Ministry for Primary Industries
Manatū Ahu Matua

Fishery characterisations and catch-per-unit-effort indices for three sub-stocks of snapper SNA 1, 1989–90 to 2009–10

New Zealand Fisheries Assessment Report 2012/29

J.R. McKenzie
D.M. Parsons

ISSN 1179-5352 (online)
ISBN 978-0-478-40009-0 (online)

July 2012

New Zealand Government

Growing and Protecting New Zealand

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:
http://www.mpi.govt.nz/news-resources/publications.aspx
http://fs.fish.govt.nz go to Document library/Research reports

Contents

EXECUTIVE SUMMARY

McKenzie, J.R.; Parsons, D.M. (2012). Fishery characterisations and catch-per-unit-effort indices for three sub-stocks of snapper SNA 1, 1989–90 to 2009–10.

New Zealand Fisheries Assessment Report 2012/29. **112 p.**

CPUE analyses conducted for the three SNA1 substocks produced variable results. The longline analyses for the Hauraki Gulf and the Bay of Plenty generally increased between 1989 and 2009. The Bay of Plenty trawl analysis also had a generally increasing trend, but suggested that CPUE decreased in the final two years of the series. In the East Northland longline analysis, CPUE fluctuated throughout the 21 year time series and ended at a similar level to when the index began in 1989.

The fit of the GLM models to the catch data was generally good and only small departures from the assumptions of this analysis method should be expected. The longline models explained a much larger amount of variation in the data than the Bay of Plenty trawl analysis (39–55% versus 21 %). The most important explanatory variables for the longline analysis were the number of hooks set and the identification number of each vessel. For the Bay of Plenty trawl analysis the most important explanatory variables were the target of the fishing event and the identification number of each vessel. A true effort variable (e.g. tow duration) had poor explanatory power for the Bay of Plenty trawl index.

1 INTRODUCTION

Snapper (*Pagrus auratus*) is a demersal finfish species of the family Sparidae (sea breams). New Zealand snapper are found around the North Island and northern South Island inhabiting estuarine and coastal waters out to a depth of 200 m. The genus *Pagrus* is not restricted to New Zealand. Representative species are found in south Australia, south America, and Japan (Paul 1976).

Snapper is New Zealand's most valuable commercial coastal marine species and, by virtue of its high abundance around the populous regions of northern New Zealand, it is also the nation's most important recreational species (Hartill et al 2007).

Under the Quota Management System (QMS) there are four snapper Quota Management Areas (QMAs) of commercial and non-commercial significance: SNA 1; SNA 2; SNA 7; SNA 8 (Figure 1). The largest volume of catch, both commercial and non-commercial, comes from the east coast Northland QMA known as SNA 1 (Figure 1).

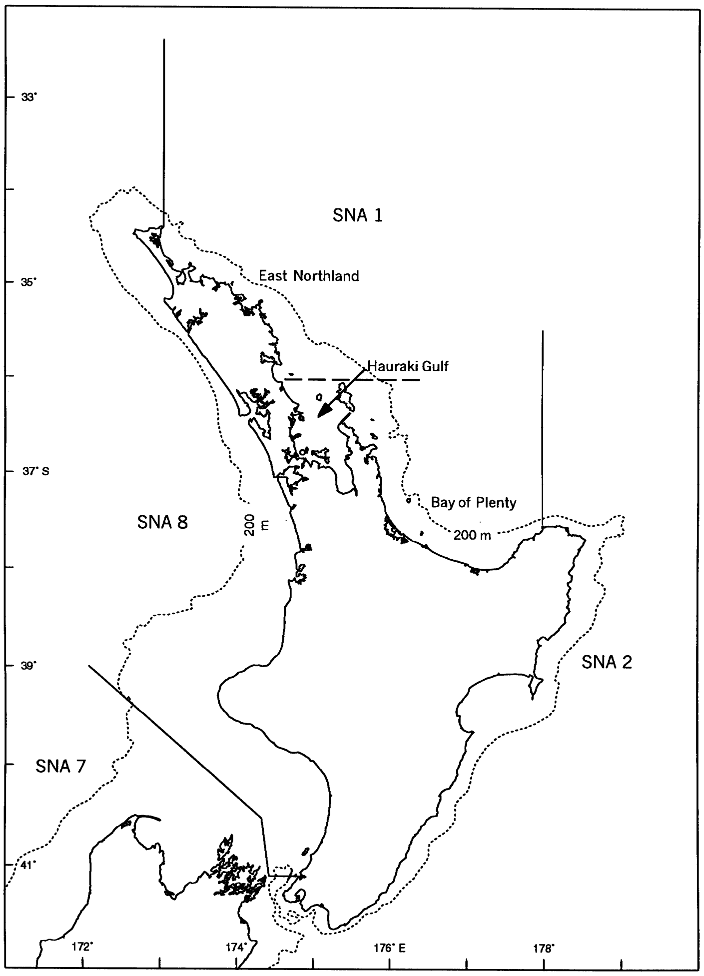


Figure 1: Snapper Quota Management Area boundaries and the SNA 1 sub-stock boundaries of East Northland, Hauraki Gulf and Bay of Plenty.

SNA 1 comprises of three sub-stocks: East Northland, Hauraki Gulf, and Bay of Plenty (Figure 1). These sub-stocks are considered separately for assessment purposes (McKenzie in pub). Longline is the predominant commercial method in SNA 1 and accounts for 60–80% of the annual commercial catch (Walsh et al. 2009).

The annual Total Allowable Commercial Catch (TACC) for SNA 1 is currently 4 500 t. The annual SNA 1 catch has closely matched the TACC since the inception of the Quota Management System (QMS) in October 1986 (Figure 2).

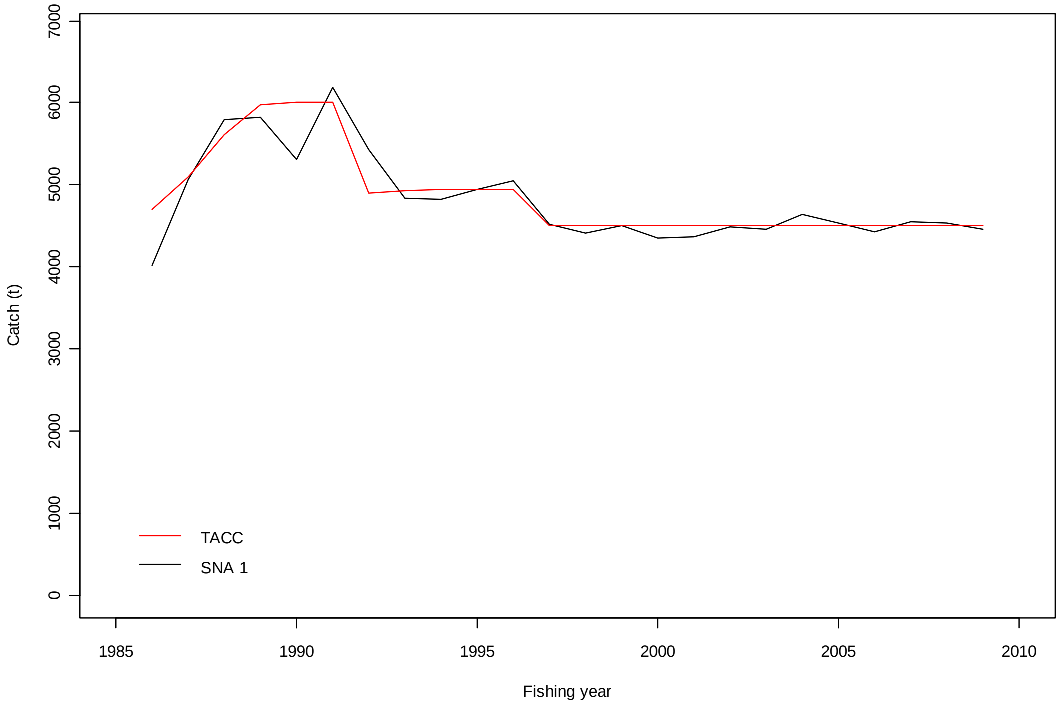


Figure 2: Reported landings of snapper (t) in SNA 1 from 1986–87 to 2009–10 and gazetted and actual TACCs (t) for 1986–87 to 2008–09.

1.1 Previous work

Literally millions of dollars have been spent monitoring SNA 1 since the early 1980s. Monitoring programmes have included commercial catch-at-age sampling, recreational harvest surveys, trawl surveys, and tagging programmes to derive biomass (McKenzie et al 2011). Given the wealth of monitoring information available from SNA 1 the preference has been to use age-structured population modelling to estimate productivity and stock status (McKenzie in pub).

The last formal SNA 1 assessment in which yield estimates were derived was undertaken in 1999 (Gilbert et al 2000). The assessment estimated the Hauraki Gulf/Bay of Plenty component of SNA 1 at 0.80 B_{MSY} in 1999–2000; this sub-stock was predicted to rebuild over the next 20 years reaching about 1.73 B_{MSY} by 2019–20. The East Northland component of SNA 1 was predicted to have been at or slightly below B_{MSY} in 1999–2000; with 95% probability of the sub-stock biomass increasing over the next 20 years (Gilbert et al 2000).

Standardised longline Catch per unit effort (CPUE) indices covering the fishing year period 1989–90 to 2004-05 were derived for the three SNA 1 sub-stocks by McKenzie (2008). In addition McKenzie (2008) derived a standardised CPUE bottom trawl index for the Bay of Plenty sub-stock, covering the fishing years 1995–96 to 2004–05. The overall trend in the standardised CPUE indices for East Northland, Hauraki Gulf and Bay of Plenty between 1989 and 2005 was increasing. However, the residuals in the log-linear model fits to the longline catch effort data were not normally distributed and therefore may not have provided an appropriate representation of the underlying variation in the data. An increasing trend was also evident in the Bay of Plenty single trawl standardised series but the increase was not as large as seen in the longline index. Residuals in the Bay of Plenty log-linear model fit to the single trawl data series were more normally distributed. In the current document we present a characterisation of the SNA1 fishery and an update of CPUE analyses up to the 2009–10 fishing year.

2 METHODS

2.1 Data sources

2.1.1 SNA 1 recent (2000–01 to 2009–10 fy) fishery profile data

A characterisation of patterns in the SNA 1 fishery over the period October 2000 to September 2010 was undertaken using data extracted from the Ministry of Fisheries commercial catch reporting system (Appendix 1). The dataset extracted included all effort details and associated catch weights (all species including snapper) from all trips landing SNA 1 catch.

Data obtained from the ministry was groomed and checked for typical reporting errors (Appendix 2). Information to perform the characterisation was compiled into two tables:

1. Landed catch weight: A file containing the verified green (unprocessed) landed weight of all SNA 1 trips.
2. Trip effort data: A file containing demographic information (location method, target species etc).

Although the Trip effort data table has information on catch, these are only fisher estimates. The process followed was to prorate the actual trip landed weight totals across the effort information (i.e. individual sets, tows or days) on the basis of the estimated catch ratios. The link between the two data tables was the common trip number field (trip_key).

2.1.2 Longline CPUE data

SNA 1 longline catch and effort data were obtained from the Ministry of Fisheries (Appendix 3). The criteria for the extract were intended to capture all effort information for all longline trips landing catch from SNA 1. The extract should have captured the SNA 1 longline positive catch and effort information, but is unlikely to have captured all legitimate "zero" catch sets, i.e. target sets from a trip where no snapper was caught. Due to the very low percentage of zero catches in the SNA 1 longline fishery (less than 2% McKenzie 2008) the exclusion of zero catches is unlikely to bias the CPUE results.

Longline CPUE data were available between the 1989–90 and 2009–10 fishing years (21 year time series) for all of the three sub-stocks of SNA1 (East Northland, Hauraki Gulf, and Bay of Plenty). During this period there was a change in catch and effort reporting requirements (Figure 3). Up until the 2007-08 fishing year the majority of longline vessels reported on Catch Effort Landing Return

(CELR) forms (except vessels greater than 28 m which reported on the more detailed Lining Catch, Effort Return (LCER) form). The CELR forms provide day aggregated catch and effort information at the spatial resolution of large statistical reporting areas. From October 2007 onwards (the 2007–08 Fishing Year) all longline vessels between 6 and 28 m in length were required to report on more detailed Lining Trip Catch, Effort Return (LTCER) forms. As a result, data from 1 October 2007 contains effort information for every line set including associated start latitude and longitude positions.

