Indian Ocean Tuna Commission Working Party on Billfish Working Paper: **IOTC-2010-WPB-04**

Swordfish Catch Rate Standardization for the Seychelles Semi-Industrial and Industrial Longline Fleets

June 2010

Dale KOLODY (Dale.Kolody@iotc.org)

Jan ROBINSON (jrobinson@sfa.sc)

Vincent LUCAS (vlucas@sfa.sc)

Abstract

This document describes some exploratory work undertaken to standardize the swordfish CPUE series from the Seychelles semi-industrial (SEMI) and industrial (IND) longline fleets for the purposes of stock assessment. The primary goal of CPUE standardization is to estimate a time series of (fishery-selected) relative abundance, and this is accomplished, or at least attempted, by identifying and removing the effects of various sources of CPUE variation that are attributable to causes other than changing abundance (e.g. changes in efficiency of the fleet due to improvements in technology or changes in targeting practices). The analysis includes: i) simple descriptive summary of spatial/temporal effort patterns and operational details, ii) attempts to identify subsets of the data that represent relatively homogeneous fleets with consistent operations, and iii) estimation and presentation of annual time series of relative abundance derived from different GLM models. The main focus was the SEMI fleet, but a brief analysis of the IND fleet was undertaken to check the consistency between the two fleets in the area of overlap.

The analysis of the SEMI fleet suggested that the CPUE series may be sensitive to individual vessel effects (which are problematic because the fleet has a high turnover with few vessels that operate for long periods of time). The analysis was robust to different assumptions about the possible influence of shark targeting, but different assumptions about tuna targeting led to somewhat different results. Swordfish CPUE may be influenced by relative changes in tuna and swordfish targeting. While we are not confident that the analyses undertaken can adequately estimate the targeting effect, operators in the SEMI fleet are known to be conservative and targeting shifts from swordfish to tuna may be relatively uncommon. Key features of the time series are robust: i) there appears to be a decline in CPUE from around 1999/2000 to 2002/2003, and ii) there is a substantial steady decline from 2005 to 2009. In contrast, the initial analysis of the IND fishery suggests that the CPUE is stable over the 2004 to 2008 period. Further analyses are encouraged.

Introduction

This paper represents a preliminary attempt to describe and analyze the historical swordfish CPUE data from the Seychelles semi-industrial (SEMI) and industrial (IND) longline fleets in preparation for the Indian Ocean Tuna Commission (IOTC) Working Party on Billfish (WPB) in 2010. The locally-owned SEMI fleet was introduced in the mid-1990s and consists of small vessels that operate around the periphery of the Mahé Plateau. Logbooks have been returned to the Seychelles Fishing Authority (SFA) since 1995. Many of the vessels switch between swordfish/tuna targeting and shark targeting, and this switching was especially evident in relation to restrictions placed by the Seychelles authorities on exports of swordfish to the EU due to concerns over cadmium concentrations in excess of EC regulations. Disruptions in the fishery began in December 2002 and continued through 2003, during which period vessels needed verbal permission to fish and export, until a formal ban on export (i.e. refusal to issue health certificates) was imposed by Seychelles in early 2004. After the EU revised its (prohibitive) regulation and raised allowable cadmium to realistic levels, the ban on export was removed by Seychelles in March 2005. The industrial fleet targets various tuna species and has a much broader range, with some logbooks reporting operations in the Atlantic and Pacific Oceans. The main focus of the analysis was the SEMI fleet. The IND fleet has a much shorter history and a much broader geographic range, but may provide an important independent source of information in the future. Recommended time series are proposed for the WPB 2010 stock assessment.

Stock assessment for most large pelagic fish species requires the interpretation of commercial CPUE as an index of relative abundance. For the Indian Ocean swordfish stock(s), the Japanese and Taiwanese longline fleets have traditionally been used to generate these abundance indices. These fleets have an extensive history, broad spatial coverage, and substantive logbook programmes. However, the operations of these fleets have changed historically, with large shifts in targeting that are poorly quantified. Large discrepancies between the estimated time trends of the two fleets (e.g. Nishida 2008) indicates that at least one series must be substantially biased. This is believed to be related to targeting changes, and there is an ongoing problem that not all of the relevant operational details are available for analysis (e.g. Hooks Per Basket is not available as a proxy for depth prior to the mid-1990s for the Taiwanese fleets). Conventional fisheries theory (i.e. stationary recruitment dynamics) suggests that the depletion estimated by the Japanese series has been more consistent with the swordfish exploitation history than the Taiwanese series, and this interpretation has generally been given more weight in the assessment and management advice provided by the WPB. However, the WPB also recognized that the Japanese fleet underwent some dramatic changes in the 1990s that might be exaggerating the estimated level of swordfish depletion at that time, particularly in the SW Indian Ocean region. It was recognized that other swordfish fisheries in the Indian Ocean (La Reunion, Seychelles and Spain), may provide additional sources of information about swordfish abundance trends.

The primary goal of CPUE standardization is to estimate a time series of (fishery-selected) relative abundance, and this is accomplished, or at least attempted, by identifying and removing the effects of various sources of CPUE variability that are attributable to causes other than changing abundance (e.g. changes in efficiency of the fleet due to improvements in technology or changes in targeting practices). Two tactics were employed in this paper: 1) attempt to identify homogeneous fleets that are likely to have consistent targeting practices, such that only data with core fishing characteristics were included (e.g. night-sets and lightsticks only, exclude boats with short history in the fishery, etc.), while other data were discarded, and 2) Generalized Linear Models (GLMs) were used to estimate the effects of independent variables which are expected to influence catchability, such that the effect of these variables can be removed to estimate a time series in which (ideally) the main source of variability is changing swordfish abundance. The first approach can be problematic if too many observations are discarded and the analysis becomes restricted to a very small subset of the fishery (e.g. which may result in a very short time series, or a fishery that does not describe important spatial changes in the population, such as a range contraction). The second approach can be problematic if everything changes at once, or there is a very unbalanced distribution of operations (e.g. changes in abundance and targeting will likely be confounded in the model if all boats face similar economic circumstances and change their operations simultaneously). Of course, neither approach will be very successful if the important variables are not available to the analysis, or the mechanism of operation is misunderstood.

Data

The SEMI logbook data used in these analyses are maintained by the SFA, and cover the period from 1995-2009. Only swordfish/tuna longline trip-types, designated 'semiLL' or 'semiLL+SHK' were provided. Each record corresponds to an individual set, with location, set-time, hooks, catch in mass by species, and vessel ID (total of 6468 observations). The logbook data are not complete - around 57% of logbooks are missing for 2002-2003, 43% in 1997, while returns exceed 75% in all other years. It is possible that additional data may be available in the SFA databases that were not used for this analysis. In particular, it is not clear whether the shark trips (which were not included in the analysis) can provide any useful information about either swordfish abundance or targeting characteristics. A number of observations were excluded from all analyses, including:

- missing start position (8 Obs)
- combined SEMILL+SHK trip type (41 Obs)
- longitude outliers (2 Obs)
- 1995 had too few data
- Trips in which lightsticks were either not used or not recorded were excluded

Note that in some of the descriptive analyses that follow, the data were being explored in an iterative process to identify the most consistent and useful subset for analysis. As such, the different plots might correspond to slightly different subsets of the data, but this should not affect the main points of discussion.

Semi-industrial fleet characteristics

SEMI Spatial and Temporal Distribution of Effort

The time series of catch, effort and nominal swordfish CPUE from the SEMI fleet (Figure 1) shows the initial development of the fishery, the decline in catch and effort from 2001 to a low in 2003-2004, and the subsequent increase in SWO catch and, to a lesser extent, effort. The inter-annual variability in the nominal time series is probably greater than one would expect purely on the basis of the reproductive dynamics of the species. Given this level of variability, it is not obvious that there is any overall trend in abundance. However, there is a severe and consistent downward trend over the past 5-6 years which should not be ignored.

The spatial distribution of SEMI (swordfish/tuna) sets over all years is shown in Figure 2. Figure 3 shows how the spatial pattern varies among years. Few sets are reported from 1995, so this year was dropped from subsequent analysis. In some years (1996, 2003-4), the SW region was lightly fished, and the SE region seems to be avoided in June in all years. Overall, the patterns do not suggest the sequential offshore expansion observed in some other coastal swordfish fisheries. Figure 5 indicates that there is no obvious seasonality in the spatial pattern of sets. Figure 6 suggests that there is seasonality in the swordfish catch rates, with a peak visible in the April-July period in most years - the latter months of this period are characterised by comparatively large catches of spawning females from northern areas of the fishing ground. Figure 7 illustrates the distribution of effort by month and year, and suggests that there is more seasonality in latter years.

Given the small area of the fishery (relative to distant water fleets), it is not clear how important the spatial heterogeneity in the fishery is. An exploratory GAM (in which Year and Season effects were essentially assumed constant) was used to examine potential spatial heterogeneity as a continuous function of space. This suggests that there may be spatial patterns in catch rates that might be reasonably well described by disaggregating the region into 4 sub-areas (Figure 4). However, it is possible that the spatial pattern could be an artefact of other effects as discussed subsequently.

SEMI Vessel Consistency

Out of a total of 23 vessels with swordfish/tuna Trip-types logged, a maximum of 10 have operated in any single year from 1995-2009 (Figure 9), and more than half of the vessels have fished for 3 years or less. This raises an important concern about the consistency of the vessels in the fleet. If the vessels have substantially different practices, it seems likely that a lot of the noise (and potentially large time series biases) in the nominal CPUE could be the result of individual vessel effects. It also seems likely that there could be a high degree of confounding among effects in the GLM models (e.g. if each vessel tends to operate in a particular quarter in a particular area, for only one or two years, it may be impossible to distinguish whether exceptional catch rates are due to the effects of the year, month, area or vessel). However, Figure 11 suggests that most vessels do not have strong preferences for particular areas, and Figure 12 suggests that most vessels do not have seasonal preferences.

It is also likely that there are learning processes going on in the fleet. Those that operate for one year only (one third of the fleet) might reasonably be expected to be less effective than the experienced boats. Figure 18 shows the distribution of vessel experience over time. It is not clear what the most appropriate measure of effort should be. The relationship between experience in sets and experience in years can be poor (e.g. Figure 17). However, both are flawed because logbook returns are incomplete, and skippers probably change vessels.

Several different approaches were employed to examine the sensitivity of the analysis to vessel effects:

- Individual vessel ID was included as a factor.
- A categorical variable for vessel experience in years was included.
- Experience in sets was included as a categorical variable with 5 levels (0-25, 26-75,7 5-150, 150-300, 300+) to recognize that learning usually tapers off with time.
- Models were refit with a subset of observations corresponding to vessels with limited experience excluded (<5 years, and <7 years).

If the analyses all produce similar results, this increases the confidence that the vessel effects may not be biasing the results very badly.

SEMI Operational data

The large majority of the SEMI swordfish/tuna sets are night sets (Figure 15) with lightsticks (Figure 13). It is curious that the few non-lightstick sets are associated with higher catch rates than the lightstick sets, especially as they mostly occur in later years, when average CPUE is down overall (Figure 14). Only sets with lightstick field = 1 (defined as lightsticks used) were included in the analyses (numbers of lightsticks used were not available).

The fishery appears to have a slight preference for setting on the full moon (Figure 16). Moon phase was included as a categorical variable with 2 levels (full = lunar day 10-20, new=everything else). Various studies have shown that moon phase has a statistically significant effect on swordfish catch rates and the timing of the effect is reported to differ among fisheries (e.g. Poisson et al, in press, and references therein). Thus, there is probably justification for including additional levels of moon phase in the analysis. However, moon phase generally has an inconsequential effect on annual time series of relative abundance, presumably because there is not much inter-annual variability in how fisheries operate with respect to moon phase.

There were no other operational data available with which targeting shifts could be further explored (e.g. with respect to hook types, bait, gear depth, etc.). However, there are data available on the catches of other species. There are various ways to use species composition as an indicator of targeting and all have problems. For example, including tuna catch as a predictive variable is problematic because it reflects not only targeting, but also tuna abundance which will change among years irrespective of targeting; the use of the ratio of swordfish/tuna represents a problem because it introduces the dependent variable into the predictive side of the equation. Similar problems can be identified for other approaches. However, ignoring targeting changes is also a problem. In this analysis, the following species-based approaches were used to consider targeting effects:

- Catches of all species other than swordfish were included as a continuous variable
- Catches of all sharks were included as a continuous variable
- All sets with positive shark catch were removed
- All sets with positive bigeye tuna catch were removed
- All sets with positive yellowfin tuna catch were removed

As with the vessel effects, if the analyses all produce similar results, this provides some indication that targeting effects may not cause large temporal biases in the time series. However, if there are important differences in the estimated time series, this should be interpreted as an avenue for further investigation, rather than a reliable quantitative result.

Methods

Generalized Linear Models (GLM) were used in this analysis, and the model notation adopted here is that used by R software (e.g. Crawley 2007). In all models explored here, the dependent variable was (a function of) CPUE by set. The independent variables used in the analysis are described in the data section above and summarized in Table 1.

Modelling the Zero CPUE observations

The distribution of swordfish CPUE observations appears to be approximately lognormally distributed (Figure 10), except for the 15% of zero observations (i.e. zero catch and positive effort). Three approaches were used to consider the influence of the zeroes (Table 2):

1. LN – traditional linear models are used with dependent variable $\log_e$(CPUE+C), where the constant C is equal to the lower 10^{th} percentile of all non-zero CPUE observations. The annual time series of the lognormal models are estimated using the standard approach (e.g. Venables and Dichmont 2004, where P is the parameter estimate, and sigma is the SE of P. Integration of parameters is required when Year-interactions are estimated.

$$I_y = \exp\left(\widehat{P}_y^Y + \frac{1}{2}[\sigma_{\mathbf{P}}^2)]\right) - \mathbf{C}$$

2. BC - traditional linear model is used in which the Box-Cox transformation is applied to non-zero CPUE observations, while the zero CPUE observations are removed.

3. DL - the two stage delta-lognormal model was used. First the probability of observing non-zero CPUE is estimated as a Bernoulli process (binomial GLM). Then, the traditional log(CPUE) approach is used to model the non-zero observations. The two models are integrated together to estimate an annual standardized time series.

These three models were compared with the simplest combination of explanatory variables (Year + Quarter + Area) (appendices 1-3). For all 3 models, all of the factors are highly significant and the residual behaviour is reasonable. Model semiBC1 demonstrates the best behaviour with respect to the normality assumptions; however, given that this model ignores the zeroes completely, this is not really a sufficient criteria for adoption.

App. 3 (semiDLN1), illustrates the estimated relationship between the probability of obtaining a non-zero observation, and the expected value of the non-zero observation. This suggests that the special treatment of the zeroes may have some justification. However, there was no appreciable difference among the final time series estimated by the three models (Figure 20), and on this basis, approach number one above was adopted for subsequent analyses.

In principle, the standard errors of the estimated parameters could be used to describe the uncertainty in the time series; however, these estimates tend to be unbelievably precise for a number of reasons (e.g. observations are not truly independent). The uncertainty is usually inflated to 'more realistic' levels when subsequently interpreted in the context of stock assessment (e.g. CV is inflated in the assessment model or estimated with a lower bound). The uncertainty estimates are not considered further at this time.

Model Selection

Model development consisted of fitting a broad range of explanatory variables in a range of combinations. Table 2 and Table 3 provide a representative indication of the model specifications examined in this work. Additional details of graphical diagnostics, statistical output, and parameter estimates are included in the appendices for a subset of models.

The following points were considered in selecting the models:

- Does the residual and diagnostic behaviour suggest that the statistical model is reasonable?
    - Normally-distributed residuals?
    - Limited leverage of outliers?
- Is the effect of a factor both statistically significant and of practical significance?
- Are the estimated parameters consistent with the proposed mechanisms?
- Do additional explanatory variables justify the additional explained variance?
- Is there obvious confounding of the factors with different implications for the relative abundance series?

- If multiple models are essentially indistinguishable with respect to these features, and they have very different implications for the final time series, they should be considered as alternative plausible series.

In addition to this iterative, manual approach to model building, automated selection procedures were tested, in which AIC or BIC were used as the selection criterion. The R function step() was used to reduce a highly parameterized model (semiLN8) down to something more parsimonious using a backward stepwise approach, and lowest AIC as the optimality criterion. The various versions of AIC and BIC are useful because they recognize that there is a trade-off between model complexity and explanatory power, i.e. adding more parameters will inevitably allow a model to fit the observations better, but adding spurious variables degrades the predictive performance of the model. AIC and BIC criteria reward a model for improved fit to the observations, but penalize it for the number of explanatory variables introduced (the relative trade-off between these two competing factors differ in different versions of AIC and BIC). During each iteration of the stepwise procedure, the model is sequentially refit with one of the variables removed relative to the full model. The explanatory term with the highest AIC (poorest fit) is removed, and the process is repeated with the reduced model. When removing terms no longer decreases the AIC, the process is complete. This process is described once in appendix 4. However, it was tested a few times, with the consistent result that AIC tended to recommend models that were highly parameterized (including dubious interaction terms), while the more restrictive BIC tended to recommend models that were in line with the decisions that were reached manually.

SEMI Results and Discussion

Figure 8 compares two common measures of annual nominal CPUE, mean(C(by set)/E(by set)) sometimes referred to as 'area-weighted'(particularly when observations are aggregated on a spatial grid) - and (sum(catch(i))/sum(Effort(i)) which is often called 'effort-weighted'. The two series are very similar in this case, and are used for comparison with the standardized series discussed below.

The important results from the models described in Table 2 are briefly summarized in the following points:

- The annual estimated time series from most models were very similar to each other and the nominal CPUE series. The main exception was probably the full model (semiLN8 – not shown). All of the models examined that included Year-Area or Year-Quarter interaction terms resulted in erratic time series, and in most cases not all interaction parameters could be estimated (due to the unbalanced distribution of observations). Interactions were not considered further at this time.

- The main space-time effects (Year, Quarter, Area) were always highly significant (except the Area main effect in some models that included Area interactions) and were included in all models.

- Moon phase was consistently statistically significant, and it was retained because it has a well-established effect in other fisheries. However, it did not have much of an effect on the estimated time series.

Vessel Characteristics - The final time series from the 5 models that focussed on vessel effects are shown in Figure 21 (e.g. semiLN5, semiLN6, semiLN7, semiLN21, semiLN22 from Table 3).

- The vessel ID and vessel experience terms were all statistically significant (when added individually to the model with Year + Area + Quarter).

- Vessel ID (semiLN5) explained more CPUE variability than experience. The estimated effects of the different vessels is large (e.g. the best boat is estimated to have swordfish catch rates roughly 4 times higher than the least effective boats, Figure 23).

- When included as a predictive variable in the model, the different measures of experience (semiLN6, semiLN7) were estimated to be statistically significant, but these terms did not explain much of the variance, and do not seem to have much effect on the final time series. In both cases, the proposed mechanism (increased experience should be associated with higher CPUE) is weakly supported by the trend in parameter estimates (Figure 24). However, this could simply be a spurious result (i.e. there will always be an increase in experience over time – if there happens to be an increase in abundance over time as well, then the model will likely estimate a positive relationship whether or not it is real). However, when model semiLN6 was refit using only the data from the last 6 years (when mean CPUE is steeply declining), there was still a positive (though weaker) estimated effect of experience on CPUE (Figure 24, middle panel).

- When vessels with less experience are removed from the time series (semiLN21, semiLN22), there is a noticeable effect on the CPUE time series (primarily a large increase in the estimated abundance in 1998), but the general features of the time series remain unchanged. Figure 25 illustrates how the vessel consistency criteria affects the number of sets each year. It is apparent that the core vessels in the fleet tended to fish throughout the lean period (2002-3) when the transient vessels departed. This finding is supported by SFA observations that a large number of vessels switched from targeting swordfish/tuna to targeting sharks during the restrictions imposed by the cadmium issue from late 2002 to early 2005.

- It seems likely that the individual vessel effects are important in this small, high turn-over fishery; however, it is not clear that these models provide the best way to quantify these effects. However, it does seem reasonable that concentrating the analysis on core vessels with a consistent history is likely to avoid some problems, and on this basis, model semiLN21 is recognized to be marginally preferable than some others (appendix 5).

Species Targeting Shifts - The final time series from the 5 models (semiLN3, semiLN4, semiLN23, semiLN24, semiLN25,) that attempted to identify and quantify the effects of species targeting shifts are shown in Figure 22.

- Inclusion of the allSharks term in the model (semiLN3) was not significant. Similarly, the exclusion of sets with positive shark catch (semiLN25) does not result in a large change to the estimated CPUE time series. This suggests that variability in shark catches does not have much effect on swordfish CPUE, and this is consistent with the perception that the SEMI fleet has distinct shark targeting practices (i.e. day sets and wire leaders), and the two types of sets are probably not mixed together in this database.

- Inclusion of nonSWO (all non-swordfish catch) as an independent variable (semiLN4) was found to be statistically significant. However, the relationship was positive, which is the opposite of the proposed mechanism. The a priori expectation was that a negative relationship might indicate switching of target species between swordfish and something (everything) else. A positive relationship simply indicates that abundance of species covaries (and it would most likely be misleading to include this term in the standardization).

- When one compares the time series based on subsets of the data that exclude positive BET catch (semiLN23) and YFT catch (semiLN24), there is a substantial effect. However, it is not clear what, if anything, this means. This may be indicative of some subtle switching between tuna and swordfish targeting over time (which is not detected in other operational factors), which appears to have been sporadic and largely exploratory based on interviews with skippers Or, it may simply be an artefact of the way in which the data were subset in the analysis. However, in both cases, the most concerning downward trend over the last few years is still strongly evident.

Model Selection for the WPB

- Fitting and evaluating a range of models has not produced an obvious candidate time series to recommend for stock assessment purposes. The models examined suggest that the analysis is somewhat sensitive to assumptions about the individual vessel operations, and possibly the differential targeting of tuna and swordfish over time.

