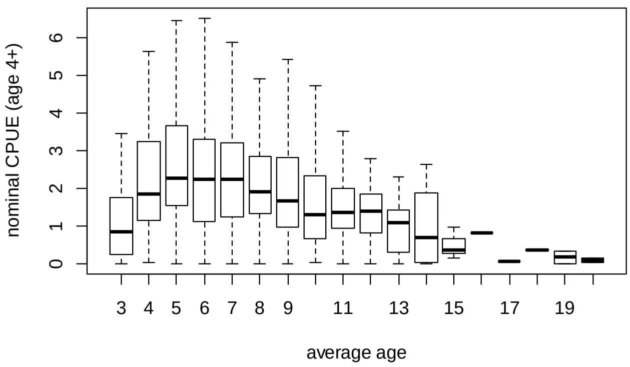


Fig. 10. Relationship between nominal CPUE for age 4+ and average age within 5x5 square. Only positive catch data were plotted due to availability of age frequency data. To easily identify the relationship, average age was treated as categorical variable for this figure (treated as continuous variable in GLM analysis).

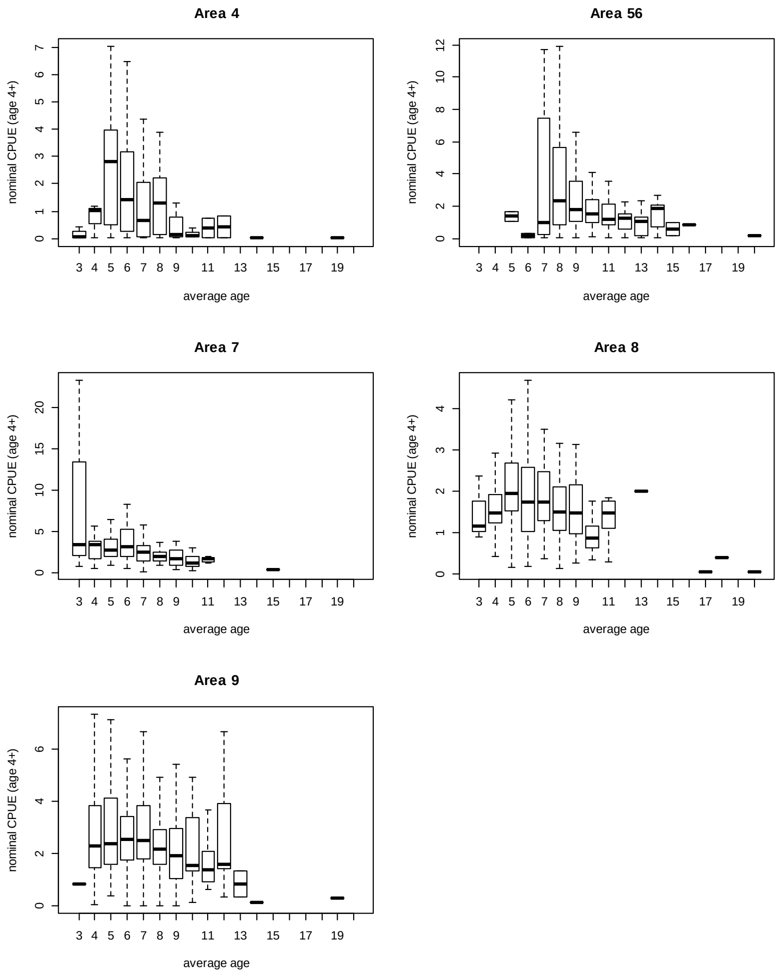


Fig. 11. Nominal CPUE by average age and area. Only positive catch data were plotted due to availability of age frequency data. To easily identify the relationship, average age was treated as categorical variable for this figure (treated as continuous variable in GLM analysis). Areas 5 and 6 were treated as one area '56' combined.