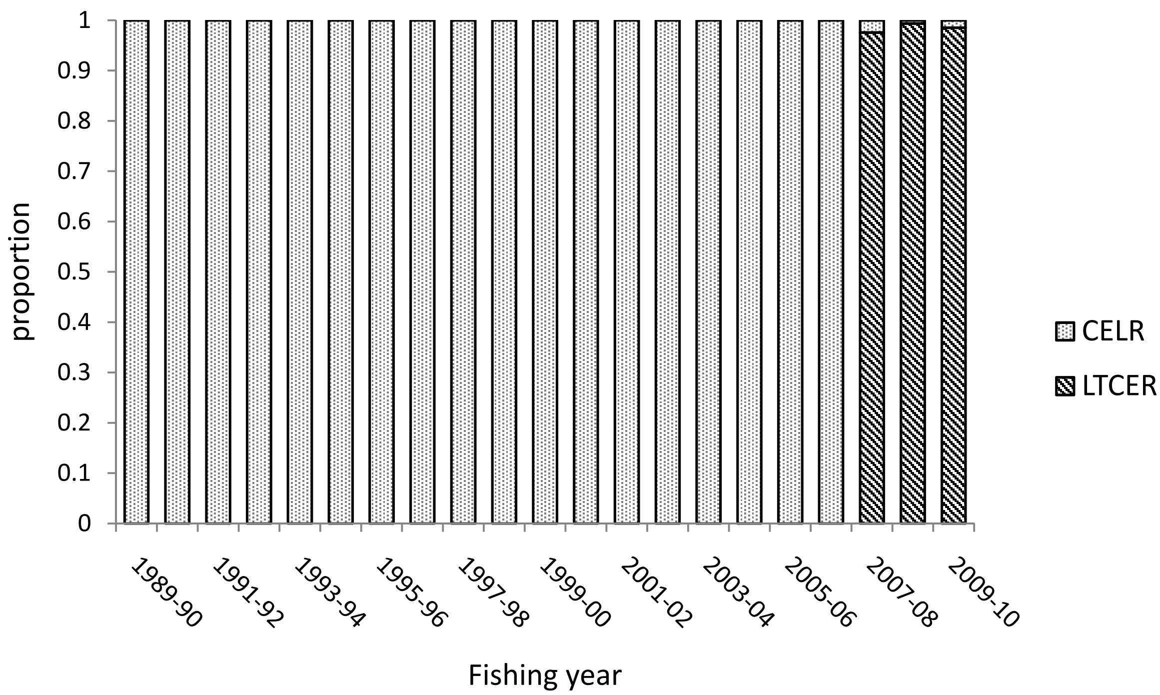


Figure 3: Proportion of trips reporting on CELR and LTCER forms in the SNA 1 longline fishery

The higher resolution of the LTCER reporting form provides greater power in the CPUE standardisations in that more spatial detail can be included in the analysis (e.g. depth). However, with only three years of higher resolution data available the Working Group deemed that a separate standardisation on these data would have limited utility at this time. The Working Group recommended that the final index should be based on combined CELR and LTCER data (i.e. one 21 year series); to do this required summing the LTCER records at the daily level to match the resolution of the CELR series. Combining data across form types had the potential to introduce bias in the analysis if effort reporting behaviour between the two form types was inconsistent. The validity of combining data from the two forms was investigated by plotting the various effort data and inspecting the annual trends for obvious discontinuities at the 2007–08 fishing year boundaries.

The longline catch and effort data were subdivided by sub-stock area to allow separate standardisations to be undertaken for each.

Data obtained from the Ministry were groomed, checked for typical reporting errors (Appendix 2), and compiled into two tables:

1. Landed catch weight: A file containing the verified green (unprocessed) landed weight of all SNA 1 trips.

2. Trip effort data: A file containing demographic information (location method, target species etc).

The Trip effort data table contained fisher catch estimates which were prorated against the actual trip landed weight totals across the effort information (i.e. individual days) on the basis of the estimated catch ratios. The link between the two data tables was a common trip number record.

2.1.3 Bay of Plenty Single Trawl CPUE data

Single trawl catch and effort data were obtained from the Ministry of Fisheries (Appendix 4). The criteria for the extract were to obtain all effort information for all single trawl trips landing catch from SNA 1 (data specific to the Bay of Plenty would then be extracted from this).

Unlike longline, the Bay of Plenty trawl fishery is a mixed species fishery; species that account for a significant proportion of the catch are: snapper (SNA); trevally (TRE), tarakihi (TAR), red gurnard (GUR), John Dory (JDO); and barracouta (BAR) (Appendix 9). Unlike the target longline fishery it was considered there was potential for the ratio of zero to positive single trawl snapper catches to hold an abundance signal. To capture zero catch effort potentially not included in the positive catch extract the Ministry was also asked to provide effort information for all single trawl tows executed in statistical areas 008 through 010 (Bay of Plenty), targeting any one of the following six species: SNA; TRE; TAR;GUR;BAR; JDO.

The Bay of Plenty single trawl CPUE data were available from fishing years 1989–90 to 2009–10 (a 21 year time series). However, three different catch effort form types have been in use during this period, partially limiting the temporal continuity of the series (Figure 4). Prior to the 1997–98 fishing year the majority of Bay of Plenty trawl fishers were using the less detailed daily CELR reporting forms. From 1995–96, however, a significant number of Bay of Plenty trawl fishers (more than 70%) were reporting on Trawl Catch Effort Processing Returns (TCEPR) that provide effort details as well as latitude and longitude information for each tow. From the 2007–08 fishing year many Bay of Plenty trawl fishers started to use the new Trawl Catch Effort Return (TCER) forms. The TCER forms are largely identical to the TCEPR forms but require catch details to be recorded of the top eight, rather than five, species by weight in the catch. It was decided not to include the CELR data in the CPUE standardisations and only to include years where a high proportion of TCEPR and TCER data were available; specifically the 1995–96 to 2009–10 fishing years (a 15 year time series).

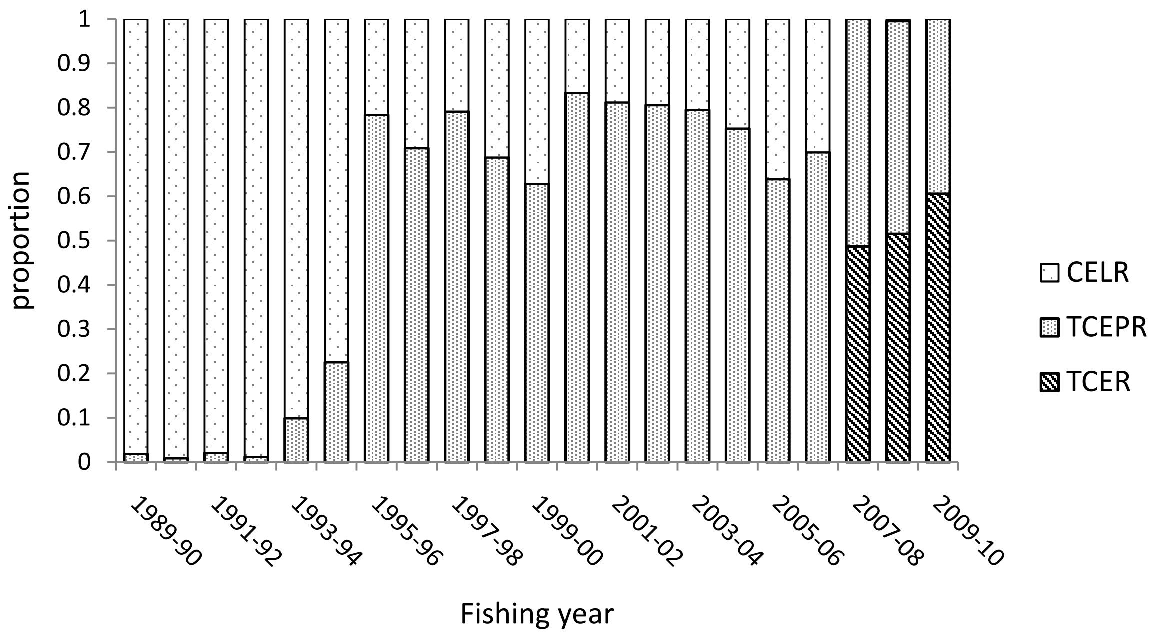


Figure 4: Proportion of trips reporting on CELR, TCEPR and TCER forms in the Bay of Plenty bottom trawl fishery.

Data obtained from the Ministry were groomed and checked for typical reporting errors (Appendix 2) and compiled into two tables:

1. Landed catch weight: A file containing the verified landed green weight (unprocessed) for all SNA 1 trips.
2. Trip effort data: A file containing demographic information (location method, target species etc).

The Trip effort data table contained fisher catch estimates which were prorated against the actual trip landed weight totals across the effort information (i.e. individual tows) on the basis of the estimated catch ratios. The link between the two data tables was a common trip number record.

2.2 CPUE standardisation methods

2.2.1 GLM regression models

Annual catch indices (assumed to represent snapper availability) were derived using generalised linear modelling (GLM) procedures (Vignaux (1994), Francis (1999)). The GLM's were conducted using the statistical software package R. The response variable in the GLM was log catch. Fishing-year was entered as a categorical covariate (explanatory) term on the right-hand side of the model. Standardised CPUE abundance indices (canonical) were derived from the exponential of the fishing-year covariate terms as described in Francis (1999).

In order to accommodate a non-linear relationship with the response variable (log catch) all continuous variables (including effort terms) were "offered" to the GLM's as third order polynomials. A forward fitting, stepwise, multiple-regression algorithm was used to fit GLM's to groomed catch, effort and characterisation data. The stepwise algorithm generates a final regression model iteratively and uses a simple model with a single predictor variable, fishing year, as the initial or base model. The reduction in residual deviance relative to the null deviance is calculated for each additional term added to the base model. The term that results in the greatest reduction in residual deviance is added to the base model if this results in an improvement in residual deviance of more than 1%. The algorithm repeats this process, updating the model, until no new terms can be added.

The approach taken with all the GLM's was to enter fishing 'effort' as a covariate (i.e., "right-hand" model term), with catch as the regressor variable. This is algebraically analogous to subtracting effort from catch in log-space.

The GLM standardisation of the zero and positive catch ratios (for the Bay of Plenty trawl analysis) was structured in a similar fashion to that described above, but used a binomial link function. The response variable in the binomial model was either "1" for a positive catch or "0" for a null catch. Indices of abundance derived from the lognormal and binomial models were also combined into a unified index using the method described by Vignaux (1994).

2.2.2 Model performance and diagnostics

The family of predictive models which attempt to "explain" variation in a response variable using a number of covariate terms or parameters can broadly be described as "regression-type" models. The family includes GLMs and analysis of variance (ANOVA) and multi-linear regression. There are a number of implicit assumptions which underlie the use of regression models which, if not met, may invalidate the use of the modelling approach or weaken the strength of its conclusions. In the case of a GLM standardisation of CPUE, serious violations to model assumptions may mean the model is inappropriate to "explain" the underlying variation in the response variable (e.g. catch or catch-per-unit-effort) and another modelling approach will need to be used (e.g. non parametric ranks). For most GLM CPUE standardisations it is more often the case, however, that violation of model assumptions (e.g. normality in the fitted residuals) are often not so severe as to warrant rejection of the model. However, in using GLM approaches for fisheries CPUE standardisation it is important to test for and describe significant violations of the modelling assumptions because the presence of such violations could mean that the GLM standardisation is "less than ideal" such that the level of variation indicated by the model is underestimated and that the trends and patterns (magnitude) in the year parameter (abundance signal) are not necessarily as strong or clear as the model indicates.

The validity of a particular GLM regression model for CPUE standardisation can be assessed specific to eight lack-of-bias criteria (Zuur et al 2009):

1. **Outliers** and errors in the raw response and covariate data values (raw observational data);
2. **Homogeneity** of variance in the response variable (CPUE) across the range of the covariate variables (e.g. between fishing years);
3. **Normality** in distribution of the response variable (CPUE)
4. Presence of a high number of **zero values** in the data (may require a different modelling approach, e.g. combined log-linear and binomial model);
5. **Colinearity** (correlation) amongst model covariate parameters;
6. **Non-linear relationships** (most GLM models assume this) between the response and continuous covariate model terms;
7. **Interactions** between fishing-year and other model terms (e.g. area or season) which may invalidate the standardisation;
8. **Independence** of the CPUE observations (lack of autocorrelation).

The diagnostic approach we have used for determining the "appropriateness" of both the raw base data and the validity of the modelling approach has been designed to specifically address Zuur's eight lack-of-bias criteria. The process is divided into steps:

1. Data grooming and data exploration;
2. Model selection and validation.

Note: Under our approach, some of Zuur's eight assumptions are examined more than once (e.g. the Normality assumption is tested in both steps).

Data grooming and data exploration

Data grooming addresses two types of error: missing values and extreme **outlier** values. In the case of missing response variables it is more often the case that the whole record has to be deleted. However, where parameter values (covariates) are missing or erroneous the more typical fix is to replace the parameter record with either the median value from the parameter distribution or with a value from an adjacent record. The parameter median and upper and lower ranges are derived from the parameter frequency distribution; typically values outside the 99% range of the data are inspected as potential outliers.

The assumption of **Normality** in the distribution of the response variable and **homogeneity** in the variance are assessed by inspecting the response variable frequency distributions; normality as an overall distribution and homogeneity as a series of comparative box plots at each category level (e.g. season).

Non-linear relationships between the response variable and continuous covariate terms and **collinearity** between the continuous covariate terms are evaluated using pairwise multi-panel plots and a crossed matrix of Pearson correlation coefficients (Pearson r). In most cases where two continuous covariate parameters are shown to be moderately correlated (0.6 – 0.9 Pearson r) they are usually both still "offered" in the model fitting process. However, where two parameters are found to be highly correlated (greater than 0.9 Pearson r) usually only the term that correlates most strongly to the raw response variable is "offered" to the model. The presence of a highly non-linear relationship between the response variable and a continuous parameter may mean Generalised Additive Modelling (GAM) is a more appropriate standardisation approach than GLM; however, our approach is to allow for non-linearity by fitting all continuous variables as third-order polynomials in a GLM (Maunder & Punt 2004).