- However, all of the estimated time series have some features in common:

    - CPUE declined prior to the fishery disruptions caused by the cadmium issue, the decline coinciding with a period of the highest effort in the fishery.

    - CPUE increased during the period of fishery disruptions caused by the cadmium issue, when effort in the fishery had declined substantially.

    - CPUE seems to drop reasonably consistently over the last 5-6 years, although effort in the fishery had recovered only partially to pre-2003 levels.

  Given the limited spatial extent of the fishery and assuming seasonal or annual migration of swordfish into the fishing grounds from the wider stock, a potential mechanism could involve localised depletion and possible gear interactions at high effort (1998-2002), a reduction in effort and gear interactions leading to higher CPUE for vessels remaining in the fishery during the cadmium-related disruptions (2003-2005), followed by a return to localised depletions and gear interactions as effort increased after the export ban was lifted. While the consistent decline in CPUE since 2005 was not commensurate with the partial recovery of effort during the period, only the most experienced vessels resumed swordfish/tuna targeting and effective effort may be compensating for the lower number of hooks compared to the 1998-2002 period. Candidate mechanisms for this trend in CPUE will require further consideration and analyses.

- The automated stepwise model selection procedure (on the full dataset) recommended model semiLN9 on the basis of the BIC criteria. This is roughly consistent with the results that were achieved through manual iterative development, and reassuring in that the interaction terms were rejected. However, we do not recommend model semiLN9 as preferable, because the appropriateness of the non-swordfish targeting index is doubtful.

- If a single time series is required for quantitative analyses in the WPB, we would tentatively recommend semiLN21. This series does not adequately account for possible effects of vessel experience or targeting shifts. However, by estimating vessel effects, and eliminating the most transient vessels, this series should account for at least some of the variability in CPUE that is not attributed to abundance.

Seychelles Industrial Fleet

The Seychelles industrial fleet (IND) is described briefly, as a possible source of additional information on swordfish abundance trends. The IND fleet consists of large, long distance vessels. The data covers 35472 sets from 2001-8. 5052 sets are in the Seychelles region as defined below, and 4308 of these sets are in the period from 2004-8.

Figure 26 shows the broad spatial distribution of sets in the industrial longline logbook data, covering large parts of the Indian Ocean, and extending into the Atlantic and Pacific Oceans. These data are partitioned by year in Figure 27. Figure 29 illustrates the catch, effort and nominal CPUE for the whole Indian Ocean region, and the Seychelles region only (as defined in Figure 26). These figures suggest that there may be enough observations to be useful in the Seychelles region for the 2004-8 period.

Figure 28 shows the CPUE spatial contours for the IND fleet from the equivalent spatial GAM that was applied to the SEMI fleet. Compared to the SEMI fleet (Figure 4), the IND fleet shows a much more uniform spatial distribution of effort and more uniform distribution of CPUE. The sub-area boundaries were retained for the IND analysis for consistency, however, this is not really a meaningful comparison because there is very little direct overlap in the spatial distribution of the two fleets within the defined region.

Figure 30 illustrates that there are a large number of zero swordfish CPUE sets, more than the semi-industrial fleet, and enough that zero-inflated or delta models should be explored.

A number of models similar to those applied to the semi-industrial fleet were applied to the industrial fleet as described in Table 4. The residual behaviour of the lognormal(CPUE+C) model indLN1 indicates that the normality assumptions are probably questionable in this case (appendix6). Application of the delta lognormal model indDL1 greatly reduces the problem with the distribution of residuals (appendix 7), however, again there really is not much difference between these two time series and the nominal series over the time period 2004-8 (Figure 31). Unlike the SEMI fleet, there is no evidence for a downward trend in SWO abundance over the period 2004-8.

Conclusions

- A number of GLM approaches have been employed to estimate a relative abundance index for swordfish in the vicinity of the Seychelles.

- The SEMI time series appeared to be robust to the treatment of the zero CPUE observations, and this is perhaps not surprising given that they only account for 15% of sets.

- The standardization models applied to the Seychelles SEMI fleet suggested that the individual vessel effects are important, and potentially problematic because very few boats have a long history in the fishery. The analyses were also sensitive to different species targeting assumptions. This is indicative of a potential problem, but the methods used are not expected to be appropriate to resolve the problem.

- Despite the sensitivities, the SEMI standardized CPUE series had some robust features (that were also evident in the nominal CPUE). In particular, the CPUE series declines from 1998 to 2002, and declines sharply over the last 5-6 years. Prior to 2005, the catch and effort series show time trends that appear to be related to the substantial economic implications of the cadmium-related export restrictions beginning in late 2002. However, the specific mechanism for the effect on CPUE will require further investigation (perhaps fine scale depletion and renewal processes are involved).

- A preliminary attempt was made to standardize the Seychelles IND fleet swordfish CPUE series from the region immediately surrounding the area where the SEMI fleet operates. While there is little direct overlap in the operation of the two fleets, it is assumed that the two fleets are catching the same swordfish population. The GLM analyses suggested that the IND time series from 2004-8 was also robust to the treatment of the zero observations, and the explanatory variables examined did not reveal any obvious biases in the nominal CPUE series. Unlike the declining trend in the SEMI fleet, the IND time series is very flat.

- We would tentatively recommend time series from model semiLN21 from the SEMI fleet, and indDL2 for the IND fleet for further consideration by the WPB 2010 (Table 5, Figure 32). These time series are not perfect, but it is not obvious that they are more biased than other series potentially under consideration in the stock assessment. Further investigation of both time series is warranted, and comparisons should be made with the Japanese and Taiwanese fleets operating in the region.

References

WPB 2009. Indian Ocean Tuna Commission Report on Billfish 2009. Report of the 7th Session of the IOTC Working Party on Billfish. Seychelles, July 6-10, 2009.

Bishop, J. 2006. Standardizing fishery-dependent catch and effort data in complex fisheries with technology change. Rev. Fish. Biol and Fisheries 16:21-38.

Crawley, M. 2007. The R Book. John Wiley & Sons, Ltd. 942.

Maunder, M.N. and Punt A.E. 2004. Standardizing catch and effort data, a review of recent approaches. Fish. Res. 70: 141-159.

Nishida, T. 2008. Notes on the standardized swordfish CPUE of tuna longline fisheries (Japan and Taiwan) in WPB6 (1980-2006 and 1992-2006). IOTC-2008-WPB-INF04.

Poisson, F., J.-C. Gaertner, M. Taquet, J.-P. Durbec, and K. Bigelow. In press. Effects of lunar cycle and fishing operations on longline-caught pelagic fish: fishing performance, capture time, and survival of fish. Fish. Bull. XXX

Venables, W.N., and Dichmont, C.M. 2004. GLMs, GAMs, and GLMMs: an overview of theory for applications in fisheries research. Fish. Res. 70: 319-337.

WPB 2001. Indian Ocean Tuna Commission Report on Billfish 2009. Report of the 2nd Session of the IOTC Working Party on Billfish. St Gilles, La Réunion, November 5-8, 2001.

WPB 2009. Indian Ocean Tuna Commission Report on Billfish 2009. Report of the 7th Session of the IOTC Working Party on Billfish. Seychelles, July 6-10, 2009.

Table 1. Explanatory variables used in the GLM models as discussed in Table 2 to Table 4.

Explanatory Variable	Description
Year (factor)	the most important factor in the standardization, the estimated year effects are required (possibly in addition to other effects and interactions) to describe the relative abundance of the population (i.e. Year effects are required whether or not they are statistically significant)
Month (factor)	Describes seasonal effects of abundance; often cannot be estimated because of seasonal distribution of effort
Quarter (factor)	Aggregate of months Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec
Area (factor)	Four sub-regions shown in Fig. 2; Given the relatively small region of the Indian Ocean in which the fleet operates, it is not obvious that spatial effects are as important as they would be for the DWF fleets that operate over much larger parts of the Indian Ocean.
Moon Phase (factor)	Either fullmoon or newmoon in this analysis. Often subdivided into new, waxing, full waning, or a continuous function of light intensity.
Vessel ID (factor)	Unique vessel identifier; may be a proxy for operational factors that are not otherwise recorded and probably largely reflect the skill of the skipper.
vessExpYears (factor)	Vessel experience in years; a priori expectation that more experience should be associated with greater skill.
vessExpSets (factor)	Vessel experience in sets; recognizes that vessel experience in years and sets can be very different in the SEMI fishery.
allSharks (continuous)	Aggregate catch of all reported shark catch (mass). This was included to see if obvious mixing of shark sets with tuna/swordfish sets could be identified in the data. It was intended as an exploratory tool , not a definitive quantitative approach.
nonSWO (continuous)	Aggregate catch of all species other than swordfish. This was intended as an exploratory tool, under the expectation that an inverse relation between SWO and nonSWO might provide evidence of other species targeting.

Table 2. Seychelles semi-industrial longline fleet CPUE standardization models comparing the treatment of zero swordfish catch observations.

Model	Dependent Variable	Explanatory Variables	Comments
Dataset = semi-industrial (5510 observations)			
semiLN0 (app. 1)	Log(CPUE+C)	Year + Month + Area	Lognormal Model; All predictive variables are significant; $R^2 = 0$
Dataset = semi-industrial, excluding SWO CPUE = 0 (4758 observations)			
semiBC0 (app. 2)	bc(CPUE>0)	Year + Month + Area	bc()=Box-Cox transformation; $R^2 = 0.$
Dataset = semi-industrial (5510 observations)			
semiDL0 (app. 3)	Pr(CPUE>0); Log(CPUE | CPUE>0)	Year + Month + Area	Delta-Lognormal Model

Table 3. Seychelles semi-industrial longline fleet CPUE models discussed in this paper. Note that the appendices describe the version of the LN models with no intercept estimated (the models in this table include the intercept so that the R^2 value is easy to interpret). All variables are categorical except for nonSWO (catches of all non-swordfish), allSharks (total shark catch), and vessExpSets (vessel experience in sets), which are continuous.

Model	Dependent Variable	Explanatory Variables	Comments
Dataset = semi-industrial (5510 observations)			
semiLN1	Log(CPUE+C)	Year + Quarter + Area (20 parameters, R^2 = 0.129)	Lognormal Model; All predictive variables are significant
semiLN2		Year + Quarter + Area + MoonPhase (21 parameters, R^2 = 0.135)	All predictive variables are significant
semiLN3		Year + Quarter + Area + MoonPhase + allSharks (22 parameters, R^2 = 0.135)	allSharks not significant
semiLN4		Year + Quarter + Area + MoonPhase + nonSWO (22 parameters, R^2 = 0.175)	All predictive variables are significant;
semiLN5		Year + Quarter + Area + MoonPhase + VesselID (42 parameters, R^2 = 0.216)	All predictive variables are significant
semiLN6		Year + Quarter + Area + MoonPhase + VesselExperienceYears (31 parameters, R^2 = 0.155)	All predictive variables are significant;
semiLN7		Year + Quarter + Area + MoonPhase + VesselExperienceSets (25 parameters, R^2 = 0.150)	All predictive variables are significant
semiLN8		Year* Quarter*Area – Year: Quarter:Area + MoonPhase + VesselID + VesselExperienceSets + VesselExperienceYears + allSharks + nonSWO (130 parameters, R^2 = 0.301)	Fullest model used as starting point for stepwise reduction (app. 4)
semiLN9		Year + Quarter + Area + moonPhase + VesselID + nonSWO (43 parameters, R^2 = 0.252)	BIC selected model
Dataset = semi-industrial, excluding: **Vessels with a total of less than 5 years in the fishery** (4037 observations)			
semiLN21		Year + Quarter + Area + MoonPhase (21 parameters, R^2 = 0.165)	
Dataset = semi-industrial, excluding: **Vessels with a total of less than 7 years in the fishery** (2708 observations)			
semiLN22 (app. 5)		Year + Quarter + Area + moonPhase (21 parameters, R^2 = 0.186)	
Dataset = semi-industrial, excluding: **Sets with positive BET catch** (2778 observations)			
semiLN23		Year + Quarter + Area + moonPhase (21 parameters, R^2 = 0.12)	
Dataset = semi-industrial, excluding: **Sets with positive YFT catch** (2542 observations)			
semiLN24		Year + Quarter + Area + moonPhase (21 parameters, R^2 = 0.091)	
Dataset = semi-industrial, excluding: **Sets with positive shark catch** (4141 observations)			
semiLN25		Year + Quarter + Area + moonPhase (21 parameters, R^2 = 0.13)	

Table 4. Seychelles industrial longline fleet CPUE models discussed in this paper. Note that the R^2 value is reported for the equivalent model with the intercept estimated (the version in the appendices are based on the table below with no intercept, in which case the R^2 value is not easy to interpret).

Model	Dependent Variable	Explanatory Variables	Comments
Dataset = industrial (4038 observations)			
indLN1 (app. 6)	Log(CPUE+C)	Year +Month + Area – 1	Lognormal Model;
indDL1 (app. 7)	Pr(CPUE>0); Log(CPUE | CPUE>0)	Year + Month + Area – 1	Delta-Lognormal;

Table 5. Time series tentatively recommended for further consideration of the WPB 2010.

Year	SEMI fleet (model semiLN6)	IND fleet (indDL1)
1996	0.72	
1997	0.73	
1998	1.55	
1999	1.37	
2000	1.10	
2001	1.12	
2002	0.92	
2003	0.99	
2004	1.20	1.14
2005	1.27	0.90
2006	0.90	1.06
2007	0.89	0.91
2008	0.54	1.00
2009	0.69	

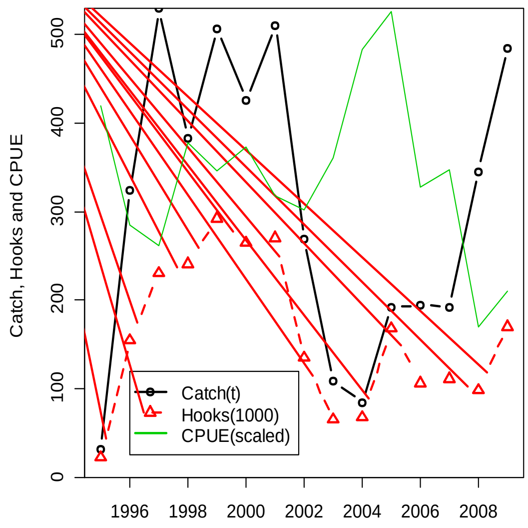


Figure 1. Time series of total catch, total hooks and total Catch/total hooks (rescaled to fit on the same axes) for the SEMI fleet.

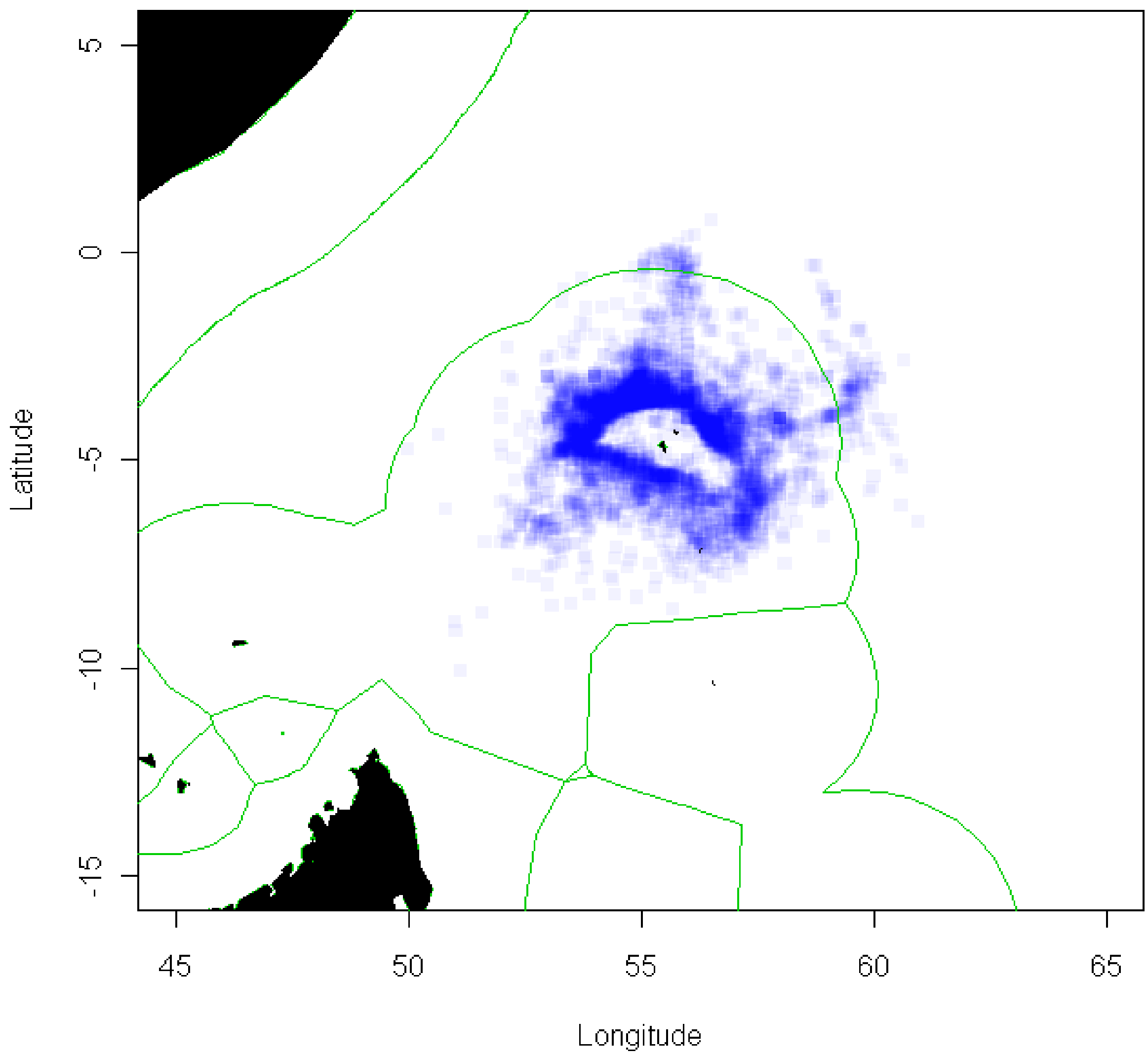


Figure 2. Location of Seychelles semi-industrial longline fleet logbook sets 1995-2009. Full colour saturation indicates 20 or more sets.

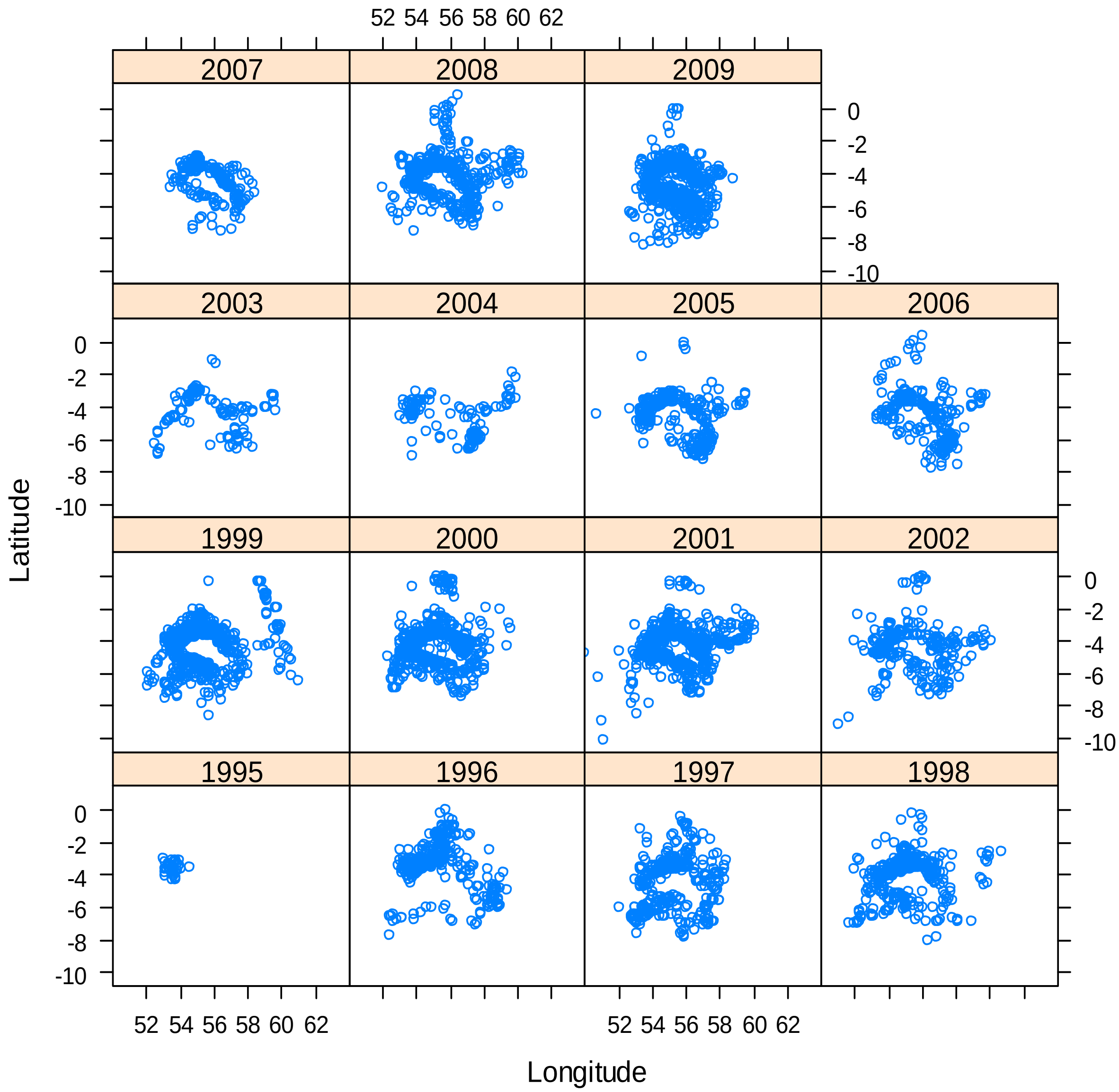


Figure 3. Location of Seychelles semi-industrial longline fleet logbook sets by year.