Model selection and validation

A stepwise parameter selection process is used to determine the optimally parameterisation of the CPUE regression model. The process involves sequentially adding and removing covariate terms from the model such that the amount of variation "explained" by the model is maximised and the number of covariate terms used in the model is minimised.

For CPUE standardisations the abundance index is given by the coefficients on the fishing-year parameter, so for the final model fishing-year is usually forced into the model regardless of explanatory power. However, for the purpose of investigating fishing-year **interactions** a stepwise model selection process is first run where fishing-year is offered as a free parameter on its own and as a series of 2^{nd}-order interaction parameters with all other covariate parameter terms.

The validity of the **normality** and **homogeneity** assumptions are again investigated by looking at the distribution of model residuals as opposed to raw response observations. The residuals from the final model fit enable an evaluation of the homogeneity assumption in respect to the continuous variables when plotted across the range of each continuous variable; any trends or patterns in these plots being evidence of a lack of homogeneity.

Plots of the model Cook's distance scores (Cook & Weisberg 1982) are used to identify data observations with undue influence on the final model fits (**outliers**); as a rule of thumb Cook's values higher than 0.5 are cause for concern.

Residual plots are also used to look for evidence of a lack of temporal **independence** (autocorrelation) in the observational data; trends or patterns in the residuals plotted in temporal order is taken as evidence of a lack of independence.

A decision to investigate the effect of **zero value** observations in the CPUE standardisations is typically made a priori on the basis of the percentage of zero sets or observations in the data. The approach taken is to derive two standardised indices; one based on positive catches (log-normal GLM), the other on the ratio of zero to positive catches (binomial GLM). Vignaux (1994) provides a method to combine these indices (if deemed appropriate to do so); however there is no straightforward way for deriving variance on the combined index.

The "appropriateness" of both the raw data and the validity of the modelling approach was assessed by investigating the assumptions of: **homogeneity of variance** in the response variable across the covariate variable range; **normality** in the overall response variable distribution; lack of **colinearity** (correlation) amongst model covariate parameters; **linear relationship** between the response variable and continuous covariate model terms; no significant **interactions** between the fishing year model term and other covariate model terms; **independence** (lack of autocorrelation) of the observational data series (Zuur et al 2009).

The level of influence that each of the fitted covariate terms had on the fishing year indices were investigated graphically using the Influ R software tools of Bentley et al. (in press).

2.2.3 Variables offered to the SNA 1 longline CPUE standardisation models.

The response variable used in the longline catch effort standardisations was the log of the prorated snapper catch (kg) at the amalgamated daily reporting level (log[kg/day]). Two effort variables were offered to the model: log number-of-sets; log number-of-hooks. Categorical variables offered to the model were: fishing-year (forced in the final model); month; vessel; and statistical-area.

A categorical variable "target" was offered for the East Northland and Bay of Plenty standardisations, this being either "1" for target and "0" for non-target. For the Hauraki Gulf longline data set, nearly all sets were listed as targeting snapper (more than 99 %), so the variable "target" was not offered.

Three continuous environmental parameters were also offered to the model. These were:

1. The Trenberth index (M1): being the average monthly sea level pressure difference between Hobart and the Chatham Islands which represents the amount of southerly flow over New Zealand;
2. The Southern Oscillation Index (SOI): being the average monthly sea level pressure difference between Darwin and Tahiti;
3. Sea Surface Temperature (SST): index as recorded daily at the University of Auckland Leigh Marine research station.

2.2.4 Variables offered to the Bay of Plenty single trawl CPUE standardisation models.

The response variable used in the Bay of Plenty trawl catch effort log-linear positive catch standardisation was the log of the prorated snapper catch (kg) at the individual tow level (log[kg/tow]). For the binomial catch standardisation the response was simply "0" (no catch) or "1" (positive catch).

Six continuous effort variables were offered to both the log-normal and binomial models. These were: Log (fishing duration); Log (net height); Log (net width); Log (gear depth); Log (engine power); Log (vessel length*depth*breadth).

Categorical variables offered to both models were: fishing-year (forced in the final model); month; season (four levels), vessel; and statistical-area.

In the Bay of Plenty trawl fishery 98% of the snapper catch is taken targeting five main species: SNA, TRE, TAR, GUR and JDO (Appendix 9). Therefore, for both of the Bay of Plenty trawl models "target" was offered as a six level categorical variable (five target species plus an "other" category).

The Trenberth index (M1), Southern Oscillation Index (SOI), and Sea Surface Temperature (SST) were offered to both models as continuous variables.

2.2.5 Core vessel selection and Data grooming

A core fleet of vessels was selected for each of the sub-stocks using the method described by Kendrick & Bentley (2011). This method attempts to restrict analysis to vessels that participated in the fishery with consistency. The selection process required maximising the number fishing-years vessels spent in the fishery, maximising the number of vessel trips per year and minimising the number of vessels used in the analysis at a 60% retention level of the total series catch. The number of vessels that had been in the fishery for different numbers of years was plotted (see Appendix 13) as three separate series (e.g. vessels that undertook 5 trips per year, vessels with 20 trips per year, vessels with 40 trips per year) that encompassed the range in the number of trips per year within a particular fishery. All vessels were then sorted on the number of trips per year they conducted, after which vessels were selectively removed until 60% of the total catch was achieved.

Detailed grooming of the core fleet data was conducted separately for each of the SNA1 sub-stocks.
Distributional frequency plots were generated for the catch response variable and effort covariates to establish plausible boundary ranges for these variables. Flags were then set up in the database that highlighted values that were in the tails of the distributions for the main response and effort variables. Each data series was then assessed line by line, paying attention to where data were flagged.

3 RESULTS

3.1 Characterisation

3.1.1 Data grooming errors

The landed catch weights derived after removing non-terminating (see Appendix 2) catch records from the landed catch data table are given in Table 1. The proportion of retained or transhipped landing records varied around 1.2% of the annual reported SNA 1 catch for most years in the series, reaching a peak of 2.39% in 2005–06 (Table 1).

For most years the amount of landed catch that could be linked directly to (prorated across) effort was very high (98 to 99%; Table 1); the characterisations in this report of the SNA 1 fishery are therefore likely to be highly representative of annual SNA 1 catches ("Landed catch" Table 1).

Table 1: Breakdown of total SNA 1 reported landed catch (t) showing Total landings (contains retained i.e. duplicate catch data); Landed catch (retained [i.e. duplicate] data removed); Effort link (catch able to be linked to SNA 1 effort data).

Fishing year	Total landings	Retained landings	% Retained	Landed catch	Effort link	% effort link
2000–01	4 441	24	0.54%	4 417	4 362	98.75%
2001–02	4 398	30	0.68%	4 368	4 334	99.22%
2002–03	4 561	62	1.36%	4 499	4 410	98.02%
2003–04	4 580	56	1.22%	4 524	4 471	98.83%
2004–05	4 666	71	1.52%	4 595	4 554	99.11%
2005–06	4 680	112	2.39%	4 568	4 485	98.18%
2006–07	4 488	68	1.52%	4 420	4 371	98.89%
2007–08	4 626	47	1.02%	4 579	4 466	97.53%
2008–09	4 586	47	1.02%	4 539	4 466	98.39%
2009–10	4 567	66	1.45%	4 501	4 434	98.51%

3.1.2 Catch by sub-stock

A large proportion of the annual SNA 1 catch was taken from the Hauraki Gulf and Bay of Plenty sub-stocks, with a consistently lower portion from East Northland (Figure 5).

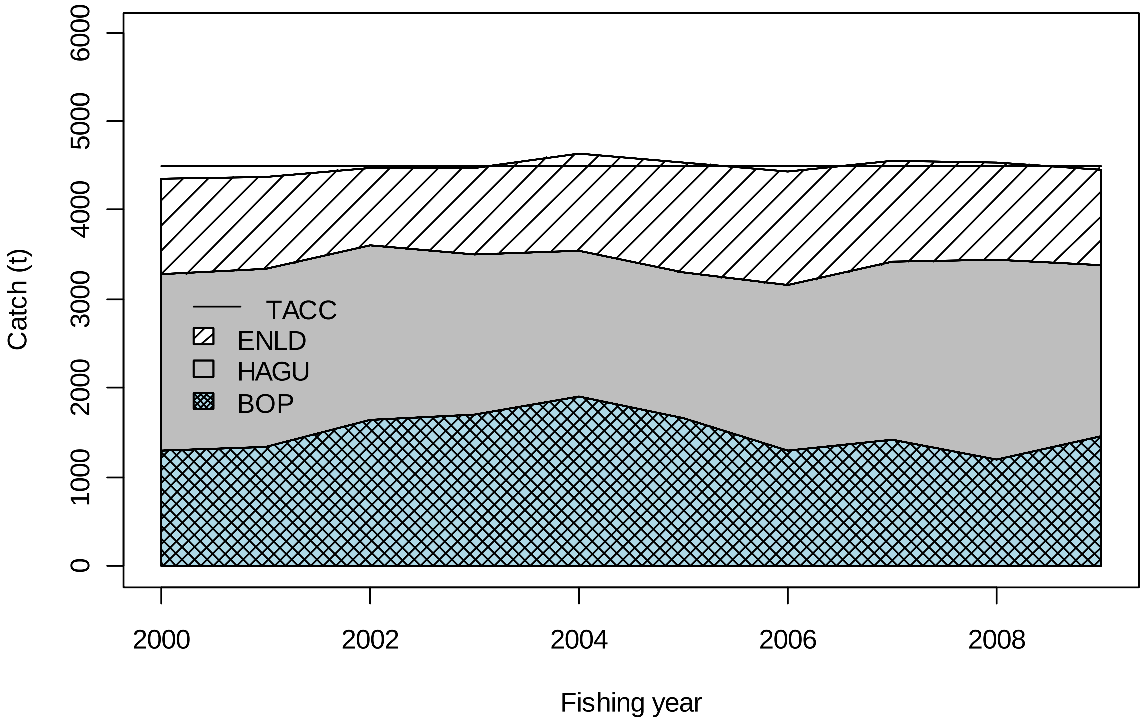


Figure 5: Annual SNA 1 catch by sub-stock (ENLD = East Northland; HAGU = Hauraki Gulf; BOP = Bay of Plenty).

3.1.3 Main methods

Very little of the East Northland SNA 1 catch was taken by methods other than longline and trawl (Figure 6; Appendix 5). Bottom long line was the dominant catching method in East Northland, with significantly more catch taken by this method than trawl over most fishing years (Figure 6; Appendix 5).

Longline, the dominant Hauraki Gulf snapper fishing method in the early 2000s, diminished in importance over the subsequent 10 years with proportionally increasing catches taken by single trawl and Danish seine (Figure 6; Appendix 5).

Bottom trawl was the dominant fishing method taking SNA1 in the Bay of Plenty, followed by Danish seine and longline (Figure 6; Appendix 5). There is little evidence for a temporal trend in the proportions of the Bay of Plenty SNA 1 catch taken by these three fishing methods (Figure 6; Appendix 5).

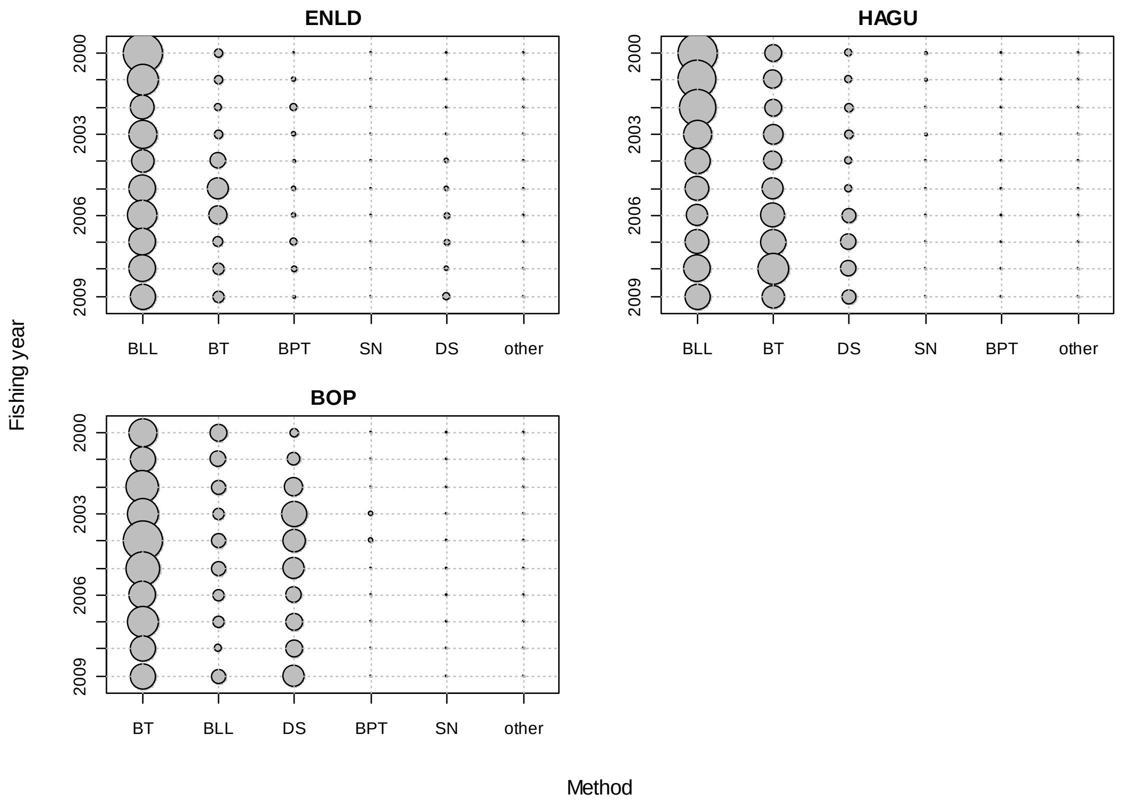


Figure 6: Relative annual SNA 1 catch by area and method (BLL = bottom longline; BPT = bottom pair trawl; BT = bottom trawl; DS = Danish seine; SN =setnet); circle area proportional to landed weight.