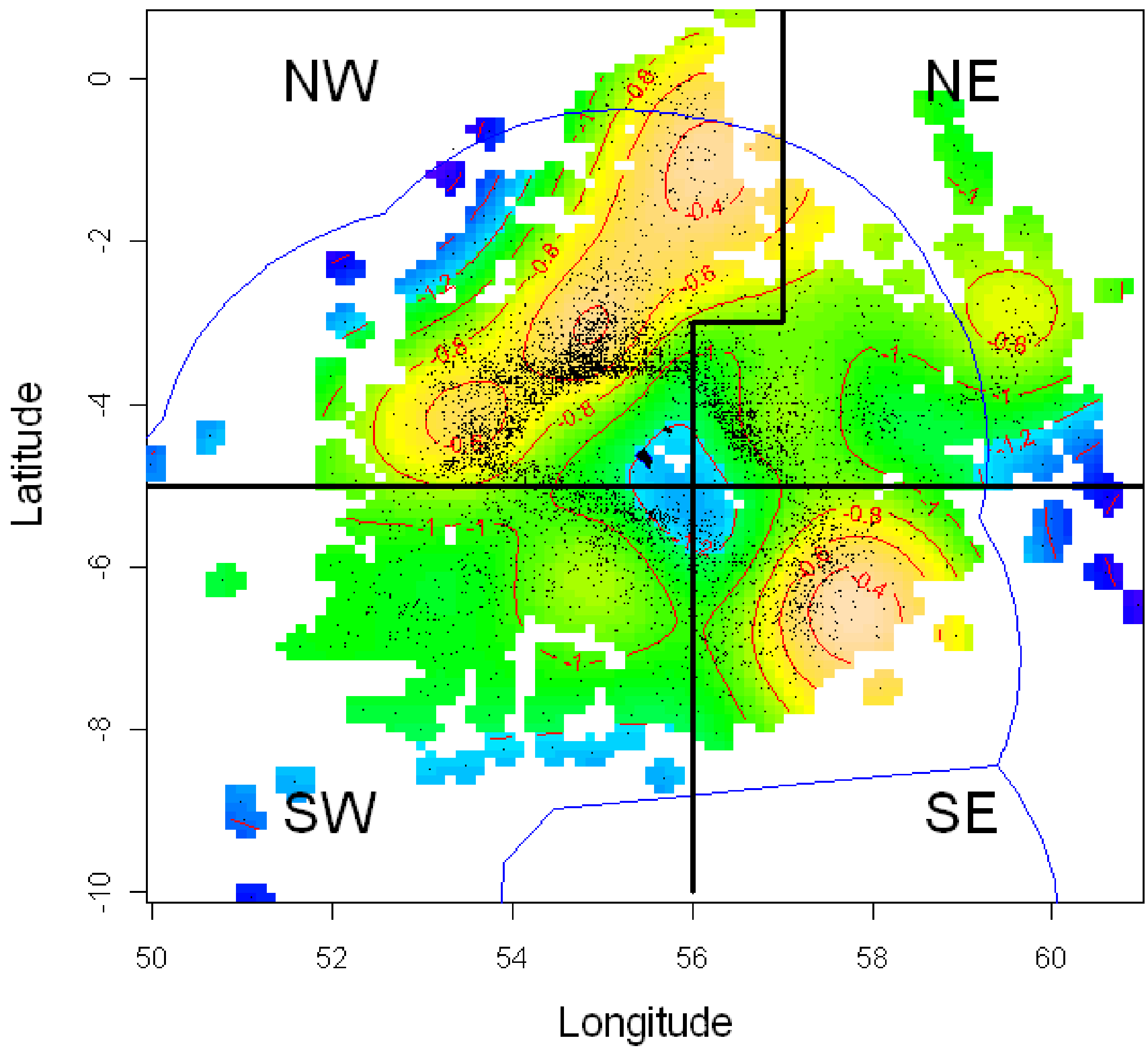


Figure 4. Spatial GAM showing the set locations (points), estimated spatial effects (log(CPUE) contours), and the 4 area structure used in subsequent analyses.

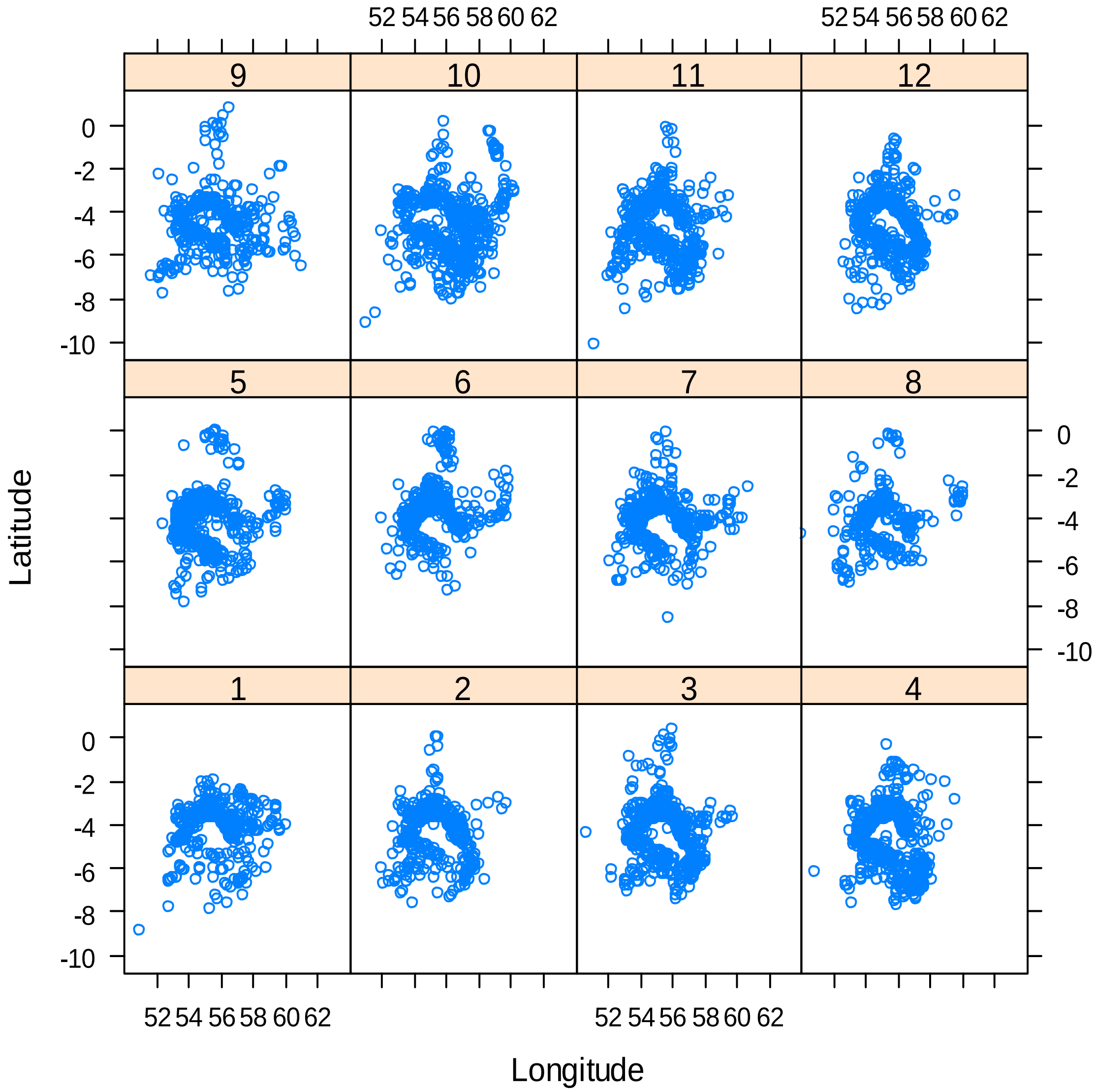


Figure 5. Location of Seychelles semi-industrial longline fleet logbook sets by Month (aggregated across all years).

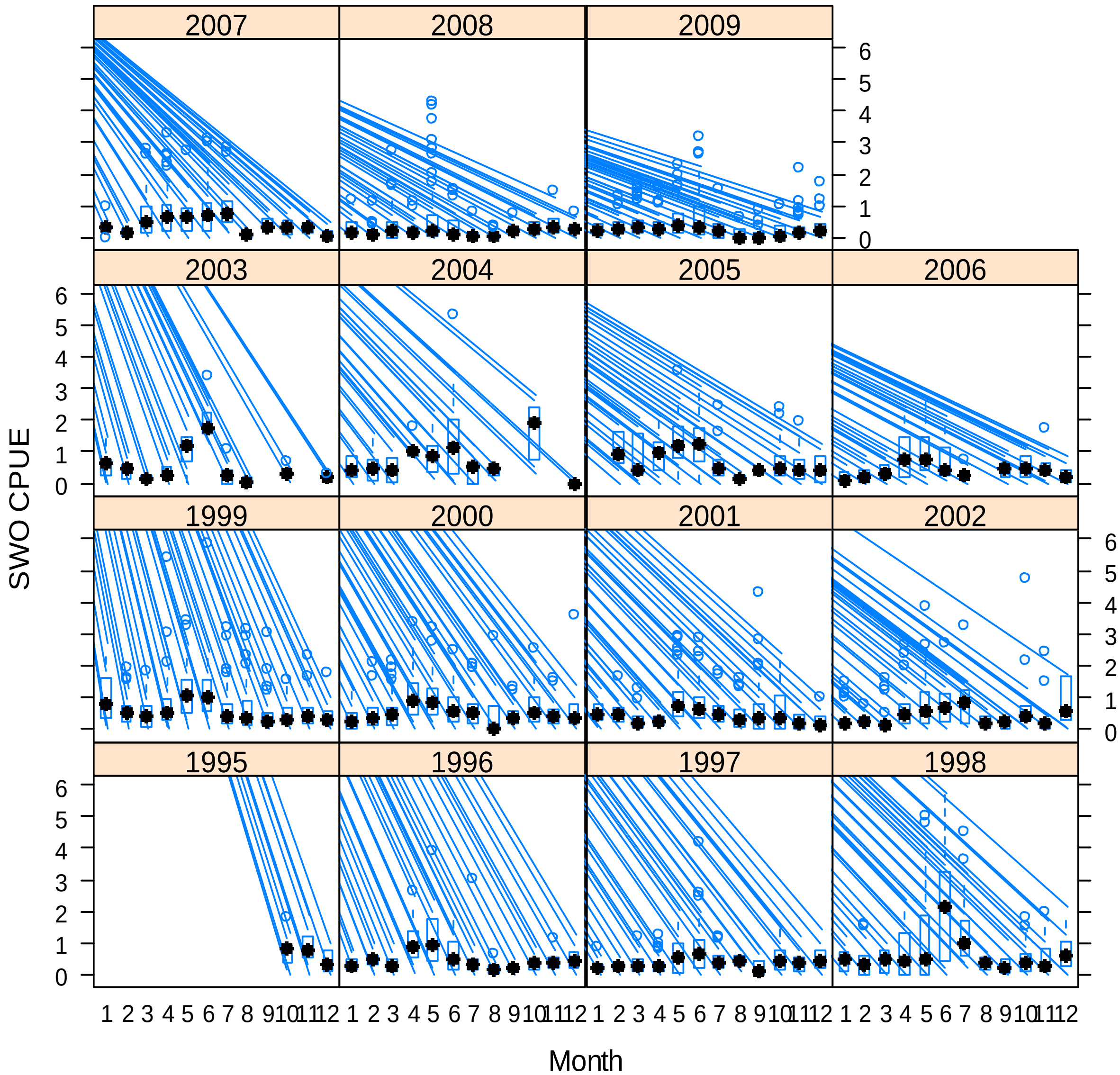


Figure 6. Monthly pattern of Seychelles semi-industrial longline fleet CPUE by year.

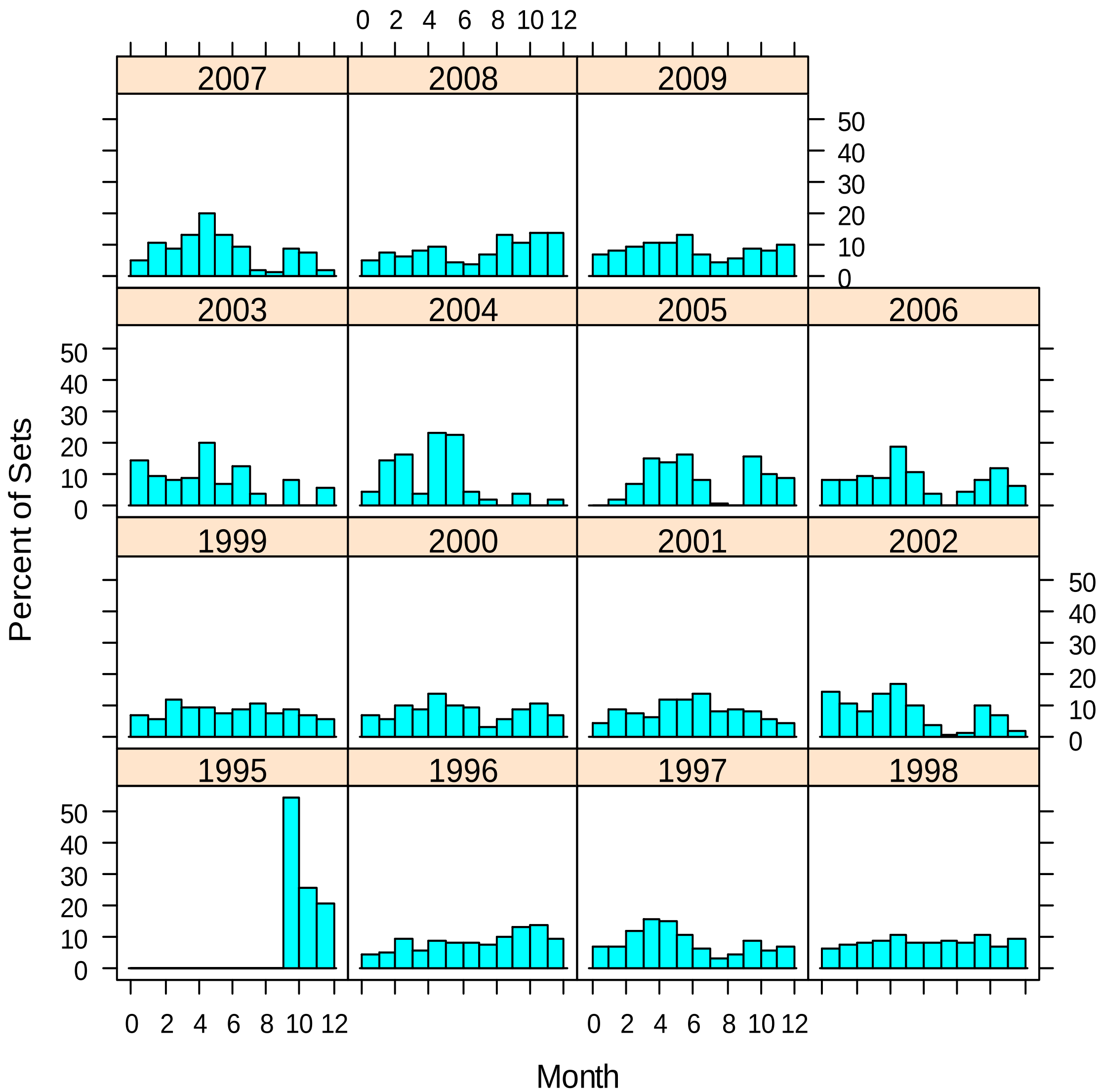


Figure 7. Monthly pattern of Seychelles semi-industrial longline fleet sets by year.

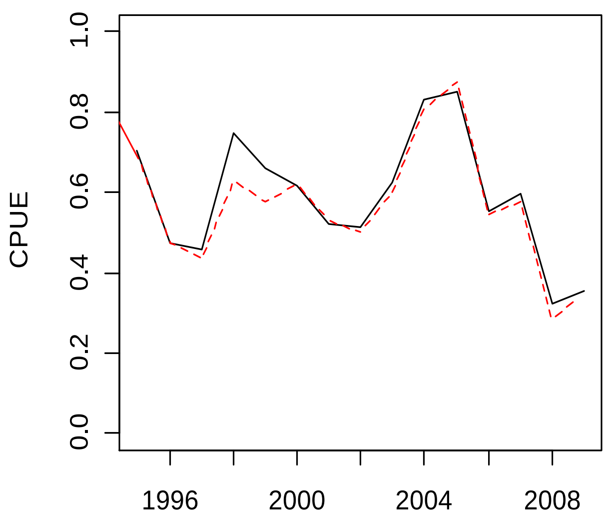


Figure 8. Two different nominal swordfish CPUE series, area-weighted (solid black) and effort-weighted (broken red).

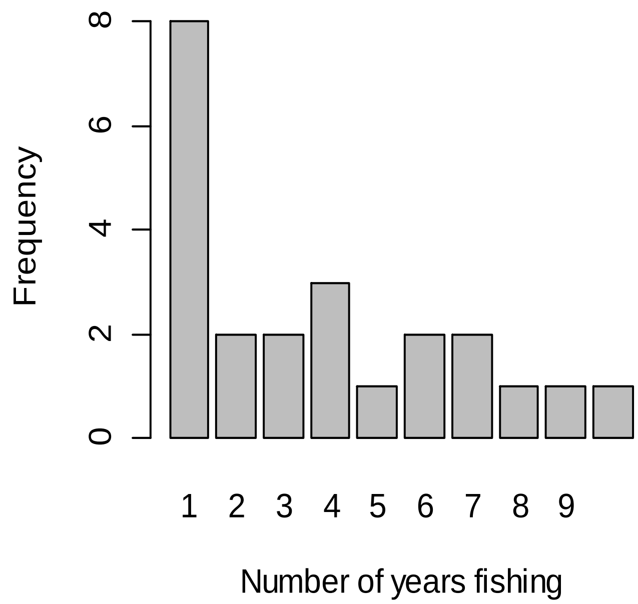


Figure 9. Frequency distribution of the number of years that vessels have fished with SEMI Trip-Types.

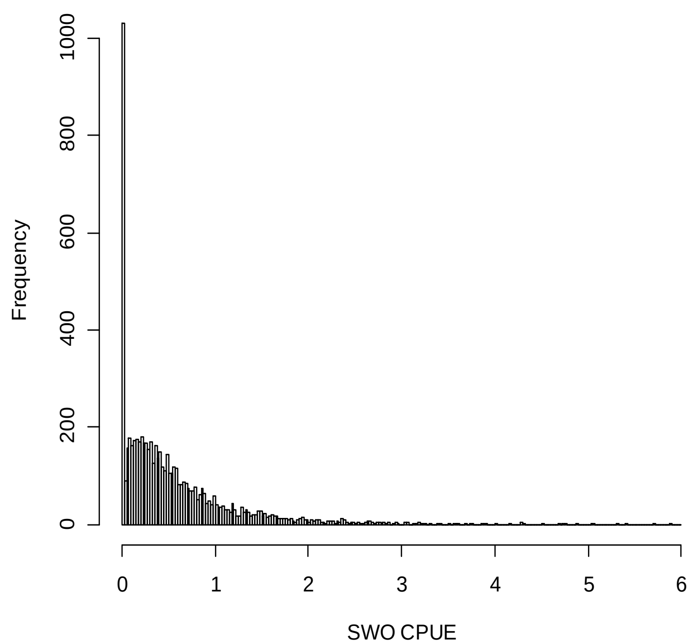


Figure 10. Frequency distribution of semi-industrial fleet CPUE indicating large number of sets with zero catch (15%).

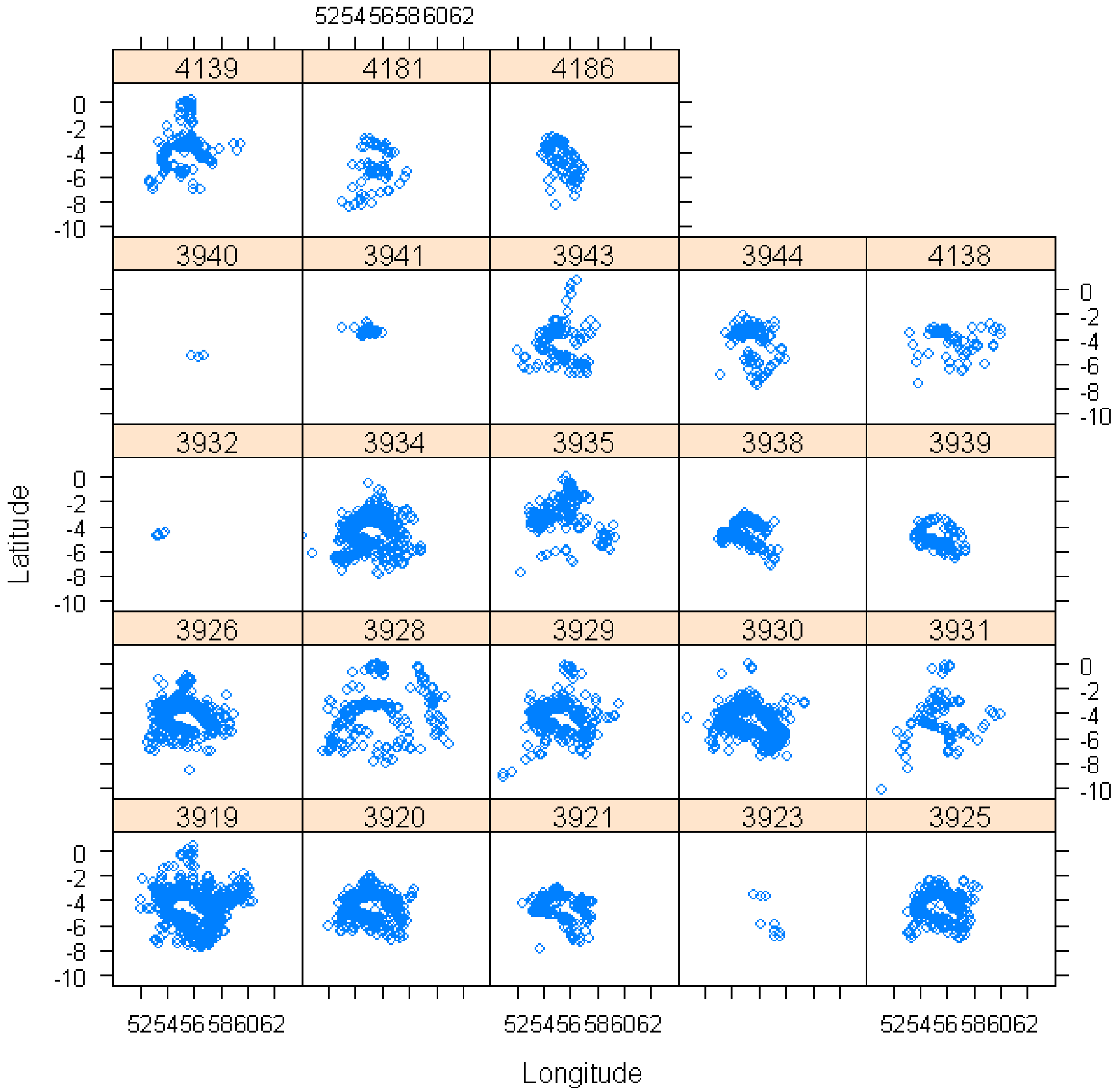


Figure 11. Set position by vessel.

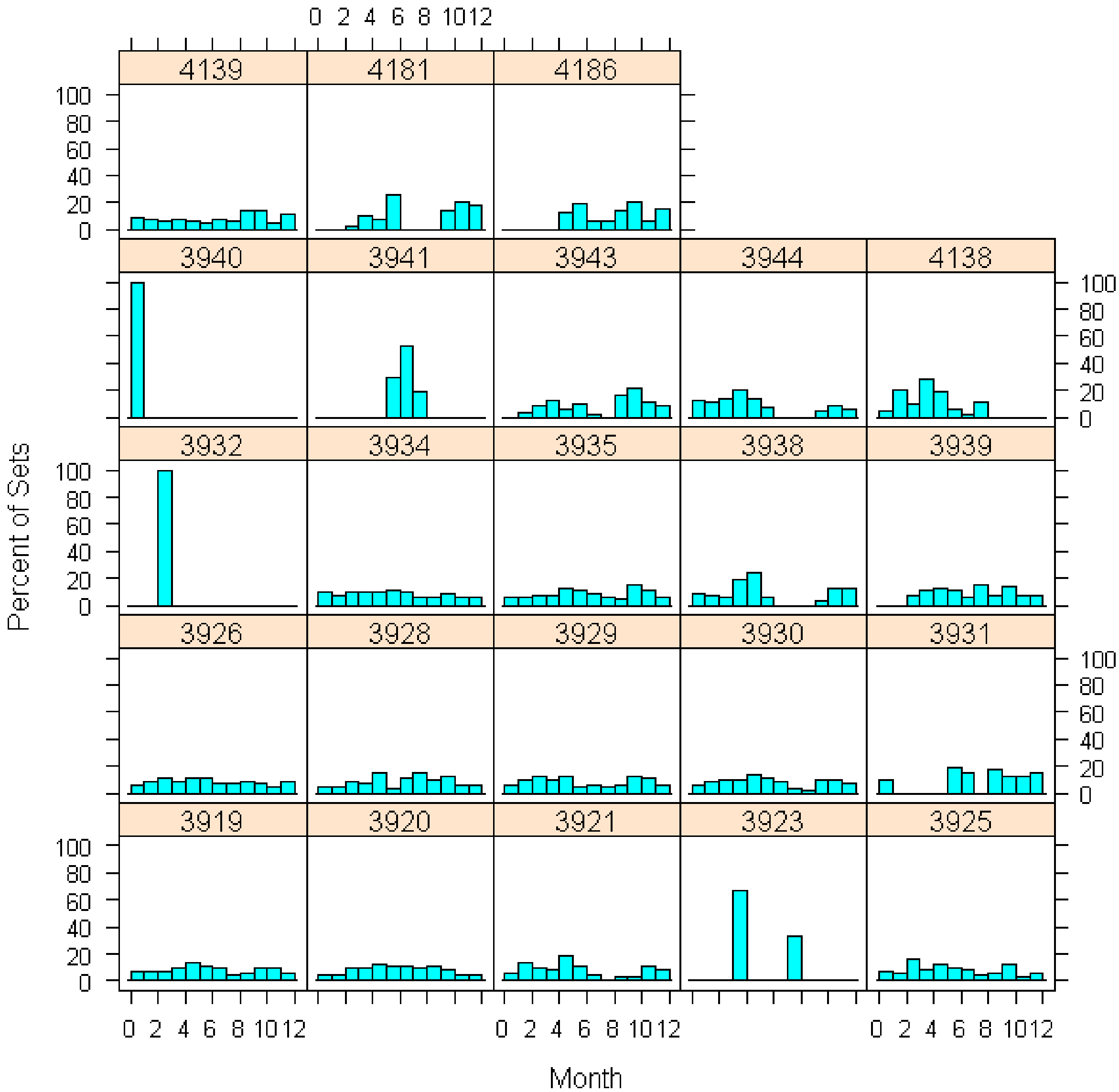


Figure 12. Month of operation partitioned by vessel.

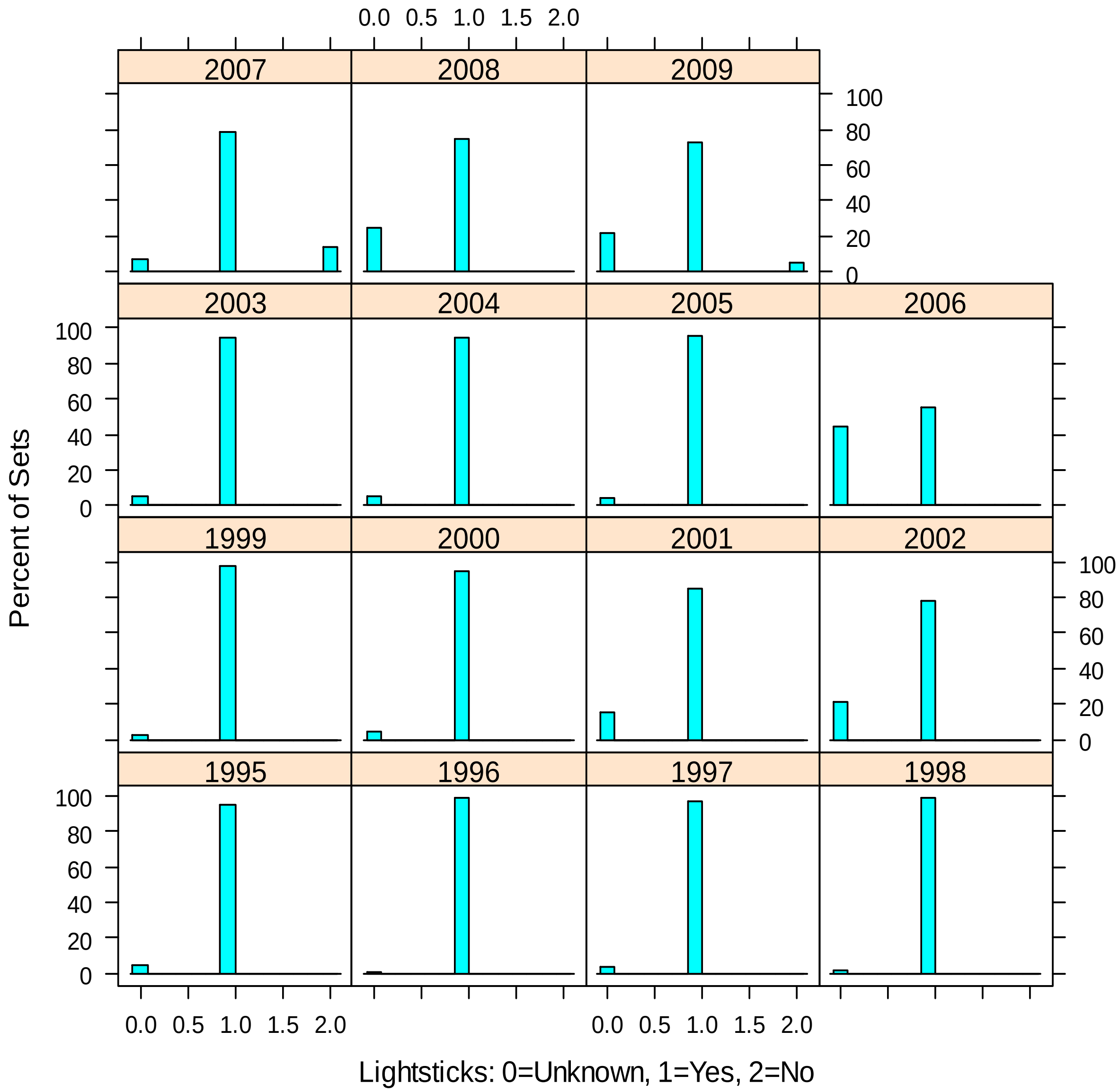


Figure 13. Lightstick usage over time in the semiLL sets.

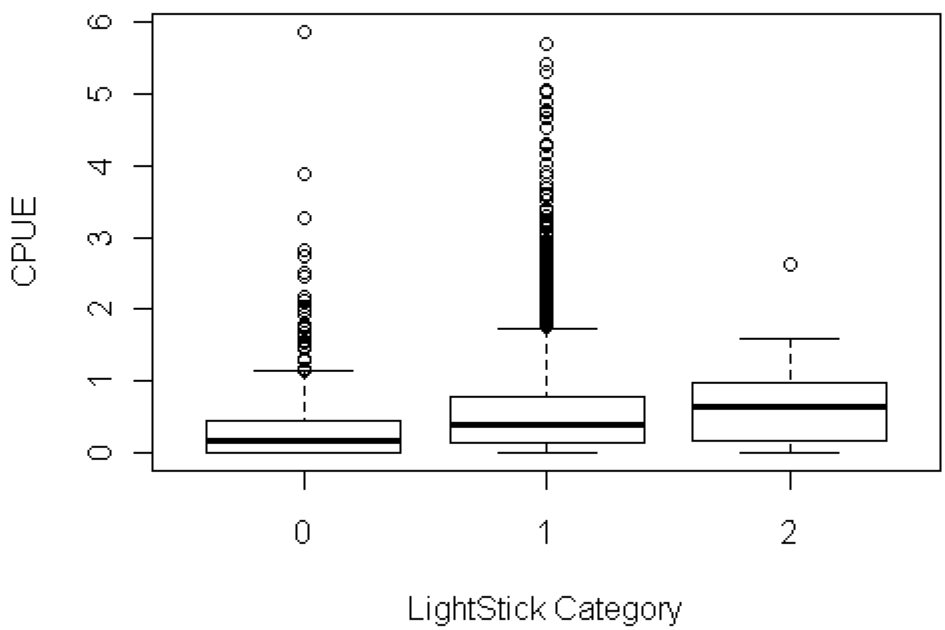


Figure 14. Boxplots of CPUE by category of lightstick usage in the SEMI fleet.

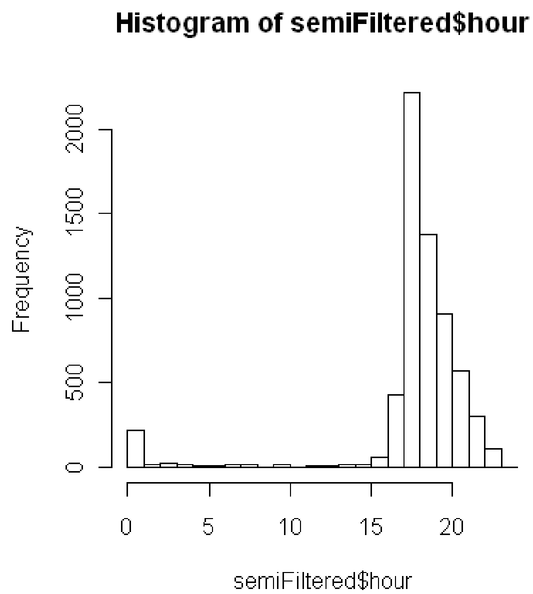


Figure 15. Hourly pattern of sets indicates virtually all sets are night sets.

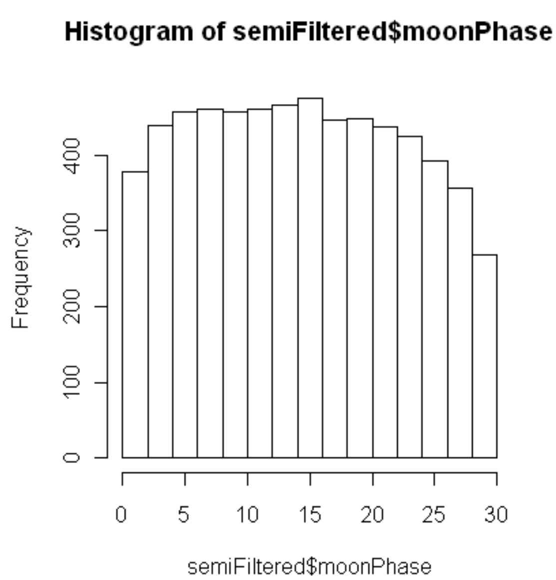


Figure 16. Daily pattern of sets around the moon phase (full moon is at 15)

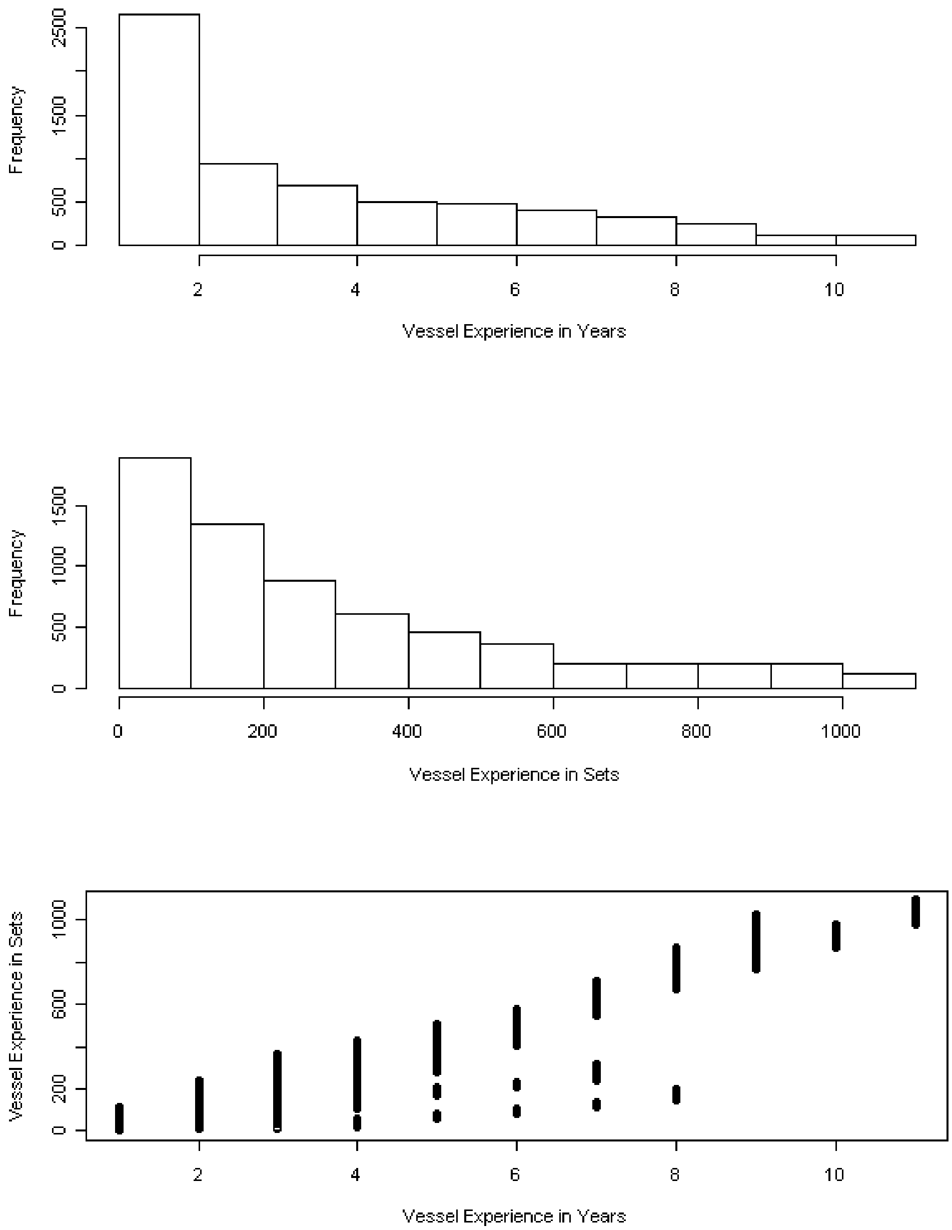


Figure 17. Individual vessel experience with "SEMI" Trip-Types by year and set.

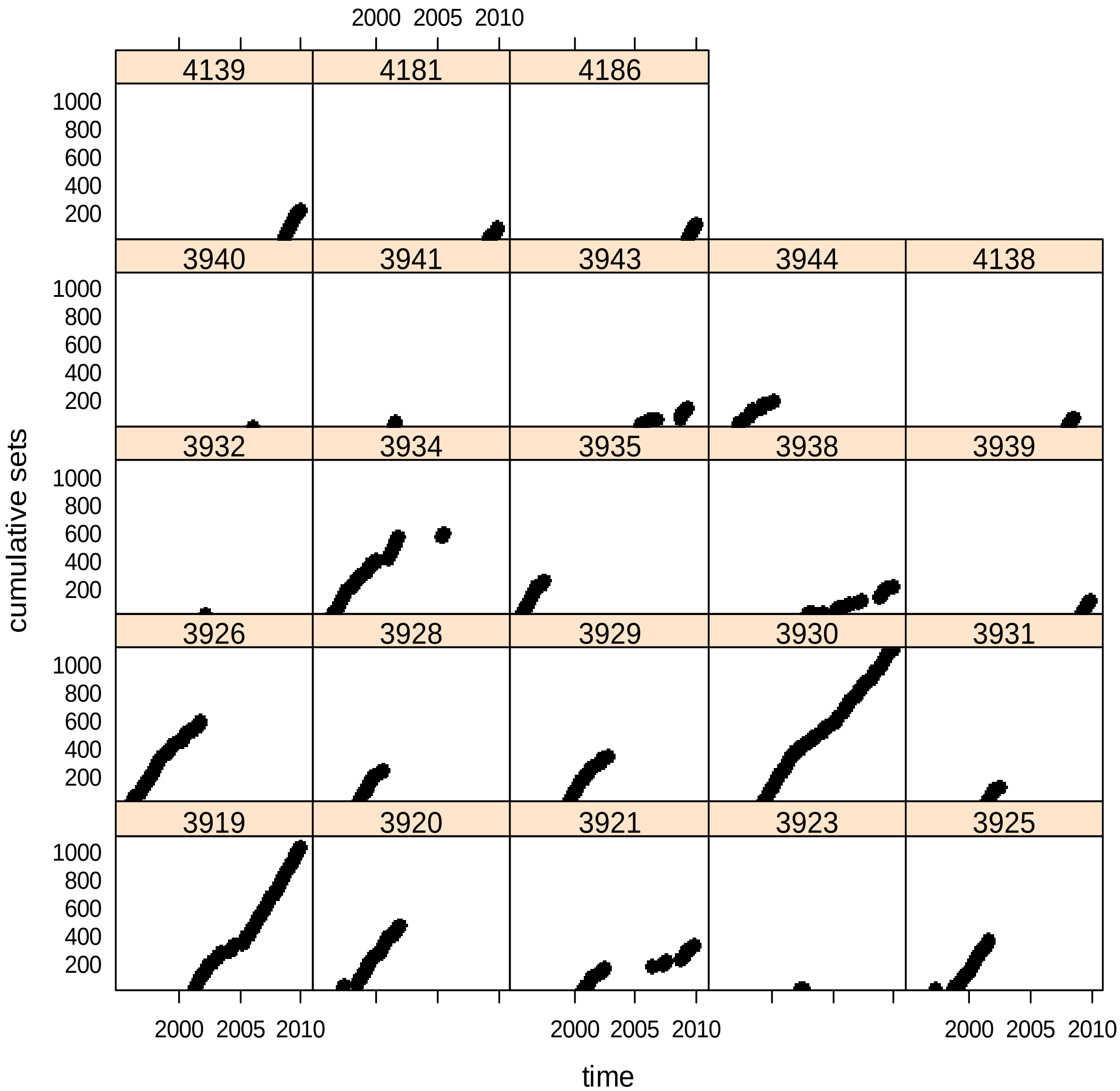


Figure 18. Individual vessel experience measures in sets (tuna/swordfish trips only), from returned logbooks.