3.1.3.1 Bottom longline (BLL)

The spatial distribution of SNA 1 longline catches in the 2009–10 fishing year shows a strong emphasis on the central Hauraki Gulf, eastern Coromandel, Bay of Islands and Great Exhibition Bay, with only sparse activity beyond Tauranga in the eastern Bay of Plenty (Figure 7). These spatial patterns in longline fishing activity are likely to have been relatively consistent through time as indicated by the coarser statistical-area catch information (Figure 8; Appendix 6).

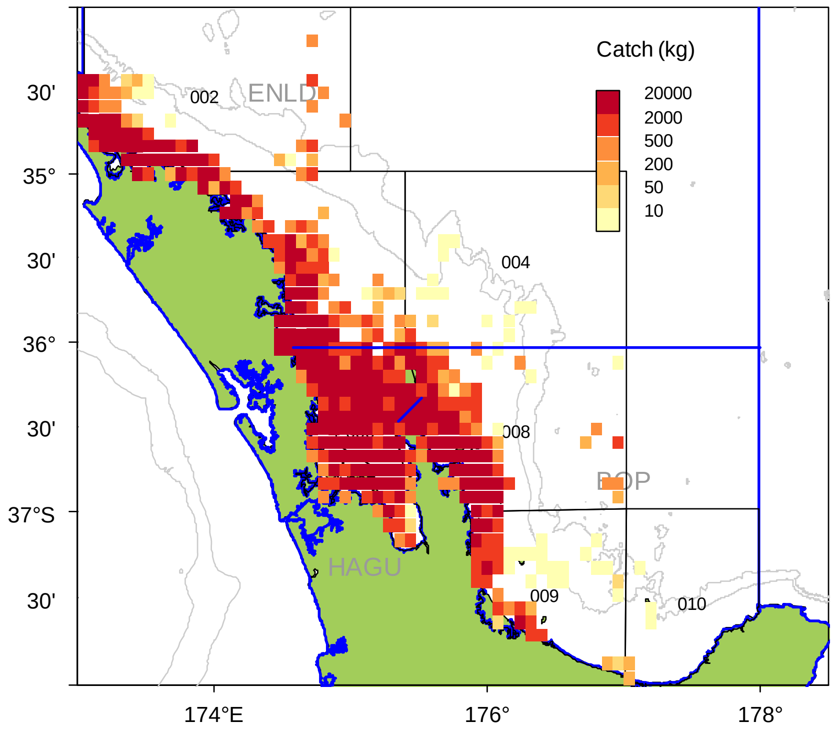


Figure 7: **Spatial distribution of the SNA 1 2009–10 fishing year longline catch**

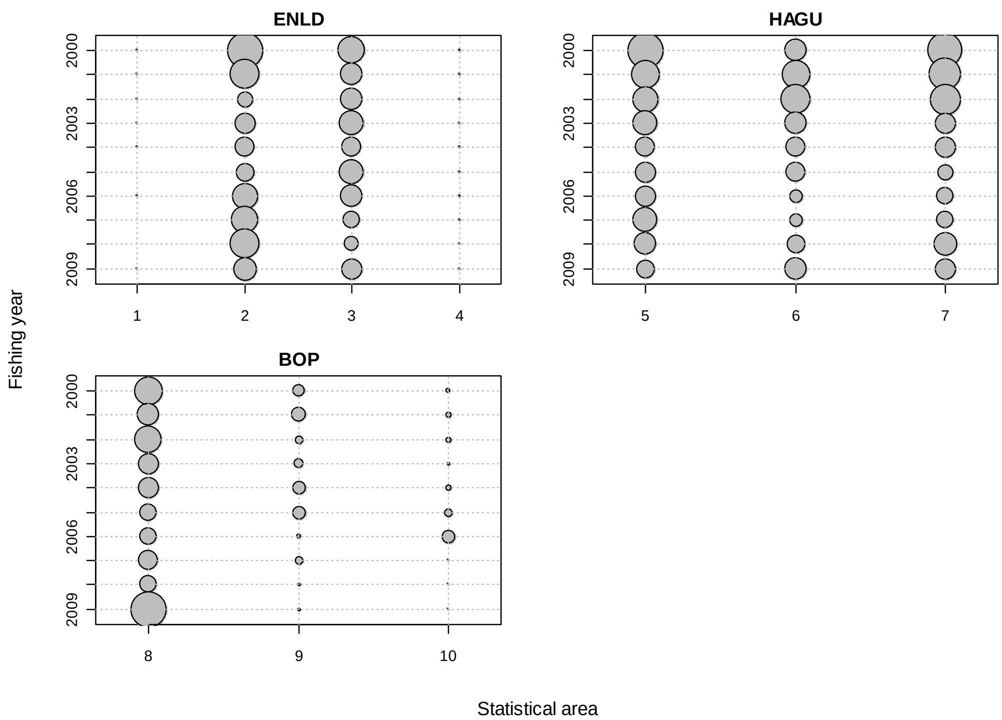


Figure 8: **Relative annual SNA 1 longline catch by statistical reporting area; circle area proportional to landed weight.**

The SNA 1 longline fisheries are predominantly target fisheries in all three sub-stocks; with a very small proportion of the SNA1 longline catch t taken when targeting red gurnard (Figure 9; Appendix 7).

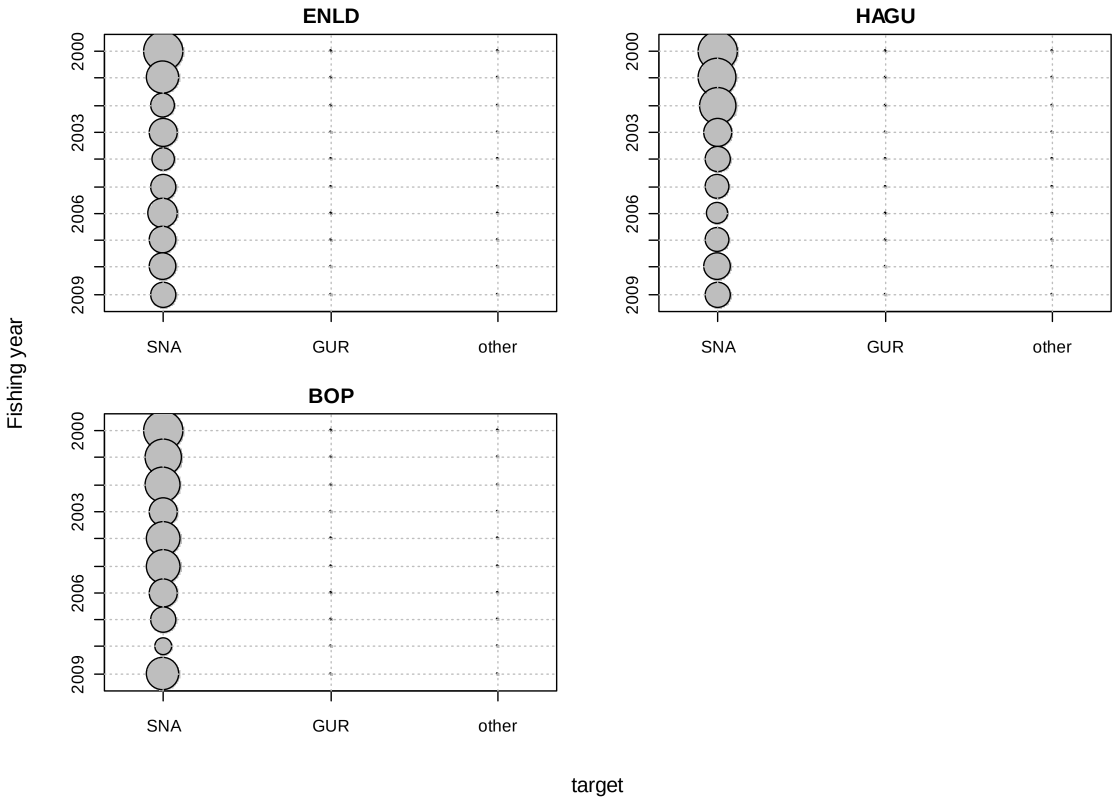


Figure 9: Relative annual SNA 1 longline catch by target species; circle area proportional to landed weight.

The longline fisheries in each SNA 1 sub-stock show different seasonal patterns (Figure 10). The fishery in East Northland appears to operate more consistently throughout the year than in the other two sub-stocks, although there is evidence of slightly higher catches in mid-summer and mid-winter (Figure 10). A pronounced seasonal trend in catches is evident for the Hauraki Gulf with larger catches taken in spring and summer (Figure 10). A less pronounced seasonal pattern in the Bay of Plenty shows that catches are slightly higher in late winter and early spring (Figure 10).

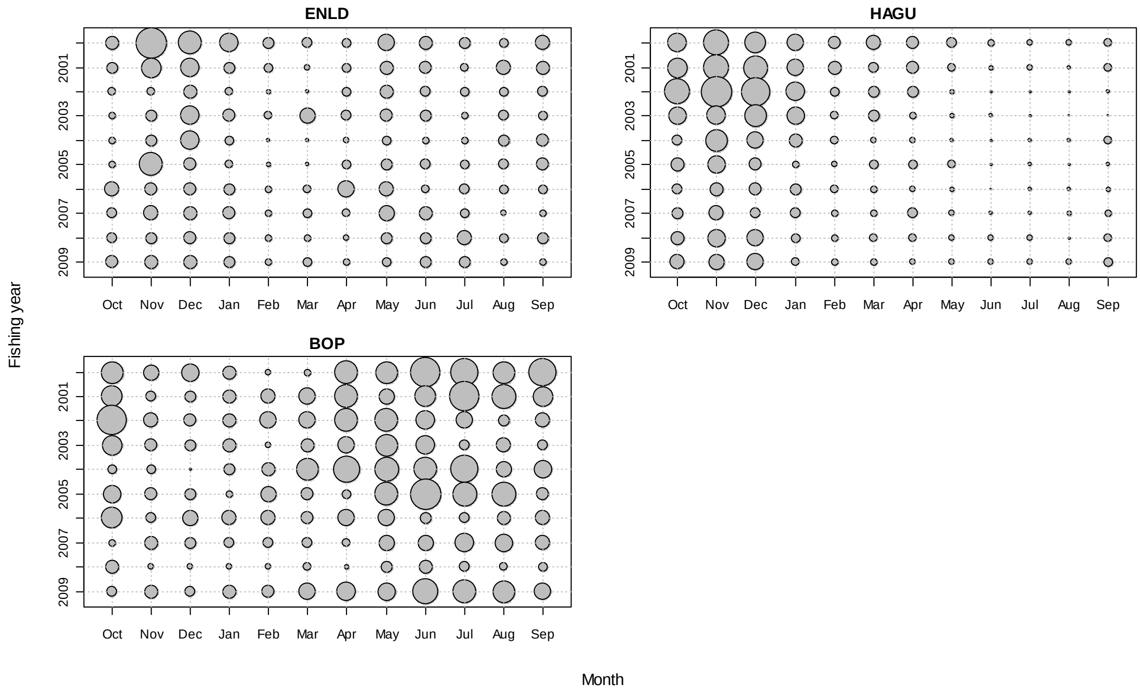


Figure 10: Relative annual SNA 1 longline catch by month; circle area proportional to landed weight.

3.1.3.2 Single Bottom Trawl (BT)

The spatial distribution of SNA 1 trawl catches in the 2009–10 fishing year shows a strong emphasis on the outer Hauraki Gulf, and central and eastern Bay of Plenty (Figure 11). These spatial patterns in trawl fishing activity are likely to have been relatively consistent through time as indicated by the coarser statistical-area catch information (Figure 12; Appendix 8).