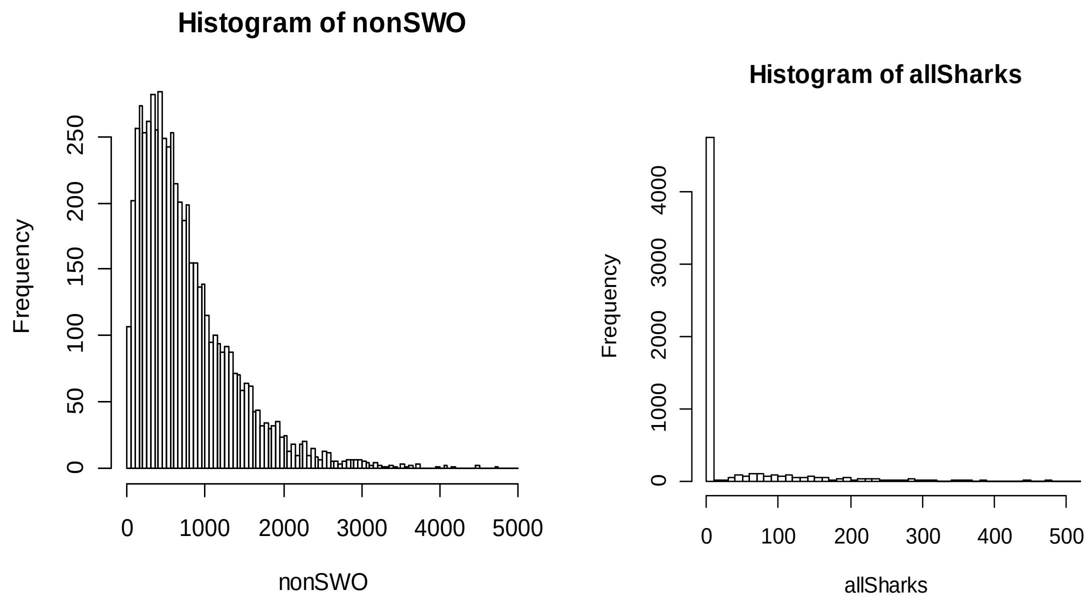


Figure 19. Histograms of other species catches by set: left panel is all shark catch (Mt), right panel is total catch less the swordfish catch.

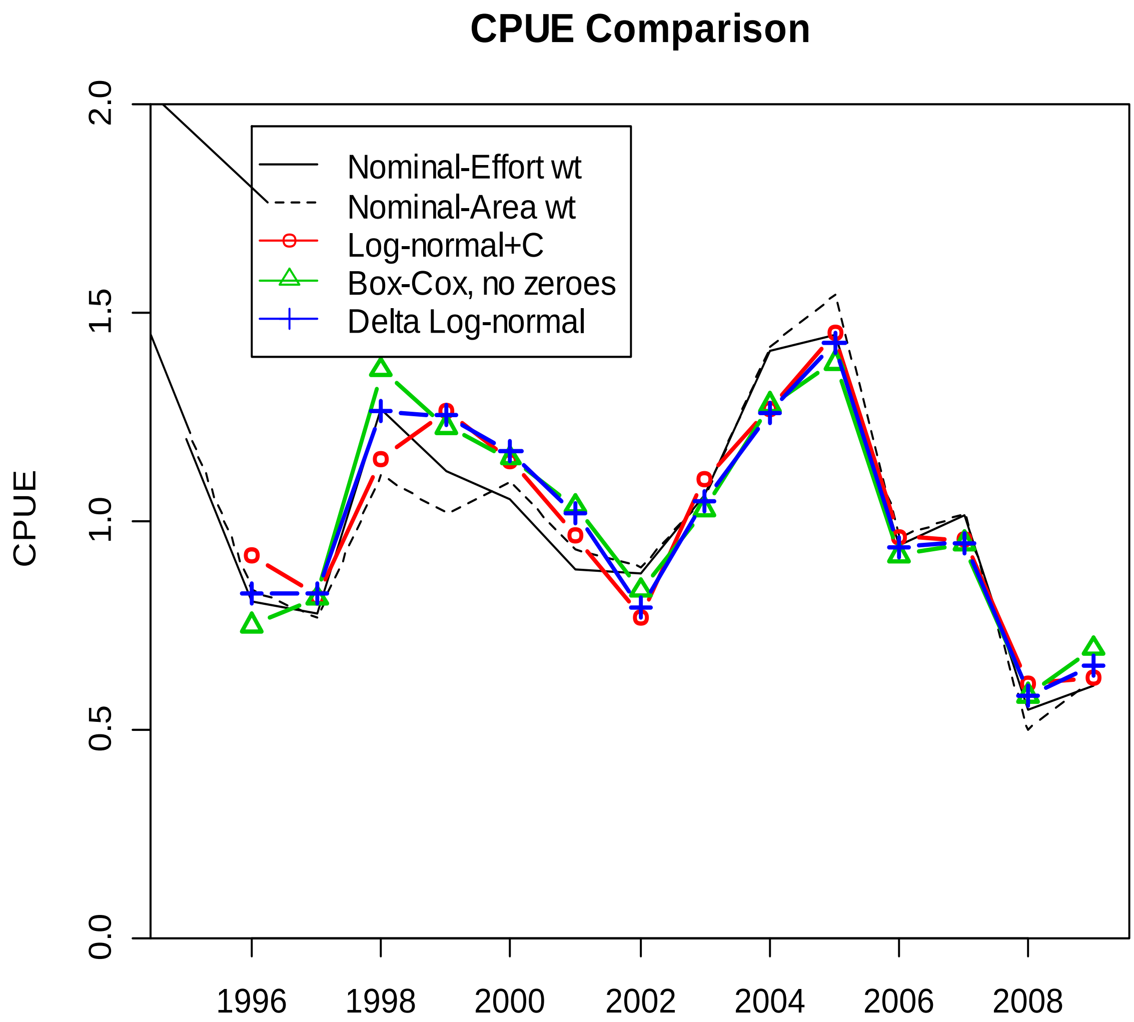


Figure 20. Comparison of SEMI CPUE series, including the nominal series and the different error models (semiLN0, semiBC0, semiDL0) with the main space-time effects.

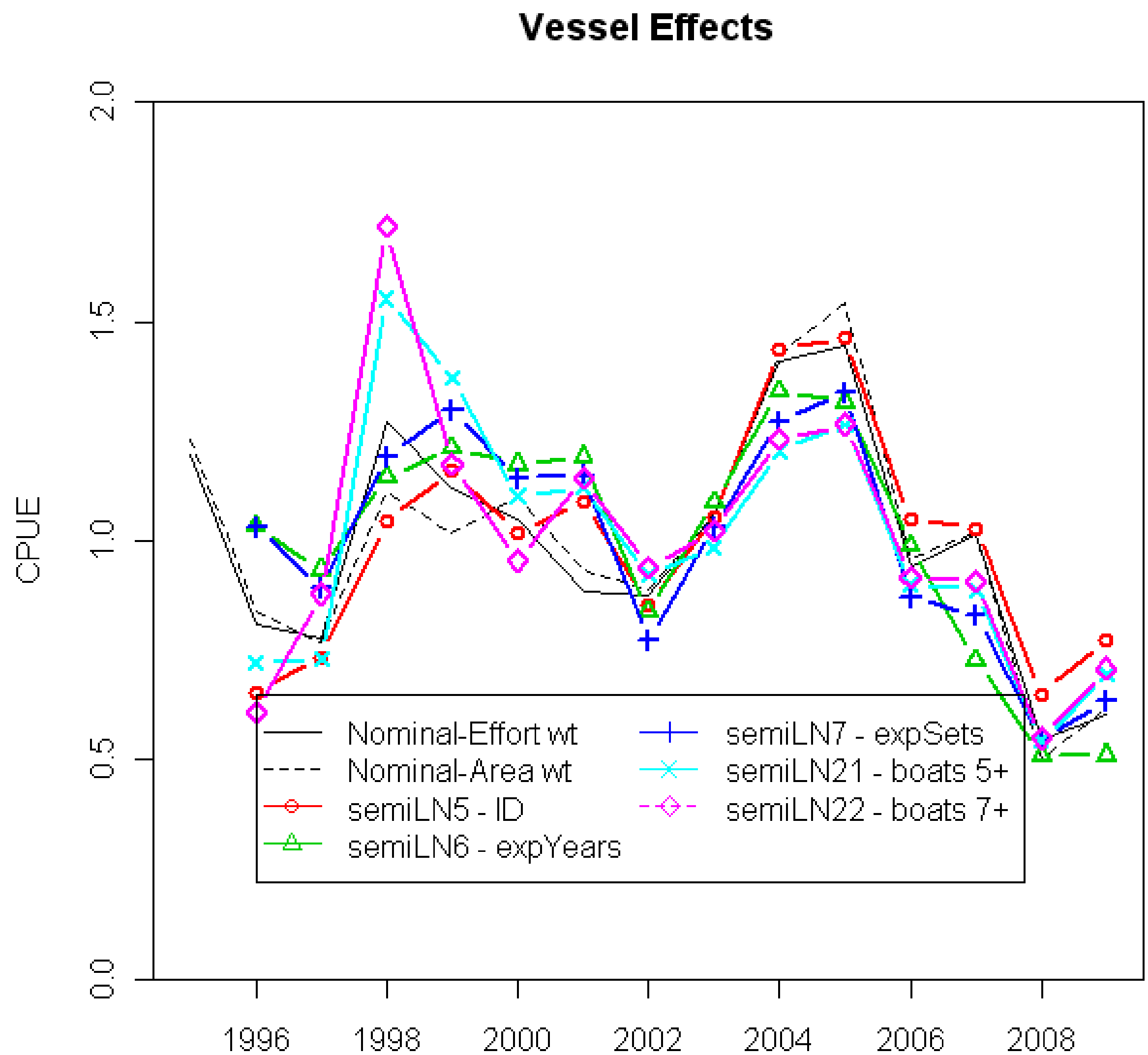


Figure 21. Comparison of Seychelles SEMI swordfish CPUE models that examine different vessel effects.

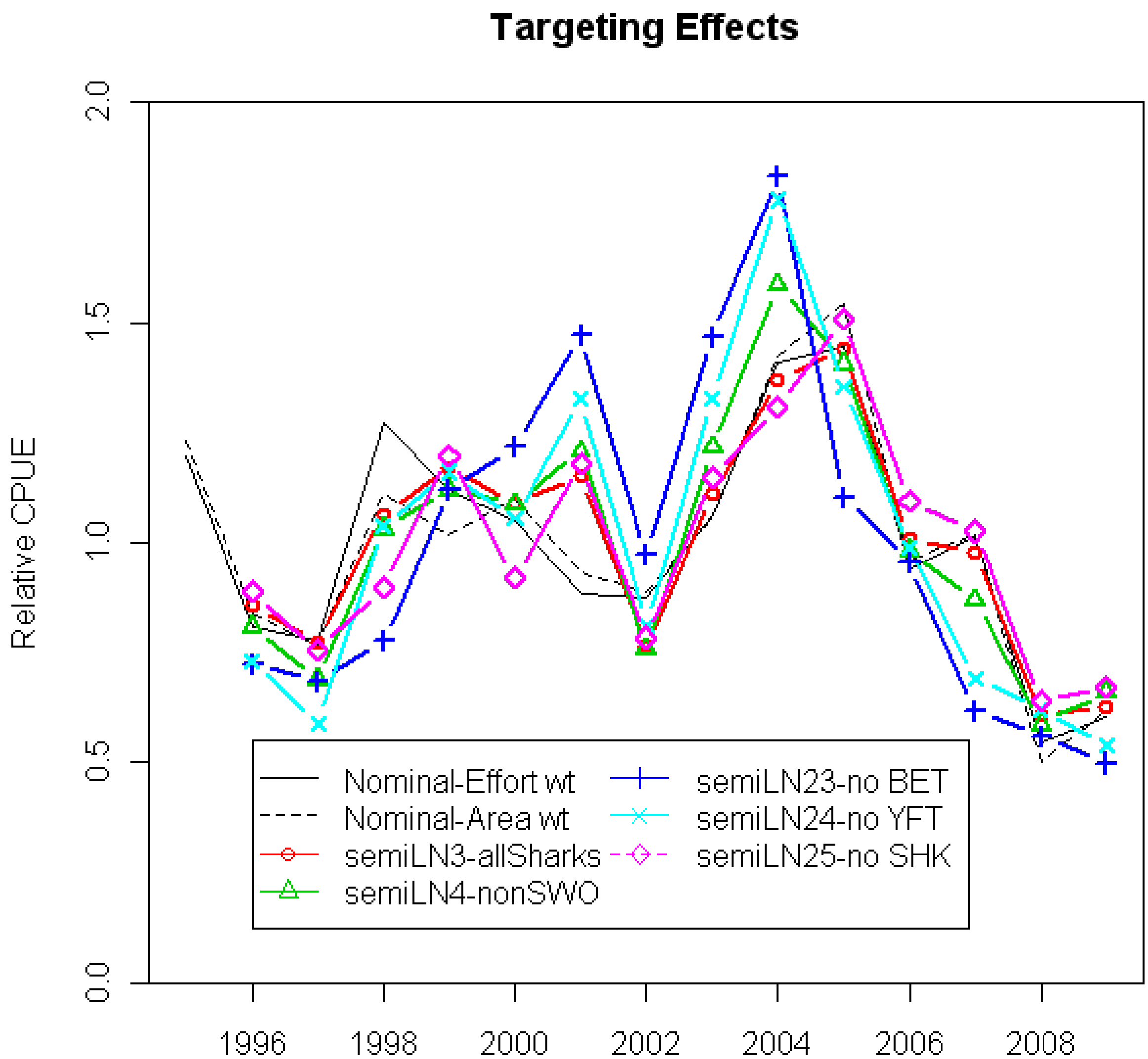


Figure 22. Comparison of Seychelles SEMI swordfish CPUE time series when different approaches are used to identify the effects of shifting targeting.

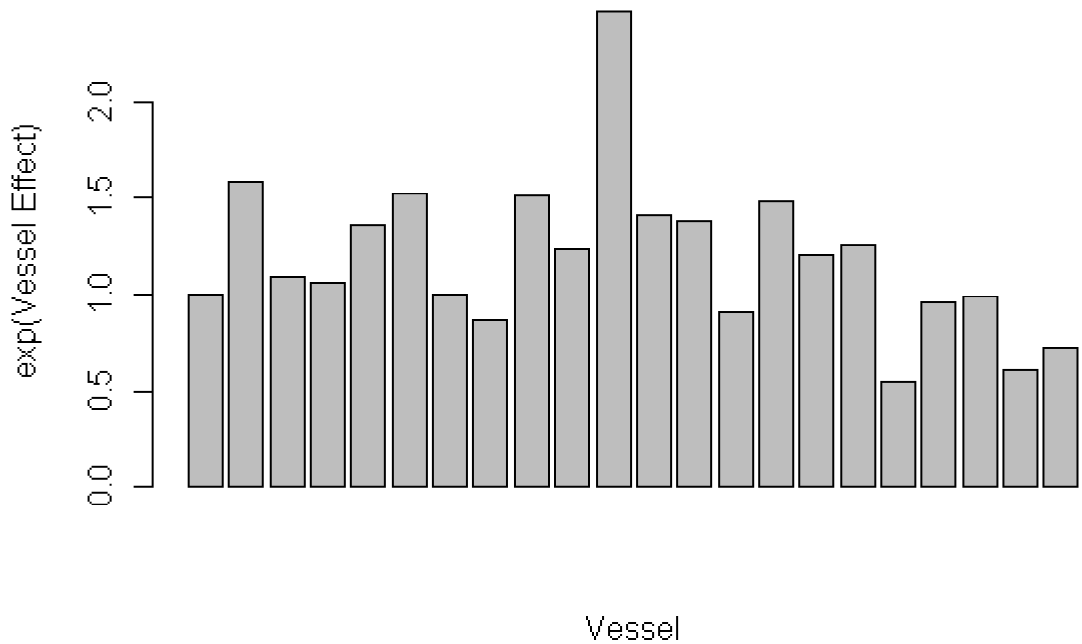


Figure 23. The effect of the individual vessel on swordfish CPUE estimated by model semiLN5.

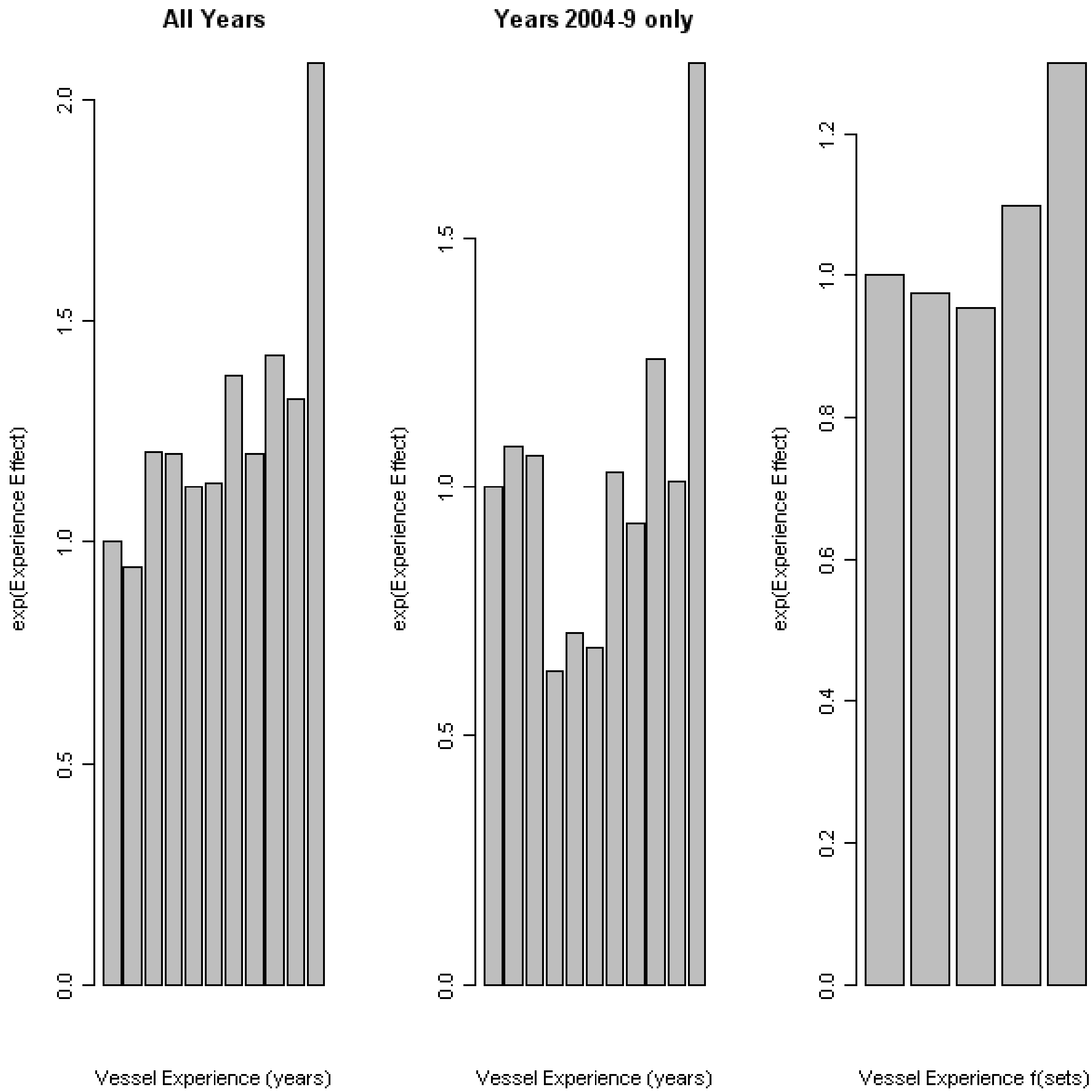


Figure 24. Estimated effects of vessel experience years (model semiLN6) and sets (model semiLN7). The middle panel was from a refitting of semiLN6, including only the last 6 years, where there is a decline in CPUE.

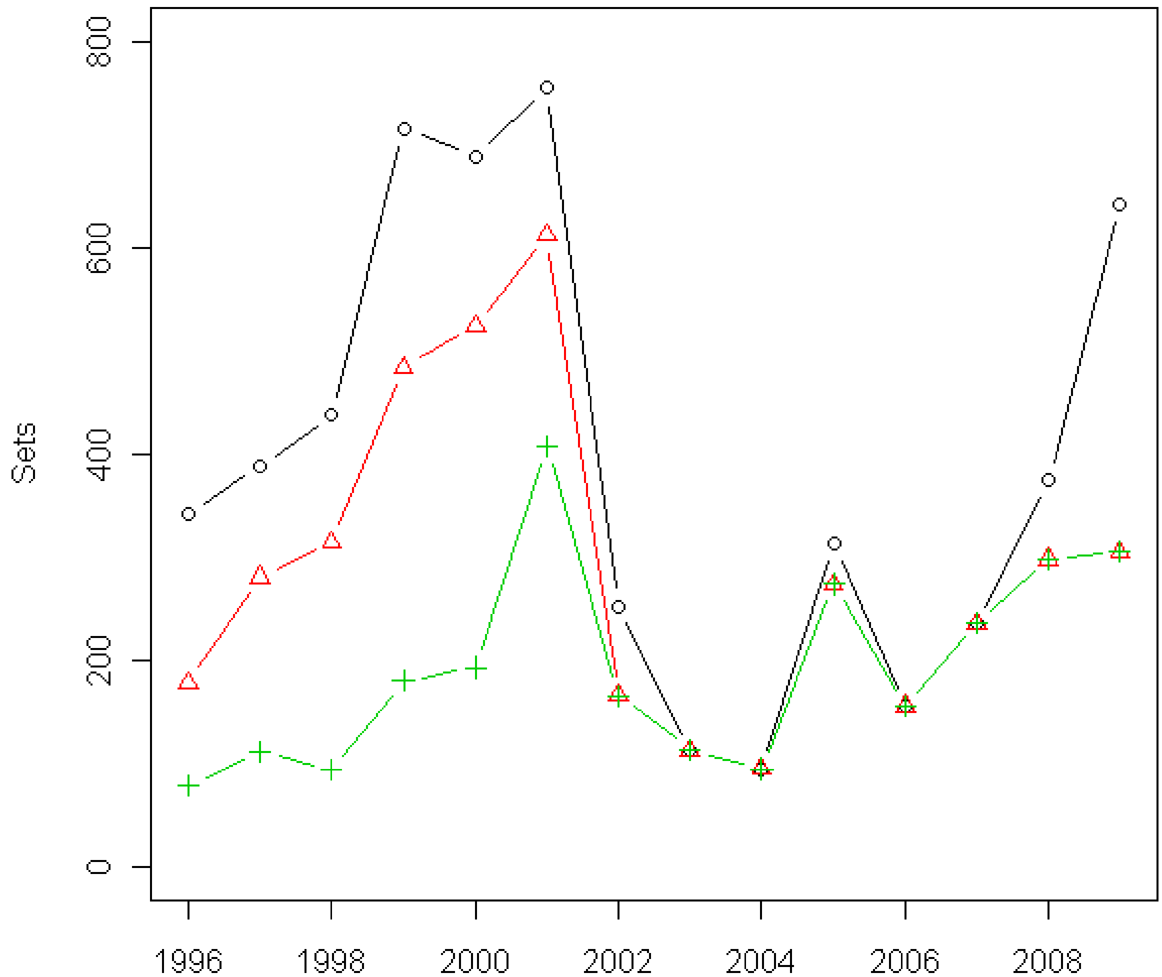


Figure 25. Comparison of the number of sets per year in the full data set (semiLN2, black upper line), the data set resulting when vessels operating for a total of less than 5 years are removed (semiLN21, red middle line) and when vessels operating for a total of less than 7 years are removed (semiLN22, green lower line).

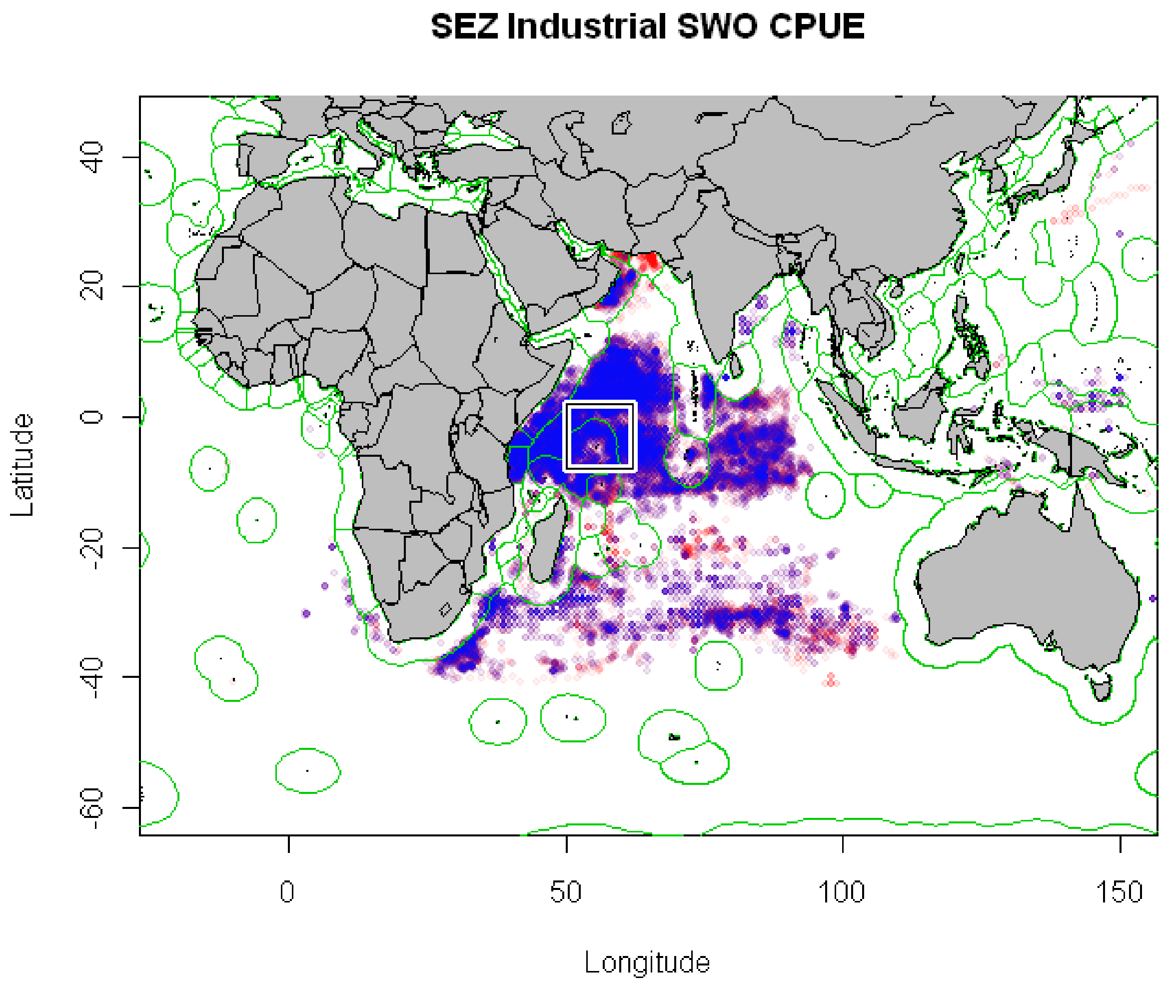


Figure 26. Location of Seychelles industrial longline fleet logbook sets 2000-2008. Red indicates positions of sets with no SWO recorded, blue indicates that SWO were caught (and are superimposed on top of the red). Full colour saturation indicates 20 or more sets. The central rectangle outlines the definition of the Seychelles Region for purposes of comparing with the semi-industrial fleet.

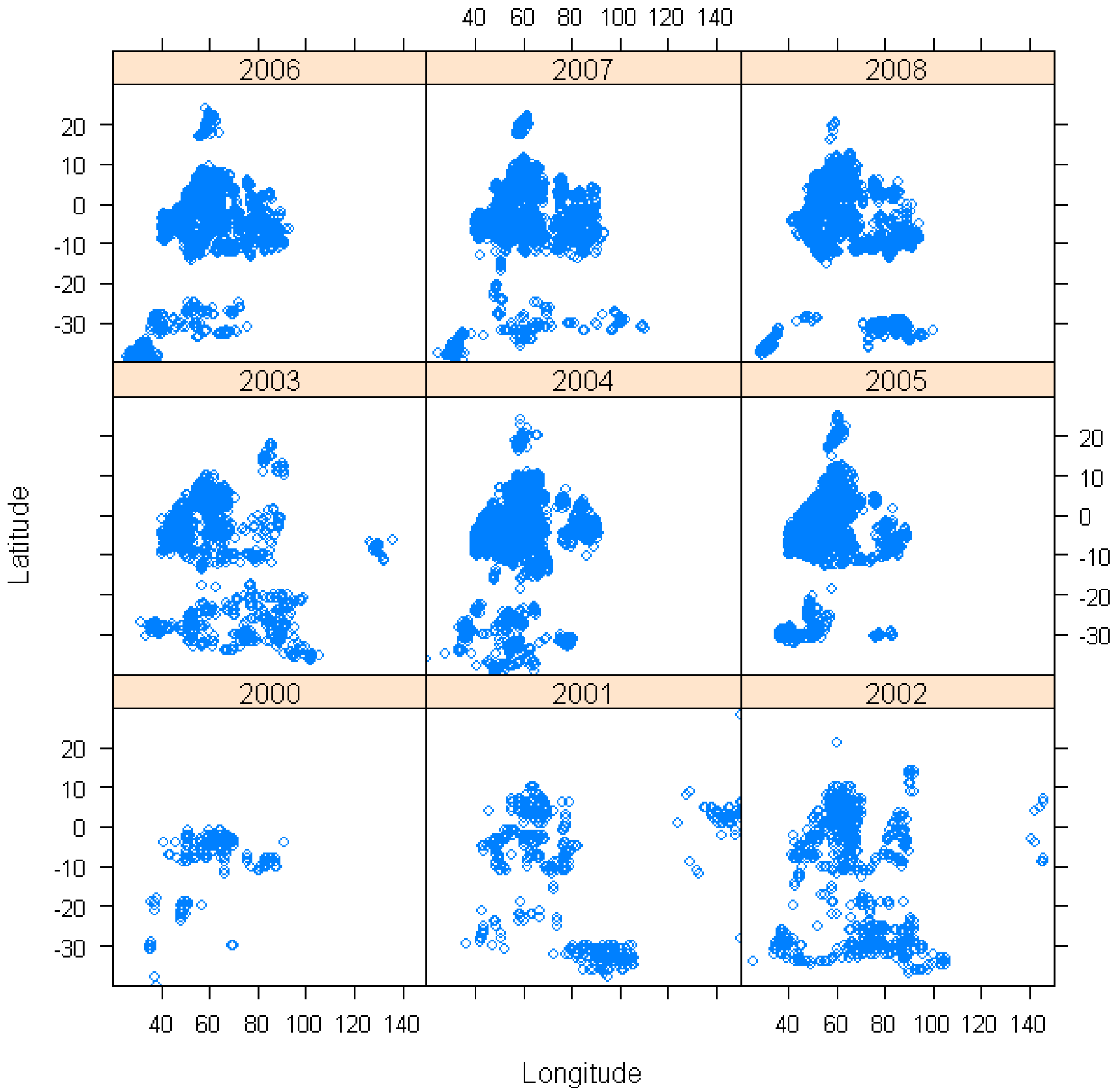


Figure 27. Location of Seychelles industrial longline fleet logbook sets by year (whole Indian Ocean).

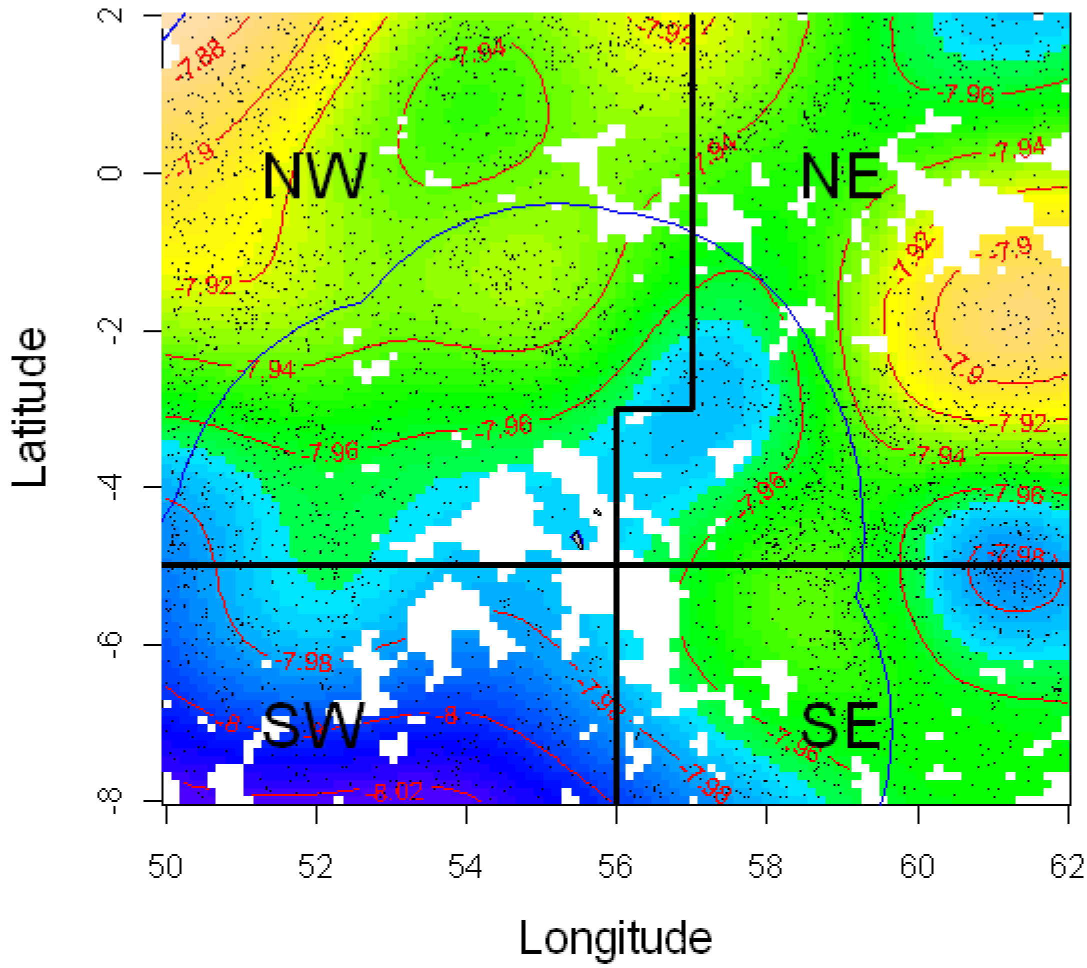


Figure 28. Spatial GAM showing the set locations (points), estimated spatial effects (log(CPUE) contours), and the 4 area structure used in the analyses.

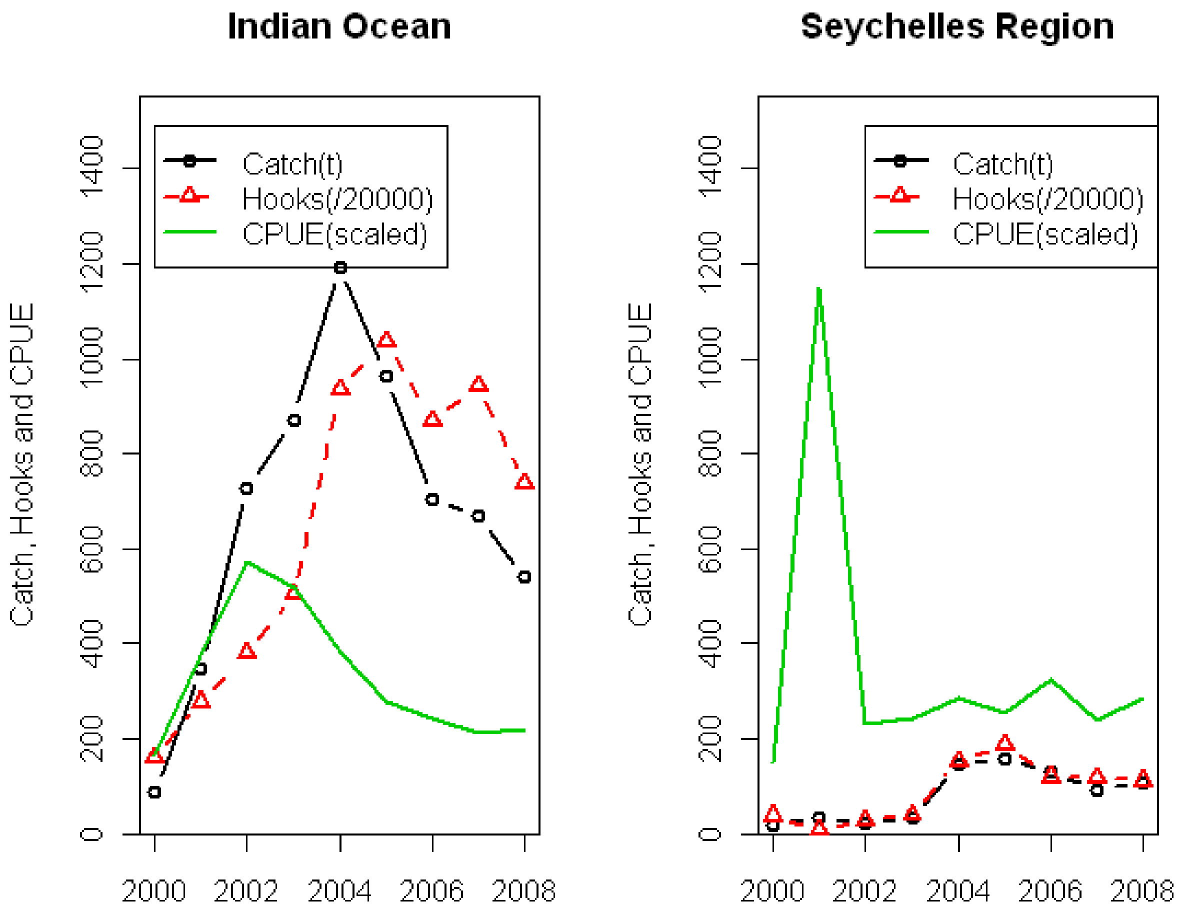


Figure 29. Time series of Total Catch, Total Hooks and Total Catch/ Total Hooks (rescaled to fit on the same axes) for the Seychelles industrial longline fleet operating throughout the Indian Ocean (left panel), and in the Seychelles region only (right panel).

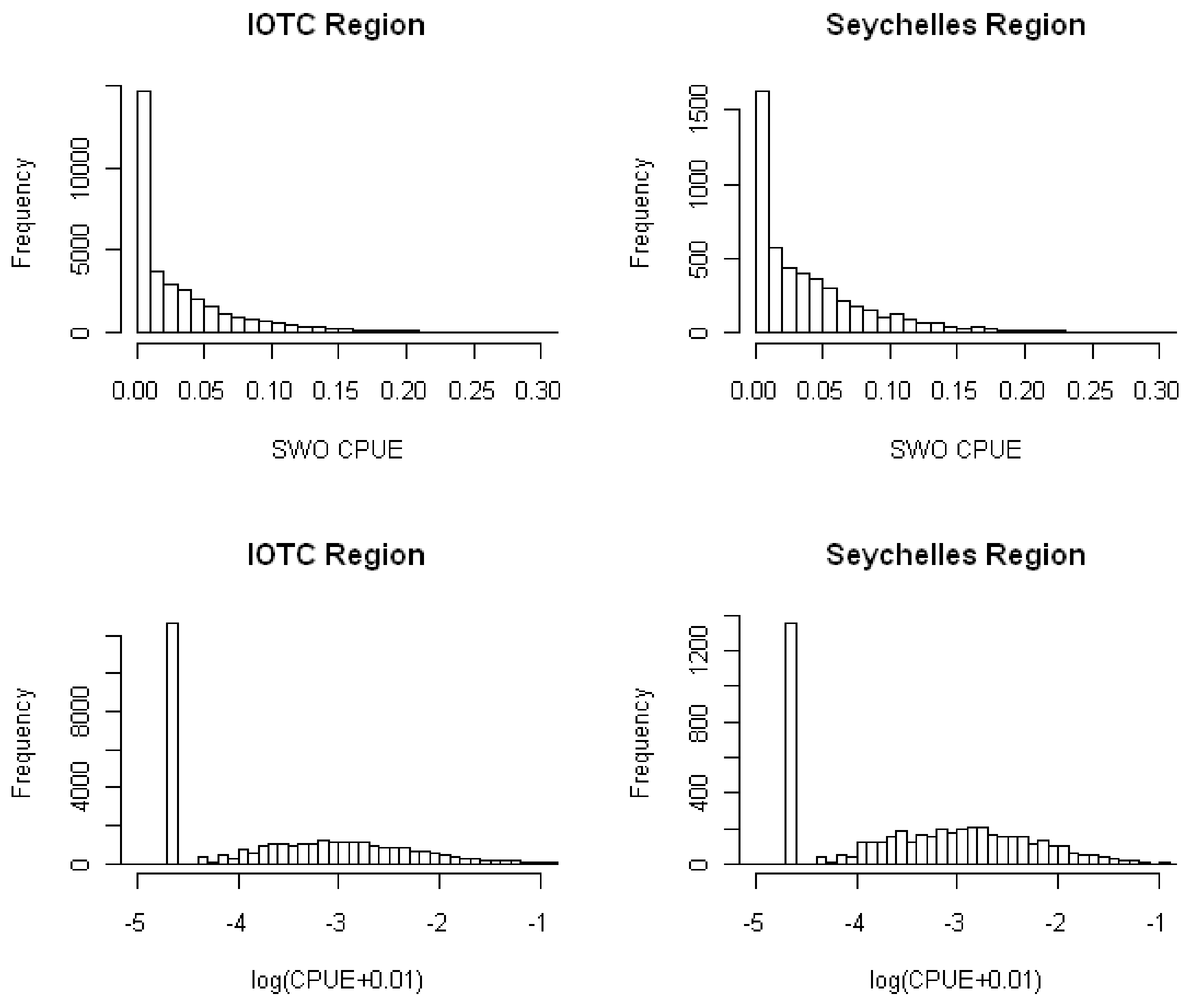


Figure 30. Frequency distribution of Seychelles industrial fleet swordfish CPUE indicating large number of sets with zero catch (35% in the IOTC region, 27% in the Seychelles region).

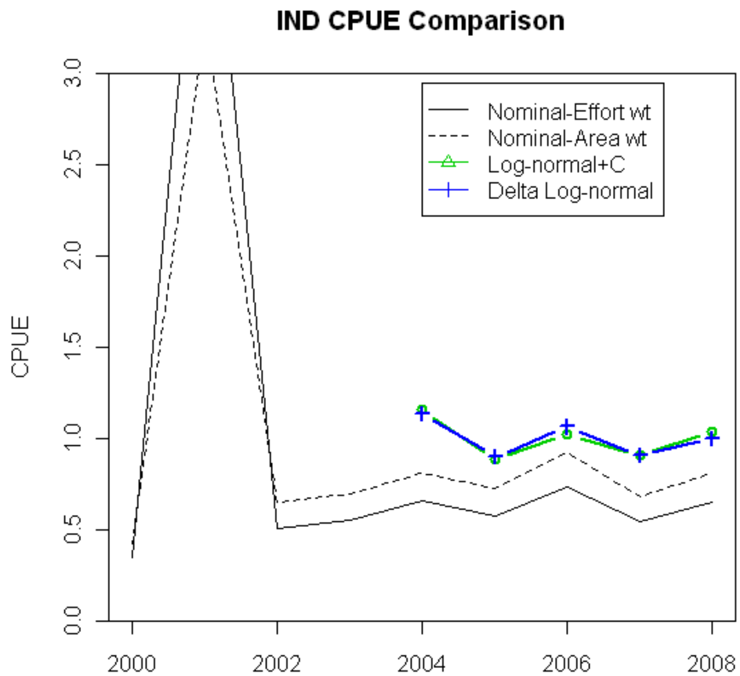


Figure 31. Comparison of the nominal and standardized CPUE series for the IND fleet in the SEZ region (normalized to a mean of unity).