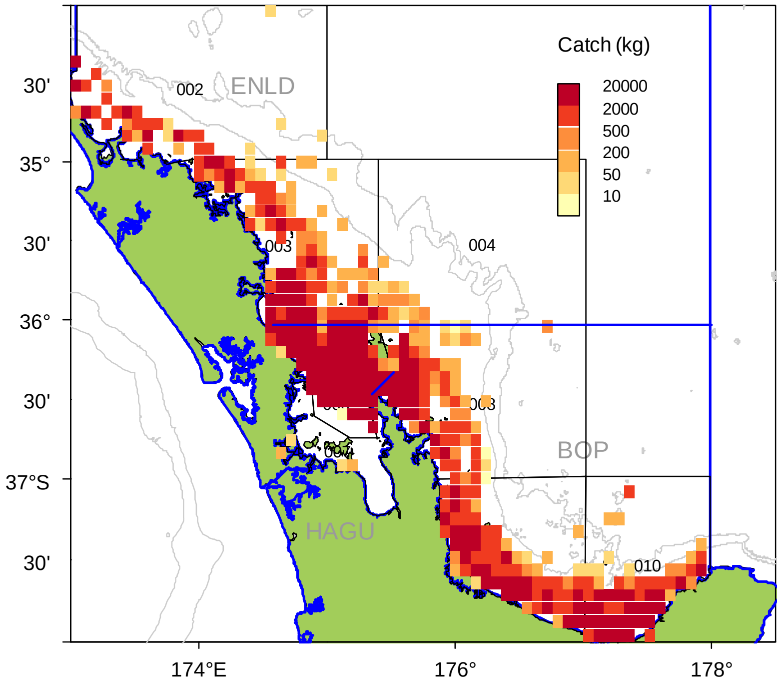


Figure 11: Spatial distribution of the SNA 1 2009–10 fishing year bottom trawl catch.

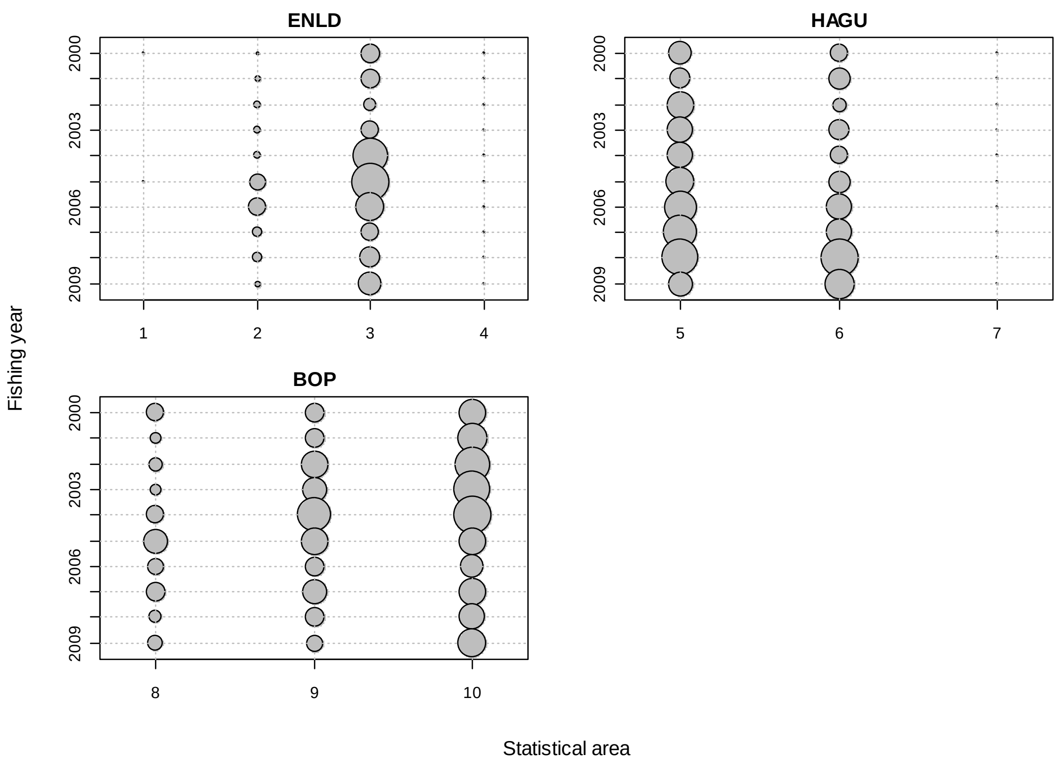


Figure 12: Relative annual SNA 1 bottom trawl catch by statistical reporting area; circle area proportional to landed weight.

The SNA 1 bottom trawl fishery is mainly a target fishery in all three sub-stocks (Figure 13; Appendix 9); however, proportionally more snapper is taken while targeting other species than is the case for longline. In East Northland other target trawl fisheries taking snapper are for trevally, tarakihi, and John dory (Figure 13; Appendix 9). In the Hauraki Gulf the only other target fishery of importance to snapper is John dory (Figure 13; Appendix 9). In the Bay of Plenty trawl fishery snapper is taken whilst targeting trevally, tarakihi, red gurnard, and John dory (Figure 13; Appendix 9).

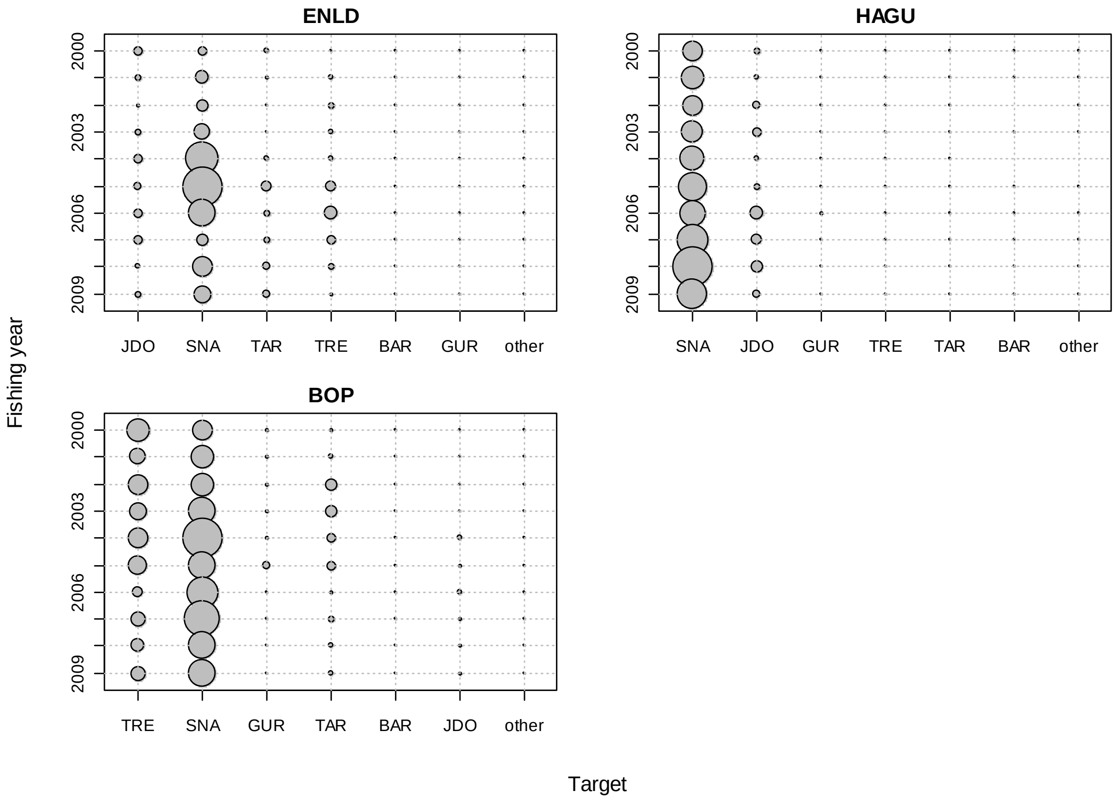


Figure 13: Relative annual SNA 1 bottom trawl catch by target species; circle area proportional to landed weight.

There is evidence of summer, autumn and spring peaks in the East Northland single trawl snapper catch series (Figure 14), although this is somewhat masked by a high degree of variation between fishing years. The Hauraki Gulf trawl fishery has moved from a predominantly summer fishery to one fishing relatively consistently throughout the year (Figure 14). Likewise the Bay of Plenty trawl fishery shows a more even spread of catch throughout the year in the latter years of the series (Figure 14).

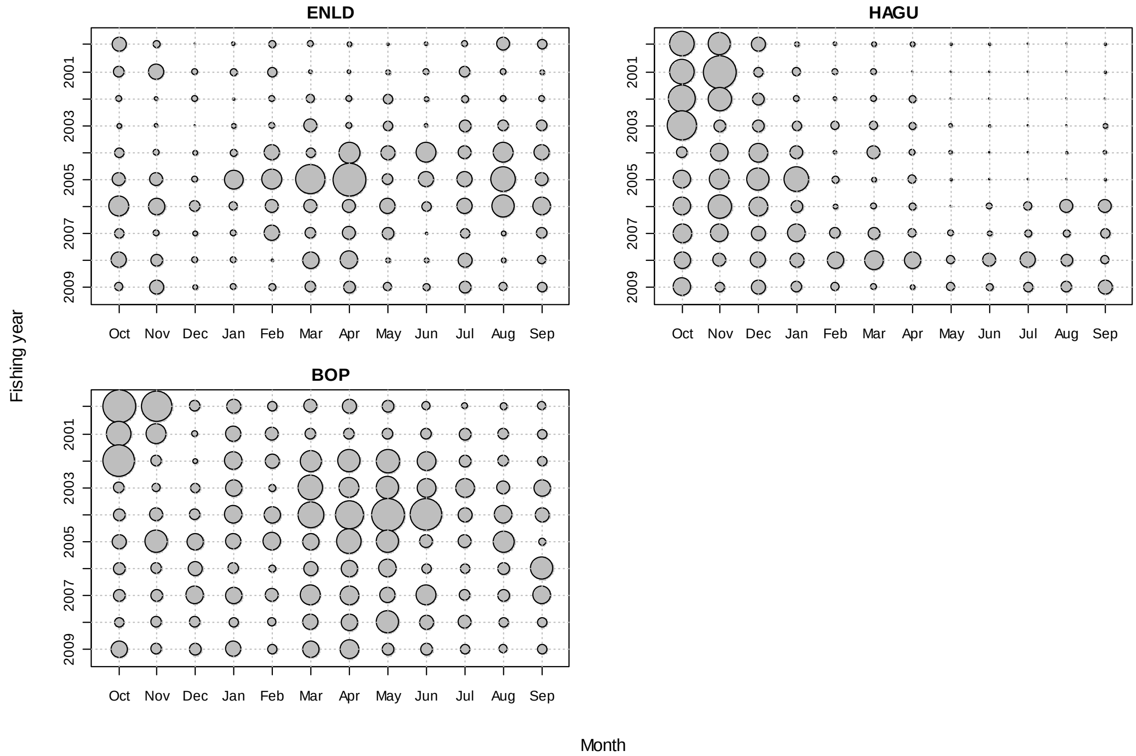


Figure 14: Relative annual SNA 1 bottom trawl catch by month; circle area proportional to landed weight.

3.1.3.3 Danish Seine (DS)

Because Danish seine vessels report predominantly on the less detailed CELR reporting forms, detailed spatial position data is largely unavailable for this method. At the broader statistical reporting area level, Danish seine predominantly operates in statistical-area 003 in East Northland; statistical-area 006 in the Hauraki Gulf, and statistical-area 009 in the Bay of Plenty (Figure 15; Appendix 10).

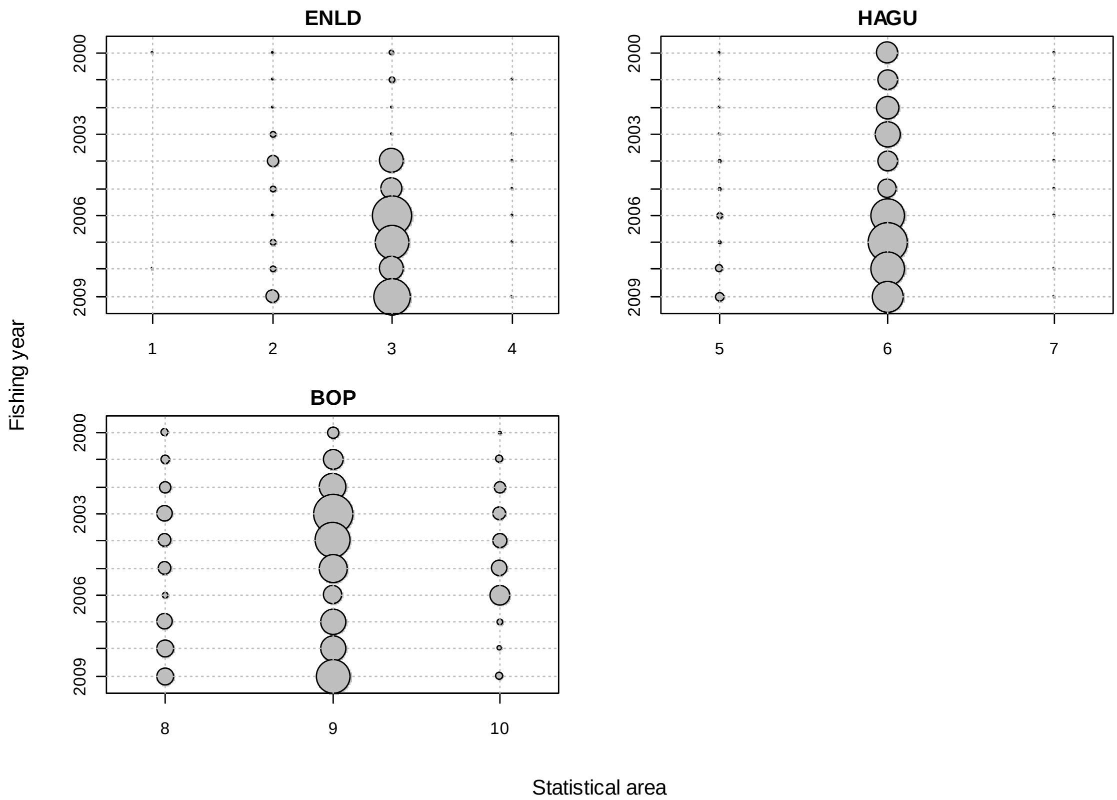


Figure 15: Relative annual SNA 1 Danish seine catch by statistical reporting area; circle area proportional to landed weight.