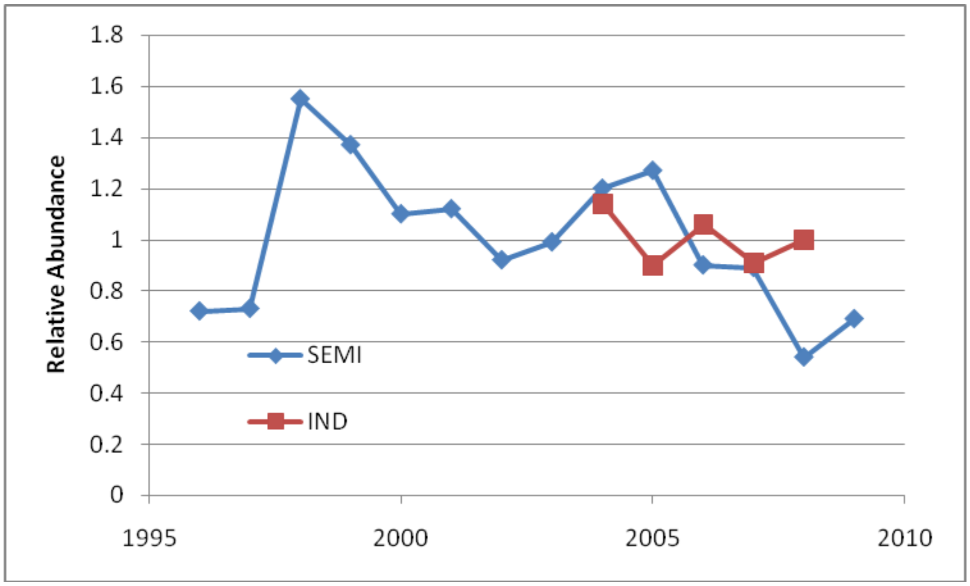
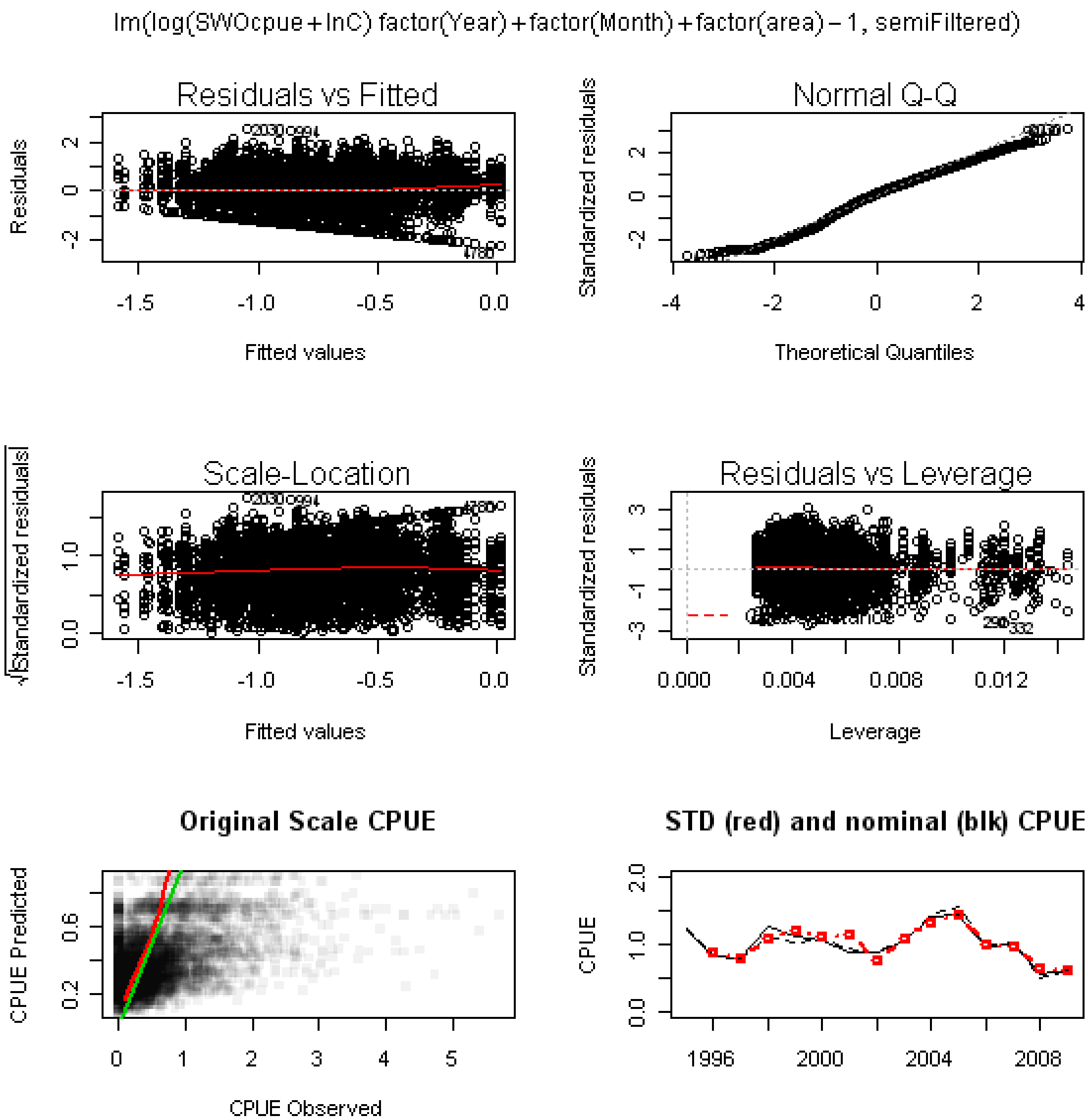


Figure 32. Comparison of the tentatively recommended swordfish relative abundance time series from the Seychelles semi-industrial (SEMI) and industrial (IND) longline fleet CPUE series (normalized to a mean of unity).

Appendix 1. Summary diagnostics and parameter estimates for model semiLN0.



Parameter estimates for Model semiLN0

```
Call:
lm(formula = log(SWOcpue + lnC) ~ factor(Year) + factor(Month) + 
    factor(area) - 1, data = semiFiltered)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.2215 -0.5312  0.1045  0.5807  2.5322 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)    
factor(Year)1996 -0.96787    0.06386 -15.155  < 2e-16 ***
factor(Year)1997 -1.03796    0.06002 -17.294  < 2e-16 ***
factor(Year)1998 -0.81042    0.05916 -13.699  < 2e-16 ***
factor(Year)1999 -0.72498    0.05350 -13.550  < 2e-16 ***
factor(Year)2000 -0.79403    0.05343 -14.862  < 2e-16 ***
factor(Year)2001 -0.76891    0.05420 -14.186  < 2e-16 ***
factor(Year)2002 -1.06928    0.06549 -16.327  < 2e-16 ***
factor(Year)2003 -0.80287    0.08727  -9.200  < 2e-16 ***
factor(Year)2004 -0.65620    0.09575  -6.853 8.02e-12 ***
factor(Year)2005 -0.58832    0.06591  -8.926  < 2e-16 ***
factor(Year)2006 -0.87488    0.07918 -11.049  < 2e-16 ***
factor(Year)2007 -0.89411    0.06932 -12.899  < 2e-16 ***
factor(Year)2008 -1.18114    0.06173 -19.135  < 2e-16 ***
factor(Year)2009 -1.20180    0.05384 -22.321  < 2e-16 ***
factor(Month)2   -0.07145    0.05996  -1.192 0.233435    
factor(Month)3   -0.05853    0.05724  -1.022 0.306593    
factor(Month)4    0.21199    0.05738   3.695 0.000222 ***
factor(Month)5    0.46390    0.05462   8.494  < 2e-16 ***
factor(Month)6    0.42624    0.05651   7.542 5.38e-14 ***
factor(Month)7    0.11369    0.05839   1.947 0.051582 .  
factor(Month)8   -0.25488    0.06458  -3.947 8.01e-05 ***
factor(Month)9   -0.15153    0.06393  -2.370 0.017801 *  
factor(Month)10   0.08685    0.05779   1.503 0.132952    
factor(Month)11  -0.07150    0.06062  -1.180 0.238235    
factor(Month)12   0.02676    0.06254   0.428 0.668790    
factor(area).L   -0.12285    0.02350  -5.228 1.77e-07 ***
factor(area).Q    0.01264    0.02752   0.459 0.645959    
factor(area).C   -0.22333    0.03133  -7.128 1.15e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.8209 on 5482 degrees of freedom
Multiple R-squared: 0.4971,     Adjusted R-squared: 0.4945 
F-statistic: 193.5 on 28 and 5482 DF,  p-value: < 2.2e-16
```

Appendix 2. Summary diagnostics and parameter estimates for model semiBC0.

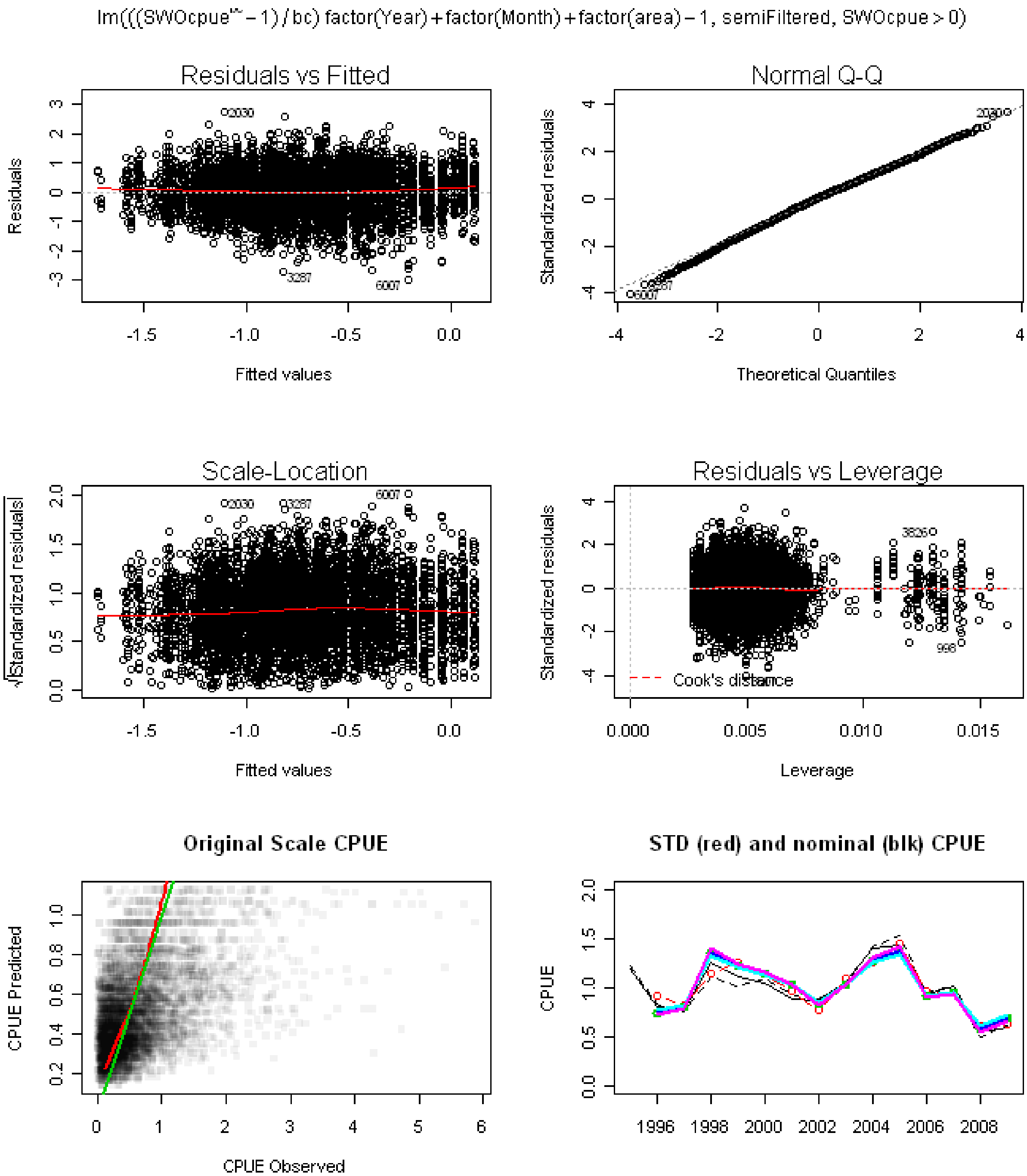


Figure A2-1. Diagnostics for model bc1. Unlike other models considered in this paper, this model is not based in log-space, the parameter effects are additive, and the final time series requires a definition of a standard set. The different coloured lines in the bottom right panel includes the standardized time series calculated for the range of the most extreme definitions adopted for the standard set (in terms of the maximum and minimum values for the Month and Area parameter estimates).

Parameter estimates for model bc1.

```
Call:
lm(formula = ((SWOcpue^bc - 1)/bc) ~ factor(Year) + factor(Month) +
    factor(area) - 1, data = semiFiltered, subset = SWOcpue >
    0)

Residuals:
     Min       1Q   Median       3Q      Max
-3.11275 -0.47846  0.04533  0.49094  2.75256

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
factor(Year)1996 -1.11016    0.06137 -18.090  < 2e-16 ***
factor(Year)1997 -1.05519    0.05859 -18.011  < 2e-16 ***
factor(Year)1998 -0.60256    0.05907 -10.201  < 2e-16 ***
factor(Year)1999 -0.70414    0.05210 -13.515  < 2e-16 ***
factor(Year)2000 -0.75275    0.05310 -14.177  < 2e-16 ***
factor(Year)2001 -0.82676    0.05271 -15.684  < 2e-16 ***
factor(Year)2002 -1.07421    0.06469 -16.605  < 2e-16 ***
factor(Year)2003 -0.83335    0.08415  -9.904  < 2e-16 ***
factor(Year)2004 -0.63000    0.09256  -6.807 1.12e-11 ***
factor(Year)2005 -0.56542    0.06381  -8.862  < 2e-16 ***
factor(Year)2006 -0.94416    0.07654 -12.336  < 2e-16 ***
factor(Year)2007 -0.85216    0.06796 -12.539  < 2e-16 ***
factor(Year)2008 -1.21973    0.06125 -19.915  < 2e-16 ***
factor(Year)2009 -1.17270    0.05373 -21.827  < 2e-16 ***
factor(Month)2   -0.08218    0.05905  -1.392 0.164117
factor(Month)3   -0.04494    0.05663  -0.794 0.427521
factor(Month)4    0.22941    0.05611   4.088 4.42e-05 ***
factor(Month)5    0.54164    0.05361  10.104  < 2e-16 ***
factor(Month)6    0.43905    0.05513   7.963 2.08e-15 ***
factor(Month)7    0.06129    0.05690   1.077 0.281488
factor(Month)8   -0.25182    0.06492  -3.879 0.000106 ***
factor(Month)9   -0.15476    0.06367  -2.431 0.015107 *
factor(Month)10   0.09754    0.05680   1.717 0.085986 .
factor(Month)11  -0.10286    0.05960  -1.726 0.084449 .
factor(Month)12  -0.05598    0.06102  -0.917 0.358997
factor(area).L   -0.17074    0.02309  -7.395 1.67e-13 ***
factor(area).Q    0.07393    0.02684   2.754 0.005902 **
factor(area).C   -0.22915    0.03066  -7.473 9.26e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7444 on 4730 degrees of freedom
Multiple R-squared: 0.5333,     Adjusted R-squared: 0.5305
F-statistic:   193 on 28 and 4730 DF,  p-value: < 2.2e-16
```

Appendix 3. Summary diagnostics and results for model semiDL0.

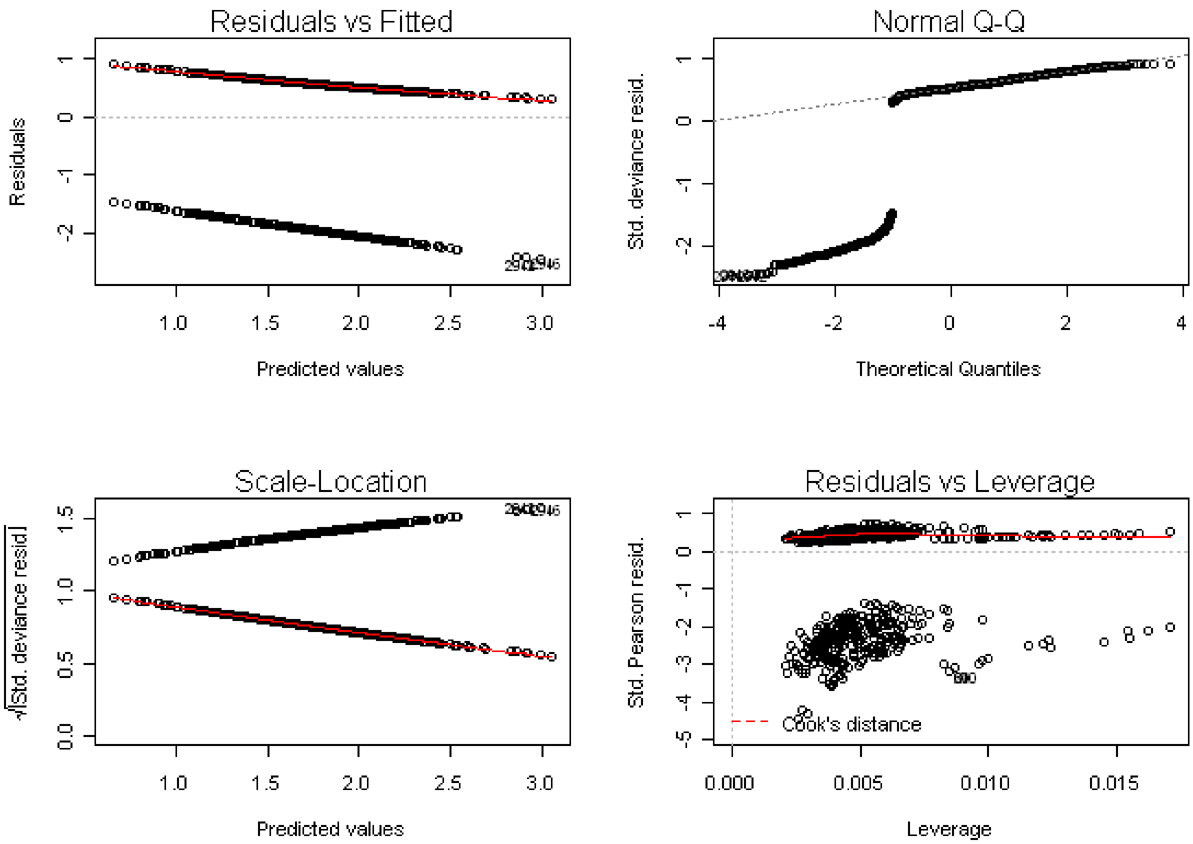


Figure A3-1. Standard diagnostics for the binomial portion of the semiDL0 model (they may not be particularly useful for the binomial model).

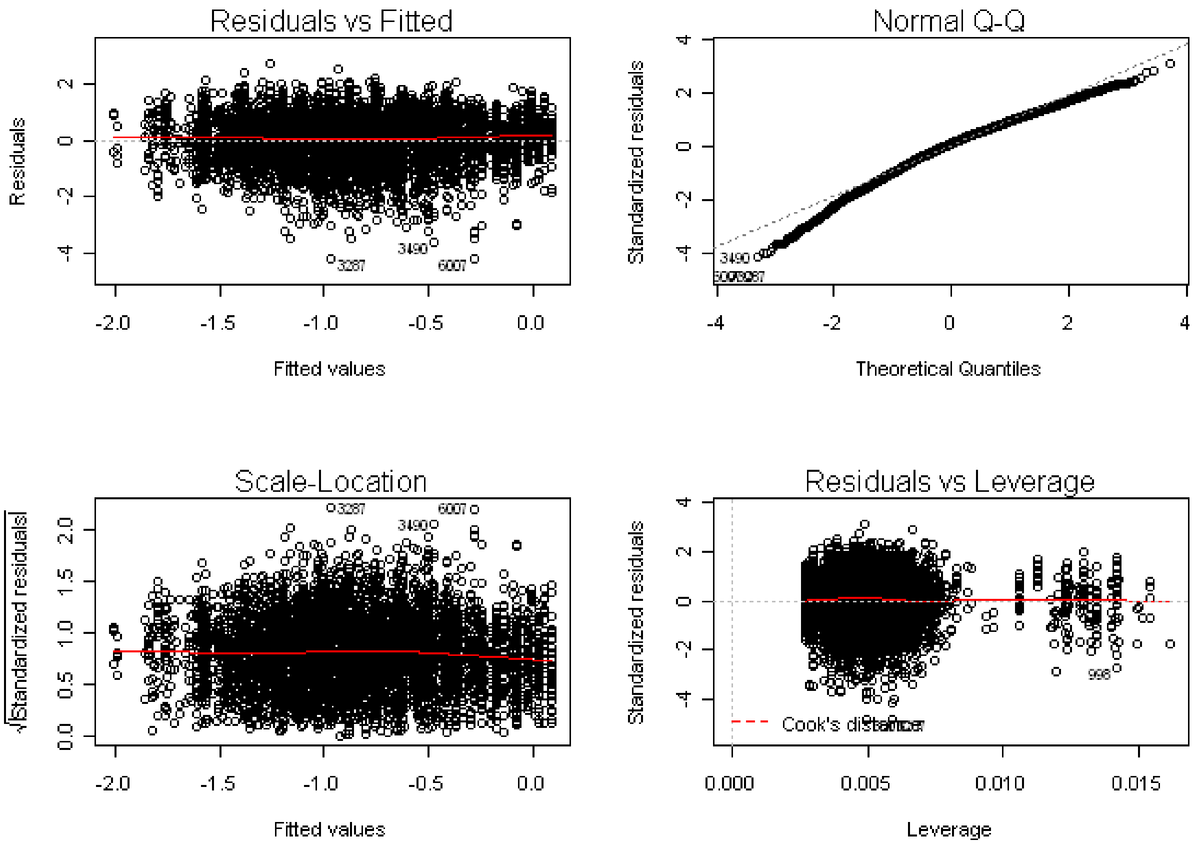


Figure A3-2. Diagnostics for the lognormal portion of the semiDL0 model.