The majority of the SNA 1 Danish seine catch is taken while directly targeting snapper (Figure 16; Appendix 11). In East Northland snapper is also taken by Danish seine when targeting red gurnard and in 2009–10 tarakihi (Figure 16; Appendix 11). In the Hauraki Gulf the only Danish seine target fishery of importance for snapper is for John dory (Figure 16; Appendix 11).The Bay of Plenty Danish seine fishery is similar to East Northland in that red gurnard is the main target fishery for snapper by-catch (Figure 16; Appendix 11).

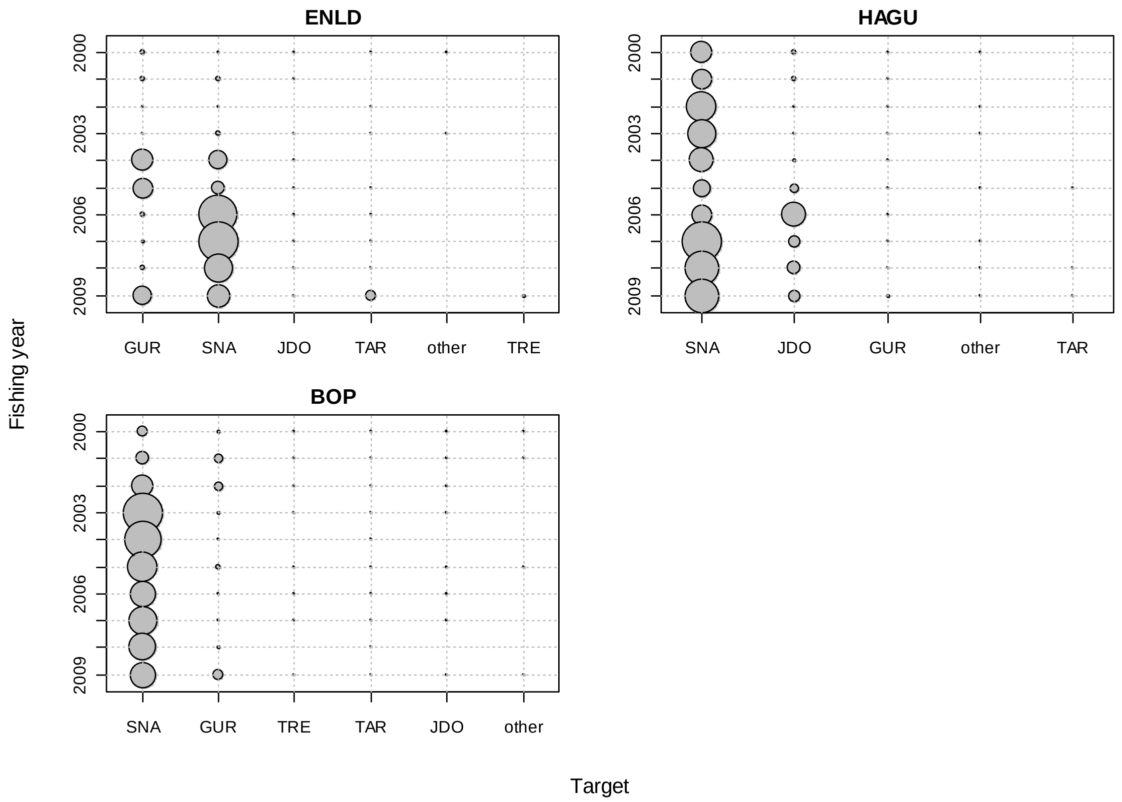


Figure 16: Relative annual SNA 1 Danish seine catch by target species; circle area proportional to landed weight.

Danish seine accounted for only a small amount of the snapper catch in East Northland prior to 2004–05 and after this date the method operated in a haphazard fashion throughout the year (Figure 17). Similar to the pattern seen in the Hauraki Gulf trawl fishery, Danish seine snapper catch in the Hauraki Gulf has moved from a predominantly summer fishery to one that fishes throughout the year (Figure 17). The operation of Danish seine in the Bay of Plenty is similar to East Northland, without a consistent seasonal pattern (Figure 17).

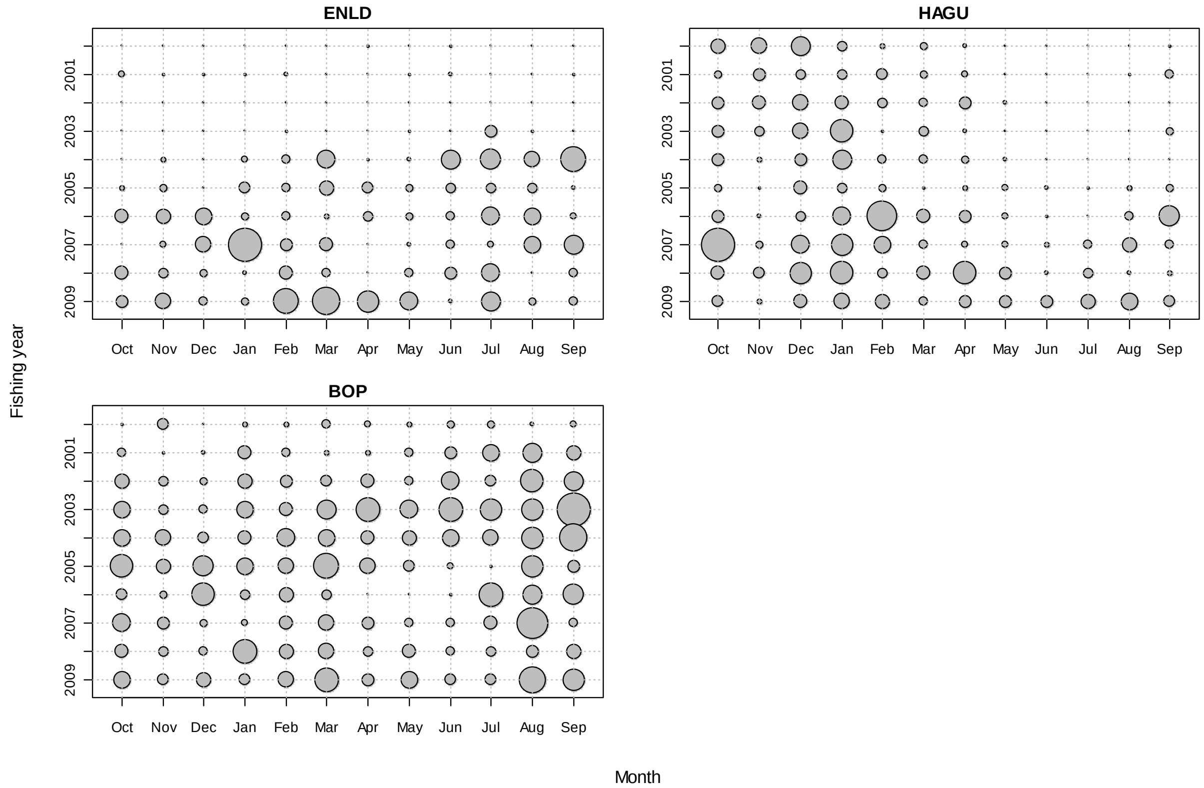


Figure 17: Relative annual SNA 1 Danish seine catch by month; circle area proportional to landed weight.

3.2 Standardised CPUE analysis

3.2.1 SNA 1 Longline

3.2.1.1 Data grooming and vessel selection

Frequency distributions of catch and the number of hooks and number of sets were used to determine plausible values for data grooming. Warning flags were set when catch was greater than 3000 kg, number of hooks less than 25 or greater than 6000 or the number of sets per day was greater than 10 (Appendix 12).

A common error in the data sets was the transposition of entries for the number of hooks set and the number of sets per day. This was often easy to identify and correct. Other common error types included missing values and typing mistakes, which in some cases were obvious and corrected to match the previous pattern for that vessel. For example, some fishers would conduct the same number of sets per day, so a missing entry or a very high value was replaced with the usual entry value for that fisher. Where entries for the catch of snapper or the number of hooks set were in doubt that entire record was deleted.

East Northland

Thirty four vessels were selected for analysis from the East Northland sub-stock (Appendix 13). The final data set consisted of 39 502 lines of data accounting for 7705 t of snapper catch. The coverage of data across the time series was reasonable, with overlapping data from multiple vessels present for all years (Appendix 14). There was no evidence of a shift in effort reporting corresponding to the 2007–08 form type change for the final selected vessels (Appendix 15).

Hauraki Gulf

Forty seven vessels were selected for analysis from the Hauraki Gulf sub-stock (Appendix 13). The final data set consisted of 60 239 lines of data accounting for 13 058 t of snapper catch. The coverage of data across the time series was reasonable, with overlapping data from multiple vessels present for all years (Appendix 14). There was no evidence of a shift in effort reporting corresponding to the 2007–08 form type change for the final selected vessels (Appendix 15).

Bay of Plenty

Thirty nine vessels were selected for analysis from the Bay of Plenty sub-stock (Appendix 13). The final data set consisted of 23 957 lines of data accounting for 4042 t of snapper catch. The coverage of data across the time series was reasonable, with overlapping data from multiple vessels present for all years (Appendix 14). There was no evidence of a shift in effort reporting corresponding to the 2007–08 form type change for the final selected vessels (Appendix 15).

3.2.1.2 East Northland longline

Data exploration

The raw log-transformed catch (per set) data did not show any major departures from that of a normal distribution (Appendix 16). The spread of catch data, across the different levels of the categorical explanatory variables, appeared to be even (Appendix 16), therefore conforming to the assumption of homogeneous variances. There were no strong correlations between the individual covariate data terms (i.e. limited collinearity; Appendix 16). There was a reasonably strong linear relationship evident between the response variable (log catch) and the number of hooks covariate (r = 0.5; Appendix 16).

Model selection allowing interactions with the fishing-year

The full stepwise regression analysis of the East Northland longline catches (i.e. a model that did not force fishing year and allowed interactions) explained 49 % of variation in the data using seven model terms (Table 2). Fishing year was selected as a significant term in this unforced model, explaining 1.7% of the variation in the data. The addition of a fishing year/vessel interaction term explained an additional 8.2% of variation (Table 2). The presence of a reasonably large explanatory interaction between these two terms indicates that abundance signals for different vessels are different; providing grounds for caution when interpreting the final model index.

Table 2: Stepwise regression results for the full model analysis (i.e. fishing year not forced and interactions allowed) of East Northland longline catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	6	7	n.s.
Num hooks*	0.249							
Vessel*	0.234	0.311						
Month*	0.036	0.290	0.353					
Target*	0.017	0.272	0.333	0.372				
Fishing year*	0.039	0.274	0.327	0.368	0.389			
Vessel: Fishing year*						0.471		
Month: Fishing year*						0.412	**0.493**	
Num sets	0.033	0.252	0.315	0.358	0.377	0.394	0.476	0.499

Model selection and CPUE standardisation

For the CPUE stepwise regression analysis fishing year was forced and interactions were not allowed. The final model explained 39 % of variation in the catch data using five variables including fishing-year (Table 3). The variable that explained the most variation was the number of hooks set (24 %), with the other four variables explaining 10% between them.

Table 3: Stepwise regression results for East Northland longline catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	n.s.
Fishing year*	0.039					
No. of hooks*	-	0.274				
Vessel*	-	0.245	0.327			
Month*	-	0.073	0.313	0.368		
Target*	-	0.060	0.300	0.350	**0.389**	
Statistical area	-	0.055	0.279	0.335	0.374	0.395

The residuals in the final model were normally distributed across plus or minus 2 standard deviations (95%) of the range of the data (Figure 18). The plot of Cook's distance scores shows no observations that had undue influence on the model fit (Figure 18). There was no evidence for failure in the homogeneity of variance assumptions as seen in the residual plots for each level of the categorical variable and across the range of the continuous variables (Appendix 17). Finally, there was no trend (i.e. significant divergence from the zero line) in the model residuals evident in the temporal order residual plot to suggest that there may have been autocorrelation in the data, i.e. a violation of the independence assumption (Appendix 18). Similarly, there was no evidence for a lack of homogeneity in the residuals plotted in temporal order (Appendix 18).