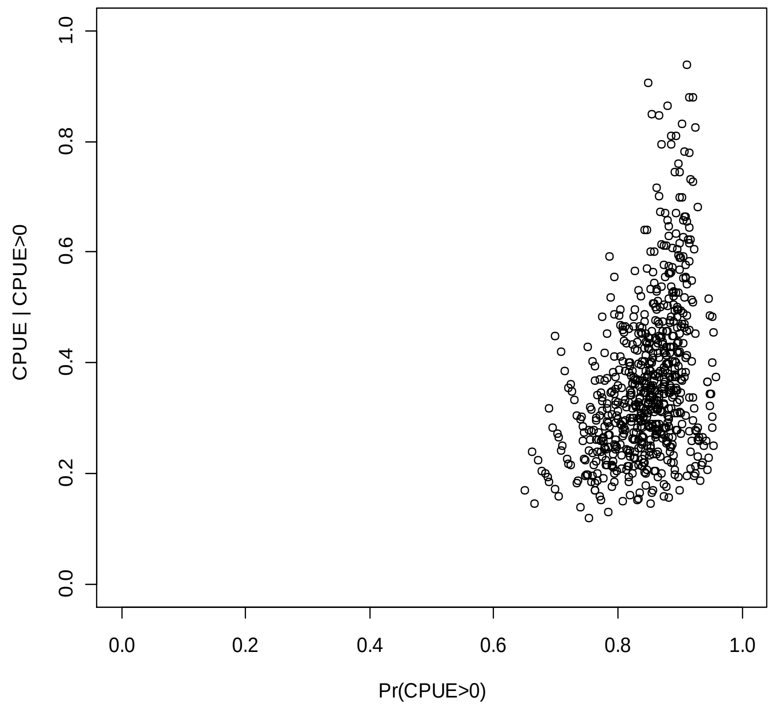


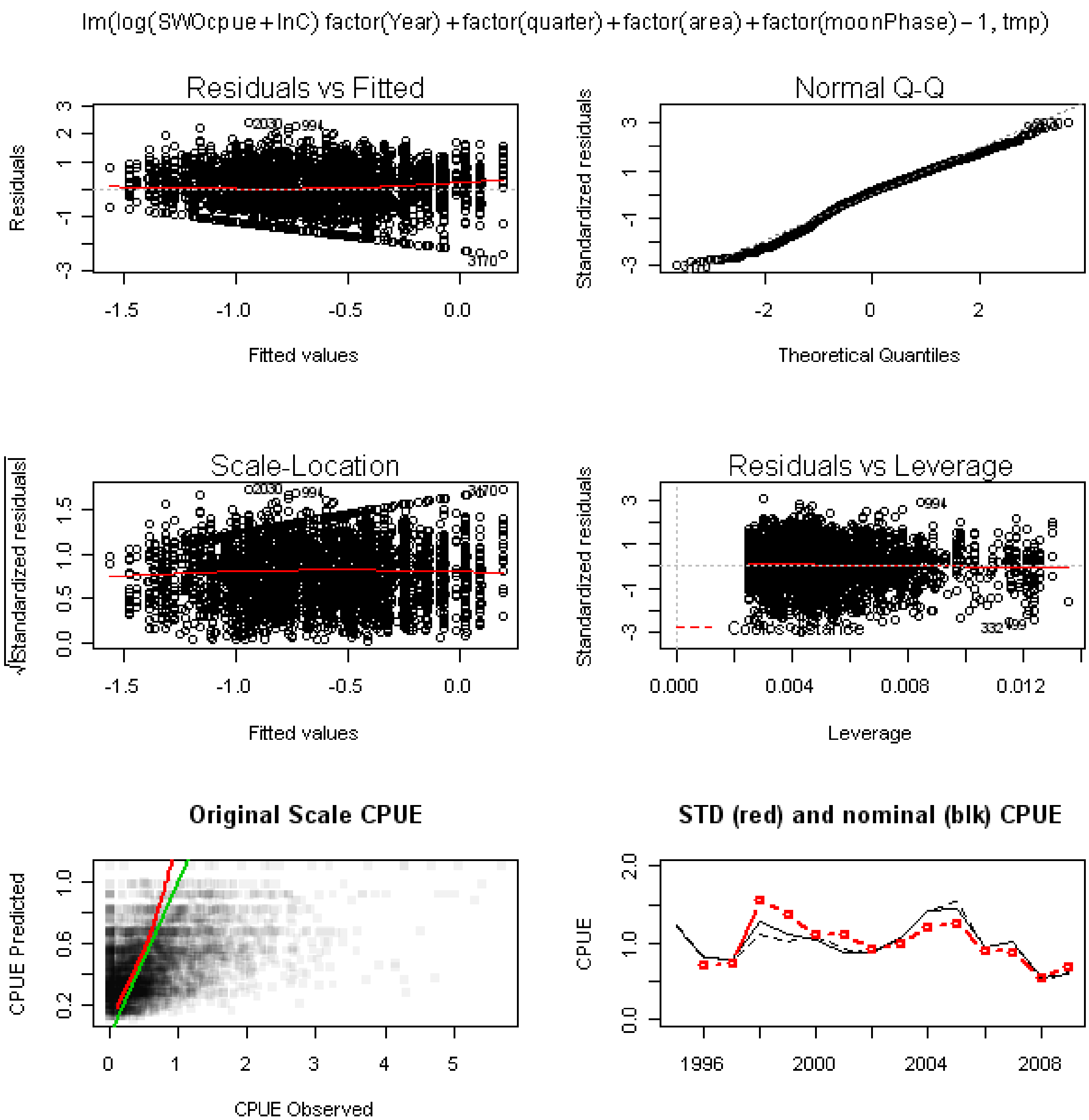
Figure A3-3. Estimated probability that CPUE>0 and estimated median CPUE given that CPUE >0 for each Year-Month-Area strata in model semiDL0.

Appendix 4. Automated model selection using R function step() and BIC criteria to remove terms from the fullest model under consideration.

```
tmp2 <- step( lm(log(SWOcpue + lnC) ~ factor(Year)*factor(quarter)*factor(area) -
factor(Year):factor(quarter):factor(area)  + factor(moonPhase) + factor(VessHistoryID)
+ nonSWO - 1, data=semiFiltered), k=log(n))

Start:  AIC=-2224.23
log(SWOcpue + lnC) ~ factor(Year) * factor(quarter) * factor(area) -
    factor(Year):factor(quarter):factor(area) + factor(moonPhase) +
    factor(VessHistoryID) + nonSWO - 1

                               Df Sum of Sq    RSS     AIC
- factor(Year):factor(area)    39    56.602 3064.4 -2457.5
- factor(Year):factor(quarter) 38    99.865 3107.7 -2371.6
- factor(quarter):factor(area)  9    38.286 3046.1 -2232.1
<none>                                      3007.8 -2224.2
- factor(moonPhase)             1     9.122 3016.9 -2216.2
- factor(VessHistoryID)        21   277.454 3285.3 -1919.0
- nonSWO                        1   187.301 3195.1 -1900.0

Step:  AIC=-2457.46
log(SWOcpue + lnC) ~ factor(Year) + factor(quarter) + factor(area) +
    factor(moonPhase) + factor(VessHistoryID) + nonSWO + factor(Year):factor(quarter)
+
    factor(quarter):factor(area) - 1

                               Df Sum of Sq    RSS     AIC
- factor(Year):factor(quarter) 38   116.724 3181.1 -2578.8
<none>                                      3064.4 -2457.5
- factor(quarter):factor(area)  9    46.491 3110.9 -2452.0
- factor(moonPhase)             1    11.060 3075.5 -2446.2
- nonSWO                        1   183.023 3247.4 -2146.4
- factor(VessHistoryID)        21   315.380 3379.8 -2098.6

Step:  AIC=-2578.83
log(SWOcpue + lnC) ~ factor(Year) + factor(quarter) + factor(area) +
    factor(moonPhase) + factor(VessHistoryID) + nonSWO + factor(quarter):factor(area)
-
    1

                               Df Sum of Sq    RSS     AIC
- factor(quarter):factor(area)  9     41.42 3222.6 -2585.1
<none>                                      3181.1 -2578.8
- factor(moonPhase)             1     13.41 3194.6 -2564.3
- nonSWO                        1    159.04 3340.2 -2318.6
- factor(VessHistoryID)        21    321.94 3503.1 -2228.6
- factor(Year)                 14    669.03 3850.2 -1647.7

Step:  AIC=-2585.07
log(SWOcpue + lnC) ~ factor(Year) + factor(quarter) + factor(area) +
    factor(moonPhase) + factor(VessHistoryID) + nonSWO - 1

                        Df Sum of Sq    RSS     AIC
<none>                               3222.6 -2585.1
- factor(moonPhase)      1     12.38 3235.0 -2572.6
- factor(area)           3     79.43 3302.0 -2476.7
- nonSWO                 1    151.54 3374.1 -2340.5
- factor(quarter)        3    192.47 3415.0 -2291.3
```

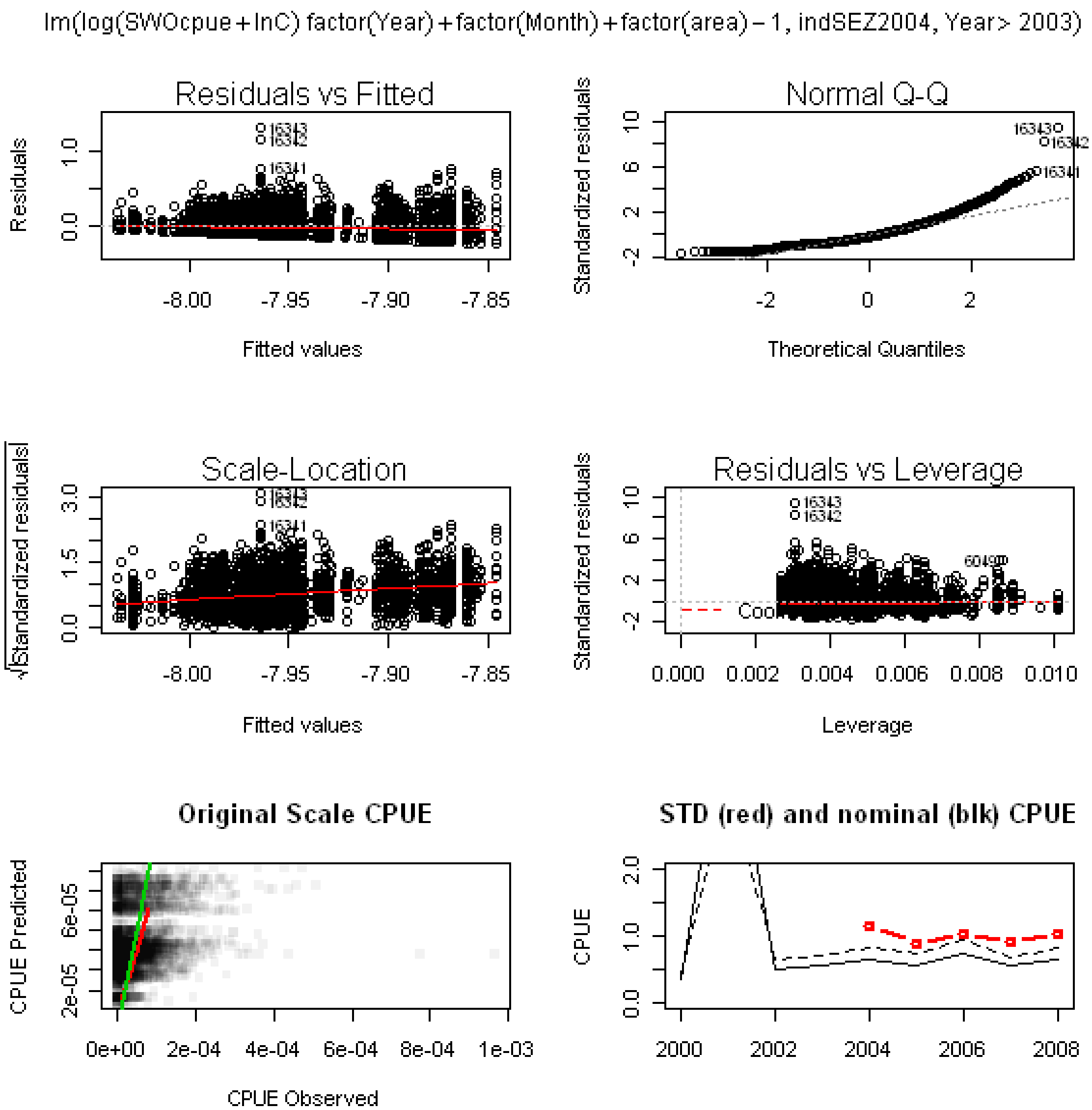
```
- factor(VessHistoryID) 21     331.07 3553.6 -2227.1
- factor(Year)          14     762.14 3984.7 -1536.0
```

Appendix 5. Diagnostics and results from model semiLN21. This is the tentatively recommended SEMI relative abundance series.

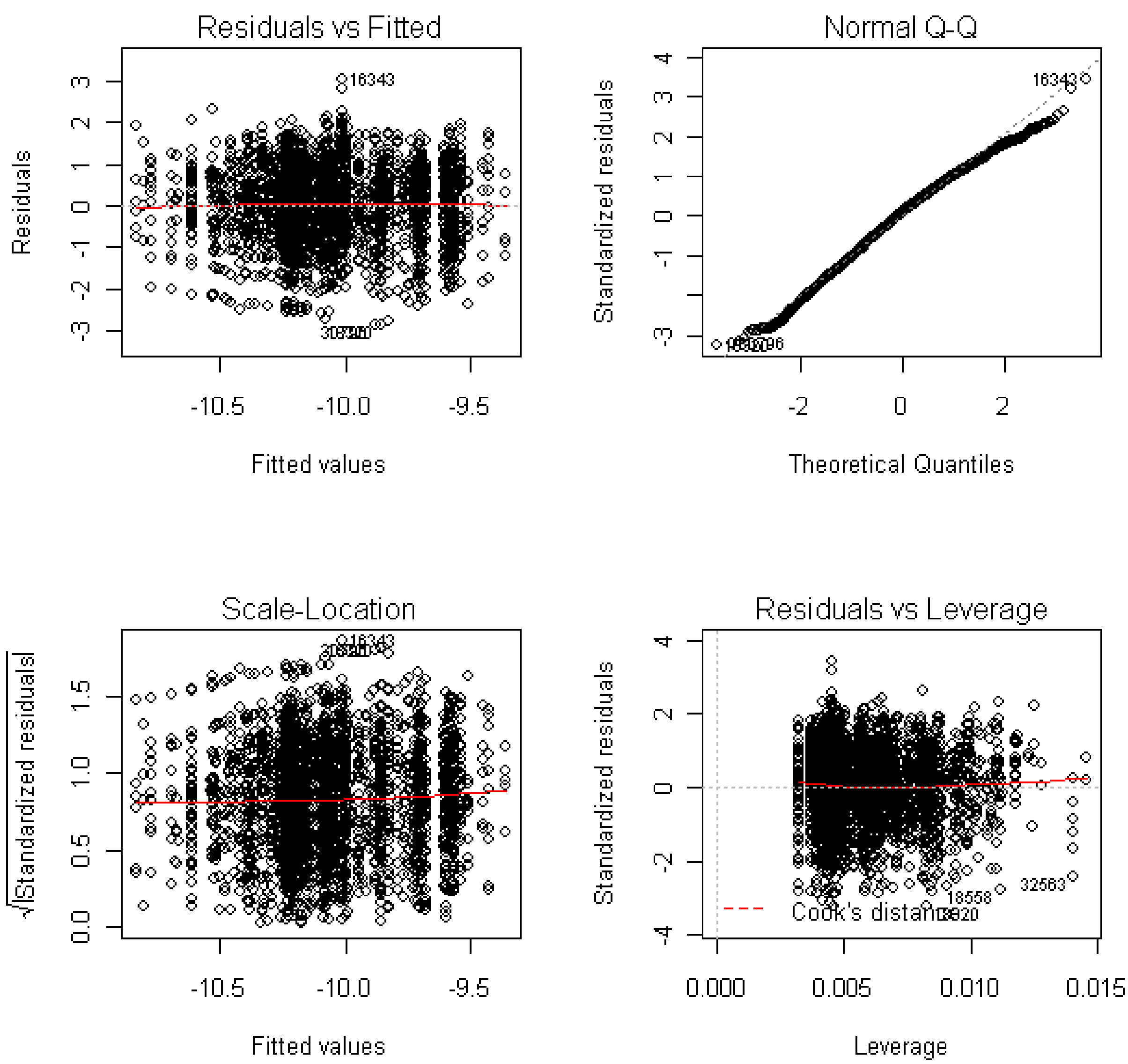


```
Call:
lm(formula = log(SWOcpue + lnC) ~ factor(Year) + factor(quarter) +
    factor(area) + factor(moonPhase) - 1, data = tmp)

Residuals:
    Min      1Q  Median      3Q     Max
-2.3967 -0.4940  0.1095  0.5565  2.4354

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
factor(Year)1996     -0.96803    0.07058 -13.716  < 2e-16 ***
factor(Year)1997     -0.95864    0.05554 -17.261  < 2e-16 ***
factor(Year)1998     -0.36875    0.05540  -6.657 3.18e-11 ***
factor(Year)1999     -0.47106    0.04811  -9.792  < 2e-16 ***
factor(Year)2000     -0.64754    0.04648 -13.931  < 2e-16 ***
factor(Year)2001     -0.63331    0.04579 -13.832  < 2e-16 ***
factor(Year)2002     -0.78739    0.06990 -11.264  < 2e-16 ***
factor(Year)2003     -0.73714    0.07984  -9.233  < 2e-16 ***
factor(Year)2004     -0.57969    0.08697  -6.666 2.99e-11 ***
factor(Year)2005     -0.53779    0.05894  -9.125  < 2e-16 ***
factor(Year)2006     -0.80598    0.07073 -11.395  < 2e-16 ***
factor(Year)2007     -0.81570    0.06070 -13.439  < 2e-16 ***
factor(Year)2008     -1.16029    0.05567 -20.842  < 2e-16 ***
factor(Year)2009     -0.99339    0.05492 -18.088  < 2e-16 ***
factor(quarter)2      0.39875    0.03485  11.443  < 2e-16 ***
factor(quarter)3     -0.03261    0.03965  -0.822    0.411
factor(quarter)4      0.05233    0.03850   1.359    0.174
factor(area).L       -0.10274    0.02638  -3.895 9.96e-05 ***
factor(area).Q        0.04077    0.03076   1.325    0.185
factor(area).C       -0.31360    0.03581  -8.756  < 2e-16 ***
factor(moonPhase)new -0.16508    0.02693  -6.129 9.68e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.806 on 4016 degrees of freedom
Multiple R-squared: 0.4542,     Adjusted R-squared: 0.4514
F-statistic: 159.2 on 21 and 4016 DF,  p-value: < 2.2e-16
```

Appendix 6. Summary diagnostics and results for the industrial fleet model indLN1.



Appendix 7. Summary diagnostics (for the lognormal portion of the delta model) and results for the industrial fleet model indDLN1.



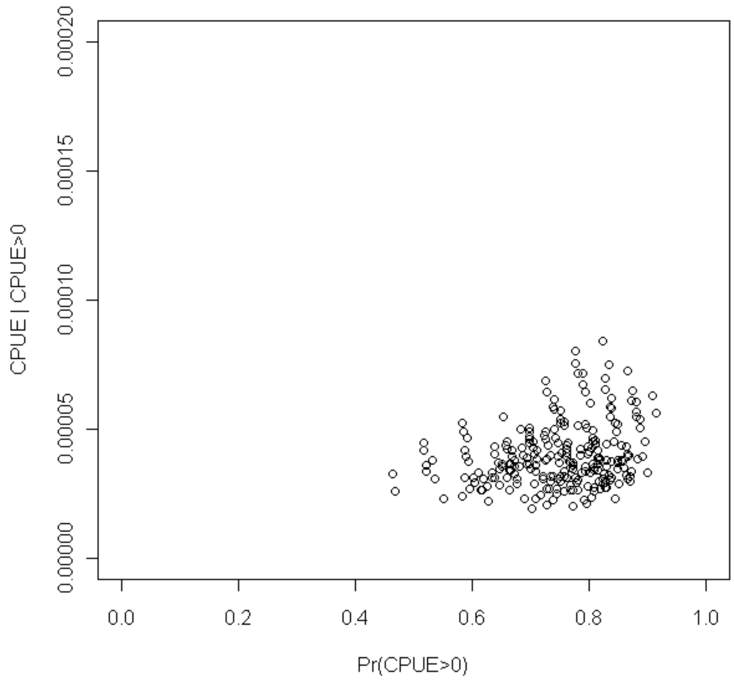


Figure A7-2. Estimated probability that CPUE>0 and estimated median CPUE given that CPUE >0 for each Year-Month-Area strata in model indDL1.