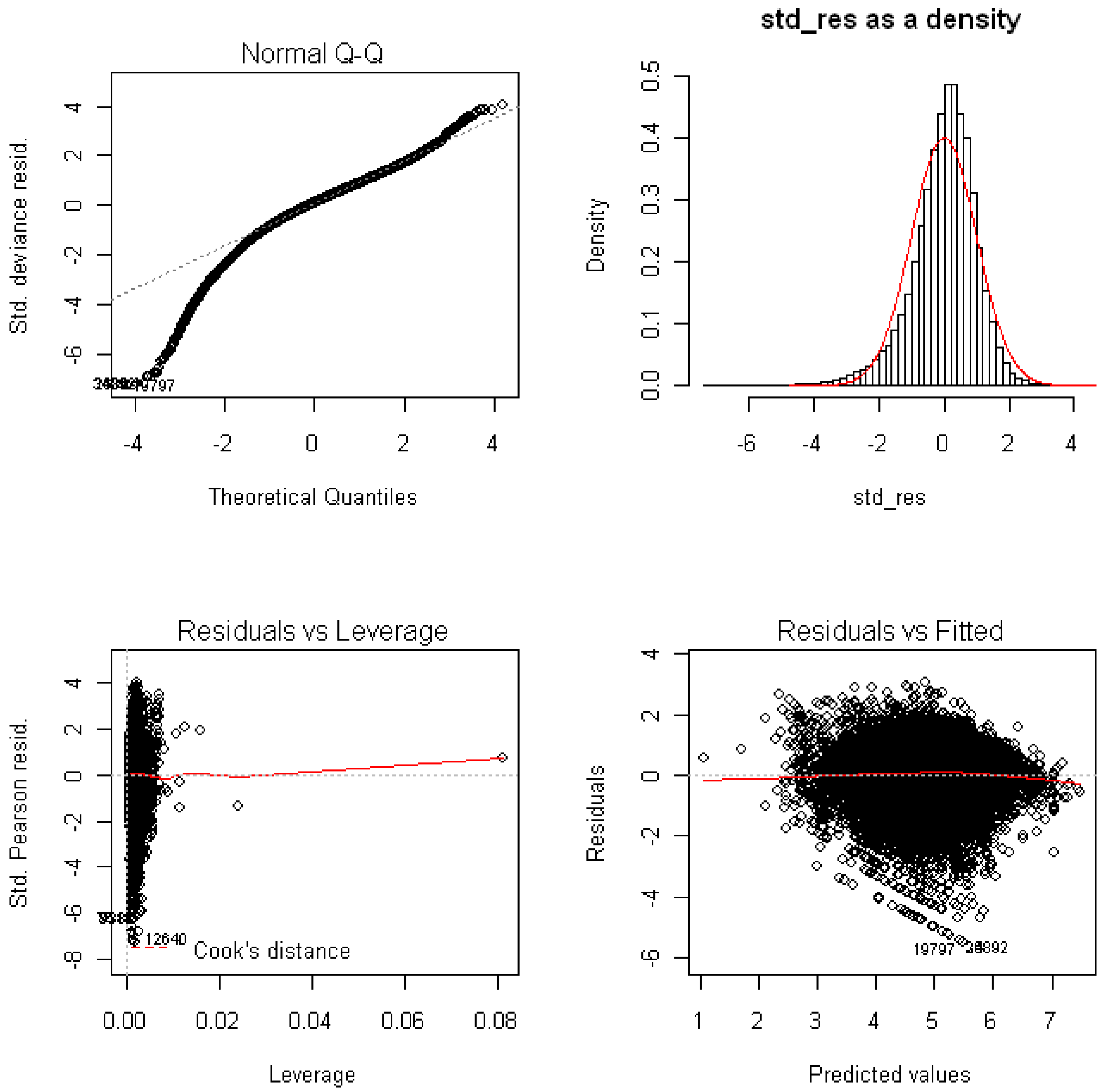


Figure 18: Analysis of regression model fit for East Northland longline catches. Top left panel is a Q-Q plot comparing the quantiles of a theoretical normal distribution to that of the residuals from the regression analysis. Top right panel is a histogram of the residuals from the regression analysis. Bottom left panel is a Cook's distance plot investigating the leverage of individual data points. Bottom right panel is a plot of model residuals against the fitted values of the model itself.

The variables most influential on the CPUE index were the number of hooks set and vessel; the addition of the other model terms had little impact on the pattern in the year index (Figure 19). The magnitude of influence the covariate parameters had on the year index is plotted in Appendix 19; the influence of number of hooks appeared to be exponential.

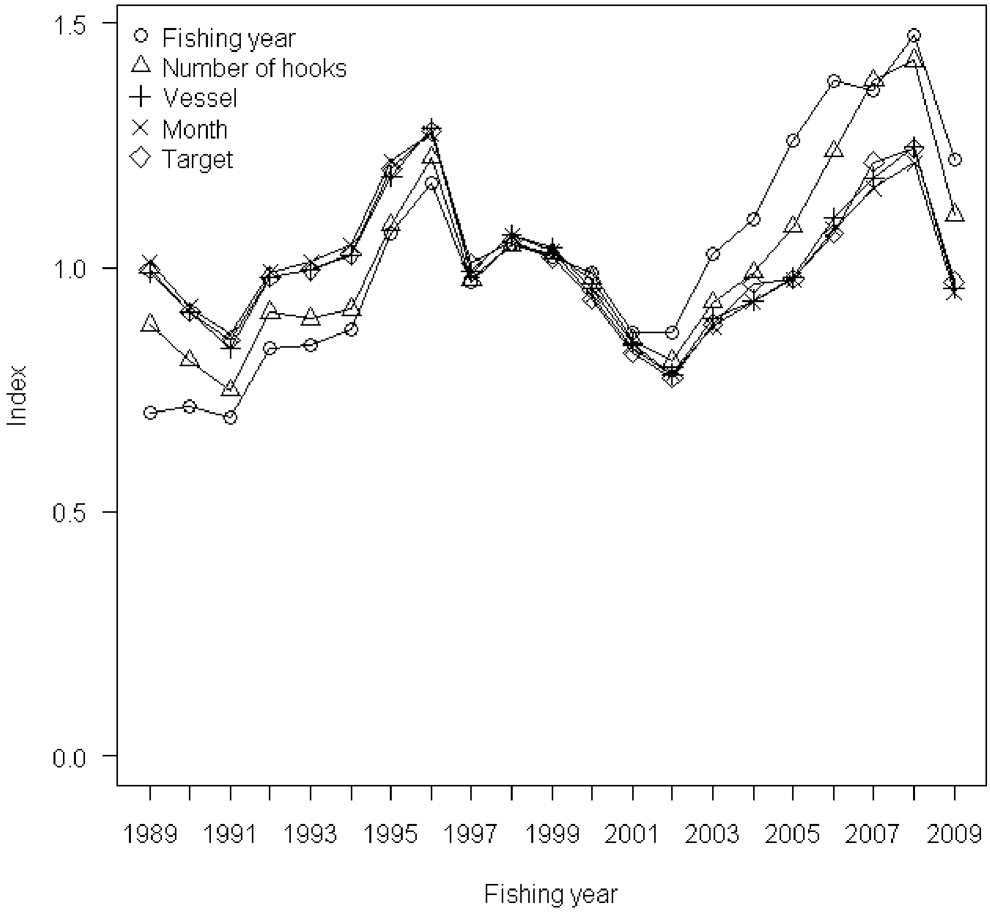


Figure 19: Step plot of the canonical index for East Northland longline catches. The model starts with just Fishing year and sequentially adds additional variables to the model (symbols lower on the legend).

The final East Northland index fluctuated without trend over the 21 year time series, with the highest CPUE occurring in the mid 1990's and around 2008 (Figure 20; Appendix 20).The final year of the index, however, saw a reduction in CPUE to a level close to that at the beginning of the time series in 1989–90. Variation around the index estimates was small (Figure 20; Appendix 20).

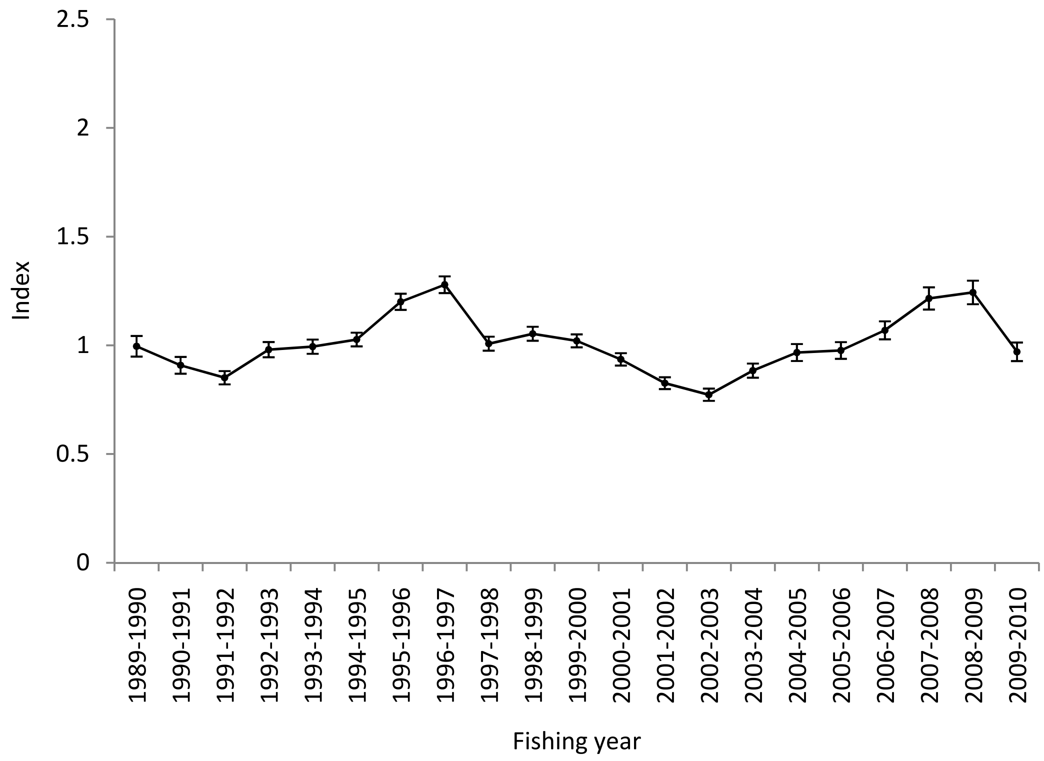


Figure 20: Canonical year indices for East Northland longline catches, error bars are 95 % confidence intervals.

3.2.1.3 Hauraki Gulf longline

Data exploration

The raw log-transformed catch data did not show any major departures from that of a normal distribution (Appendix 16). The spread of catch data, across the different levels of the categorical explanatory variables, appeared to be even (Appendix 16), therefore conforming to the assumption of homogeneous variances. There were no strong correlations between the individual covariate data terms (i.e. limited collinearity; Appendix 16). There was a reasonably strong linear relationship evident between the response variable (log catch) and the number of hooks covariate ($r = 0.5$; Appendix 16).

Model selection and CPUE standardisation

The full stepwise regression analysis of the Hauraki Gulf longline catches (i.e. a model that did not force fishing year and allowed interactions) explained 56 % of variation in the data using six model terms (Table 4). Fishing year explained 3.1 % of variation in the data. The addition of a fishing year/vessel interaction term explained an additional 4.8% of variation (Table 4). The presence of a reasonably large explanatory interaction between these two terms means that abundance signals at the vessel level are different; grounds for caution in the interpretation of the final model index.

Table 4: Stepwise regression results for the full model analysis (i.e. fishing year not forced and interactions allowed) of Hauraki Gulf longline catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	6	n.s.
Vessel*	0.292						
Num hooks*	0.028	0.410					
Month*	0.007	0.356	0.464				
Fishing year*	0.010	0.328	0.444	0.495			
Vessel:fishing year*	-	-	-	-	0.543		
Month:fishing year*	-	-	-	-	0.512	**0.558**	
Num sets	0.003	0.327	0.418	0.475	0.505	0.553	0.568

Stepwise regression analysis of the Hauraki Gulf longline catches when fishing year was forced and interactions were not allowed produced an index that explained 51 % of variation in the catch data using five variables (Table 5).The variables that explained the most variation were the number of hooks set (25 %), the identification number of each vessel (10 %) and Fishing year (10 %). The two other variables explained 5 % between them.

Table 5: Stepwise regression results for Hauraki Gulf longline catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	n.s.
Fishing year*	0.099					
No. of hooks*	-	0.346				
Vessel*	-	0.328	0.444			
Month*	-	0.166	0.399	0.495		
No. of sets*	-	0.132	0.350	0.451	**0.505**	
SST	-	0.131	0.376	0.471	0.500	0.511

The residuals in the final model were normally distributed across plus or minus two standard deviations (95%) of the range of the data (Figure 21). The plot of Cook's distance scores shows that no observations had undue influence on the model fit (Figure 21). There was no evidence for failure in the homogeneity of variance assumption as seen in the residual plots for each level of the categorical variable and across the range of the continuous variables (Appendix 17).

Finally, there was no trend (i.e. significant divergence from the zero line) in the model residuals evident in the temporal order residual plot to suggest that there may have been autocorrelation in the data, i.e. a violation of the independence assumption (Appendix 18). However pattern is seen in the residual spread when plotted in temporal order consistent with a failure in the homogeneity of variance assumption (Appendix 18). The consequences of this violation on the CPUE standardisation are unclear, but the magnitude of the index and the estimated variance are unlikely to be correct.

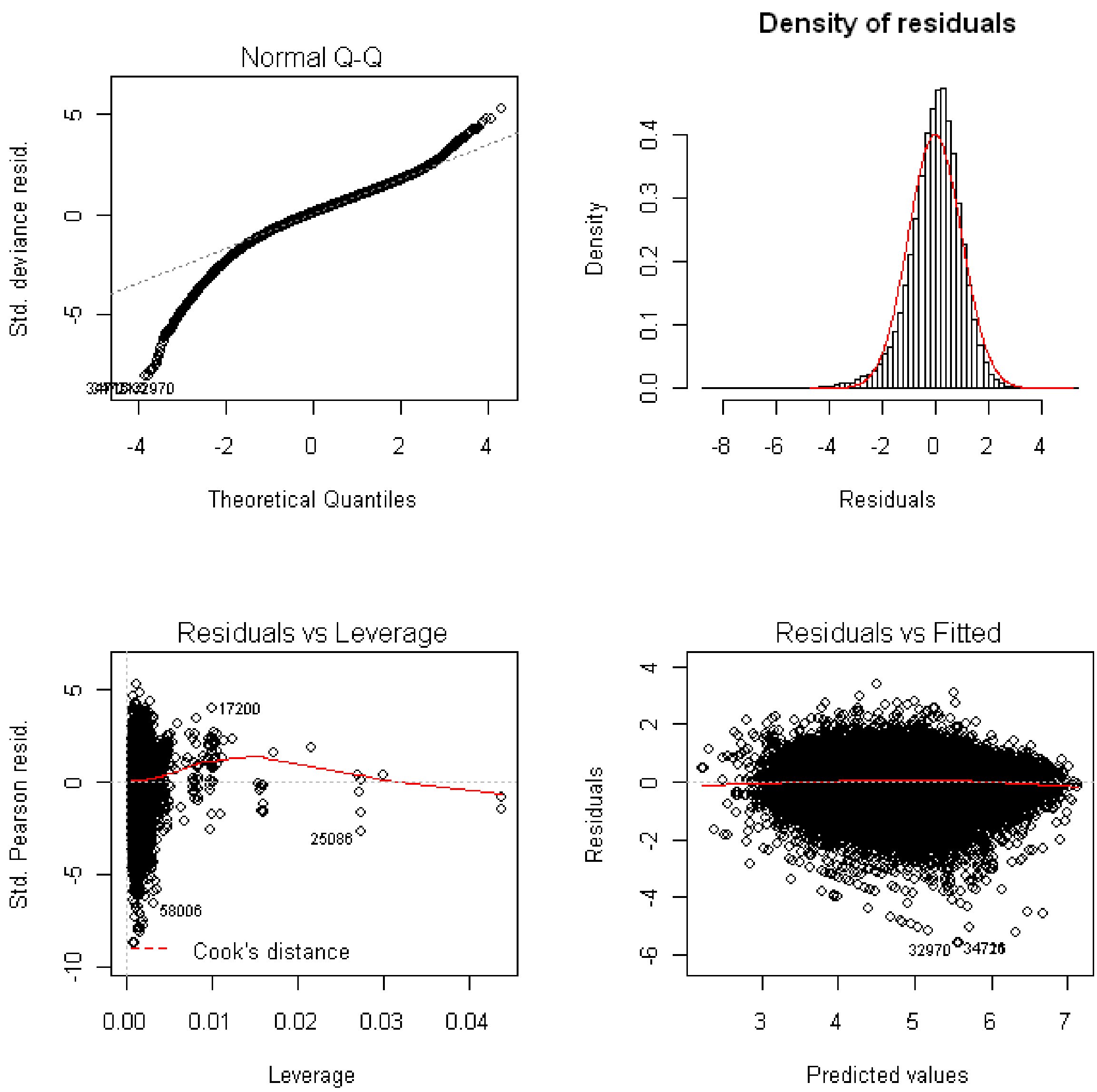


Figure 21: Analysis of regression model fit for Hauraki Gulf longline catches. Top left panel is a Q-Q plot comparing the quantiles of a theoretical normal distribution to that of the residuals from the regression analysis. Top right panel is a histogram of the residuals from the regression analysis. Bottom left panel is a Cook's distance plot investigating the leverage of individual data points. Bottom right panel is a plot of model residuals against the fitted values of the model itself.

The most influential covariate on the CPUE index was the number of hooks; the addition of the other model terms had little impact on the pattern in the year index (Figure 22). The magnitude of influence the covariate parameters had on the year index is plotted in Appendix 19; the influence of number of hooks again appeared to be exponential.

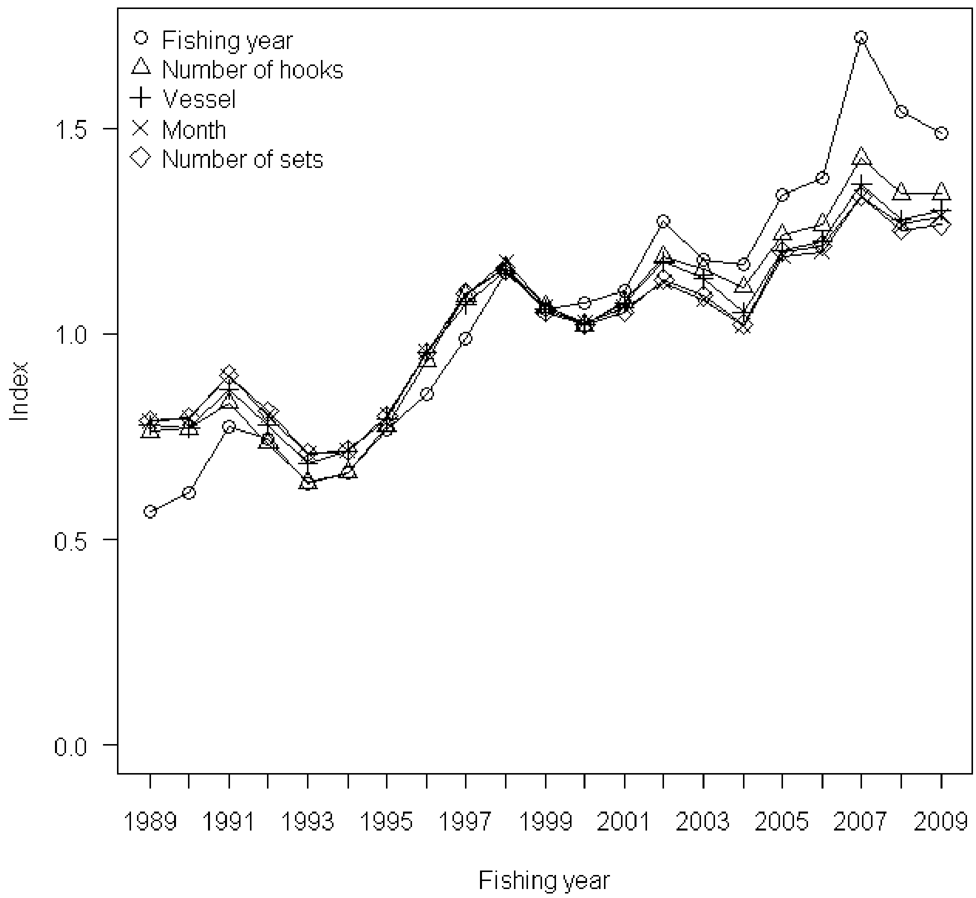


Figure 22: **Step plot of the canonical index for Hauraki Gulf longline catches. The model starts with just Fishing year and sequentially adds additional variables to the model (symbols lower on the legend).**

The final Hauraki Gulf index shows a general increase over the 21 year time series, with the CPUE at the end of the time series being 160 % of that at the beginning of the time series in 1989 (Figure 23; Appendix 20). Variation around the index estimates was small (Figure 23; Appendix 20), however evidence of a violation in the homogeneity assumption means the true precision is unlikely to be as high.

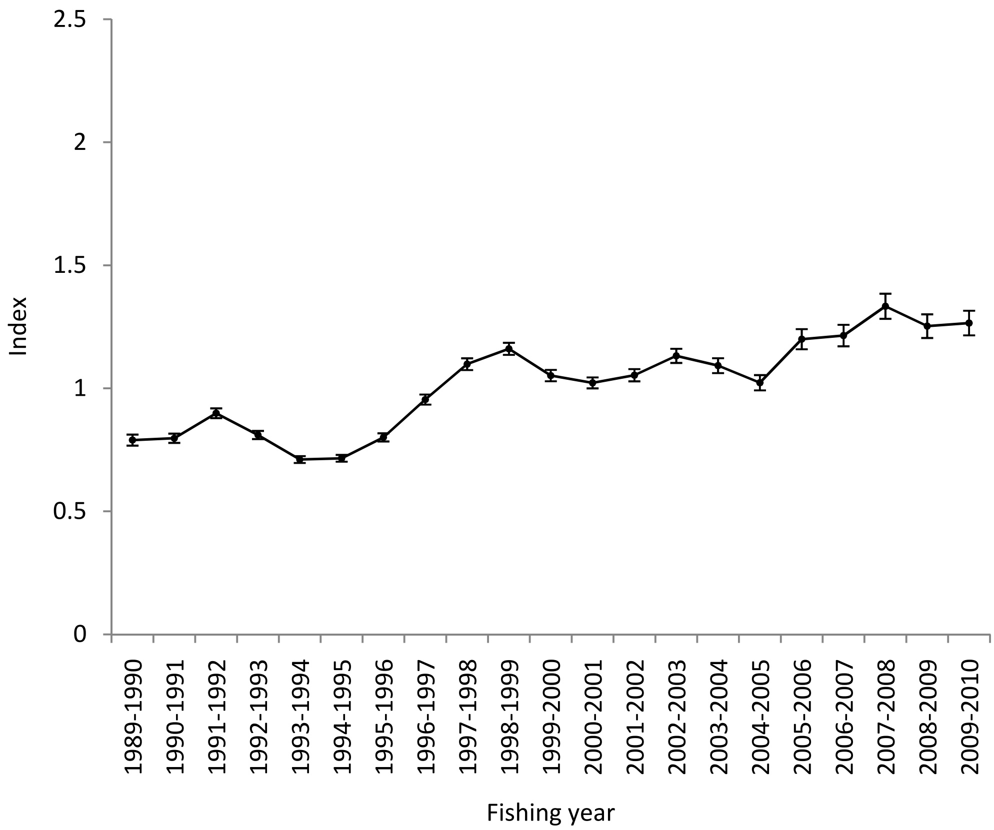


Figure 23: Canonical year indices for Hauraki Gulf longline catches, error bars are 95% confidence intervals.

3.2.1.4 Bay of Plenty longline

Data exploration

The raw log-transformed catch data did not show any major departures from that of a normal distribution (Appendix 16). The spread of catch data, across the different levels of the categorical explanatory variables, appeared to be even (Appendix 16), therefore conforming to the assumption of homogeneous variances. There were no strong correlations between the individual covariate data terms (i.e. limited collinearity; Appendix 16). There was a reasonably strong linear relationship evident between the response variable (log catch) and the number of hooks covariate ($r = 0.6$; Appendix 16).

Model selection and CPUE standardisation

The full stepwise regression analysis of the Bay of Plenty longline catches (i.e. a model that did not force fishing year and allowed interactions) explained 59 % of variation in the data using five model terms (Table 6). Fishing year explained 1.7 % of variation in the data. The addition of a fishing year/vessel interaction term explained an additional 3.7% of variation (Table 6). The explanatory power of the vessel interaction term is not as high as seen in the East Northland and Hauraki Gulf models and is of less concern.

Table 6: Stepwise regression results for the full model analysis (i.e. fishing year not forced and interactions allowed) of Bay of Plenty longline catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	n.s.
Vessel*	0.377					
Num hooks*	0.307	0.472				
Target*	0.105	0.429	0.535			
Fishing year*	0.066	0.405	0.504	0.552		
Vessel: Fishing year*	-	-	-	-	**0.589**	
SST	0.013	0.389	0.483	0.544		0.595

Stepwise regression analysis of the Bay of Plenty longline catches when fishing year was forced and interactions were not allowed produced an index that explained 55 % of variation in the catch data using four variables (Table 7). The variables that explained the most variation were the identification number of each vessel (34 %) and the number of hooks set (10 %). The two other variables explained 10 % between them.

Table 7: Stepwise regression results for Bay of Plenty longline catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	n.s.
Fishing year*	0.066				
Vessel*	-	0.405			
No. of hooks*	-	0.375	0.504		
Target*	-	0.136	0.448	**0.552**	
Month	-	0.076	0.416	0.513	0.561

The residuals in the final model were normally distributed across plus or minus two standard deviations (95%) of the range of the data (Figure 24). The plot of Cook's distance scores shows that no observations had undue influence on the model fit (Figure 24). There was no evidence for failure in the homogeneity of variance assumption as seen in the residual plots for each level of the categorical variable and across the range of the continuous variables (Appendix 17). Finally, there was no trend (i.e. significant divergence from the zero line) in the model residuals evident in the temporal order residual plot to suggest that there may have been autocorrelation in the data, i.e. a violation of the independence assumption (Appendix 18). Similarly, there was no evidence for a lack of homogeneity in the residuals plotted in temporal order (Appendix 18).

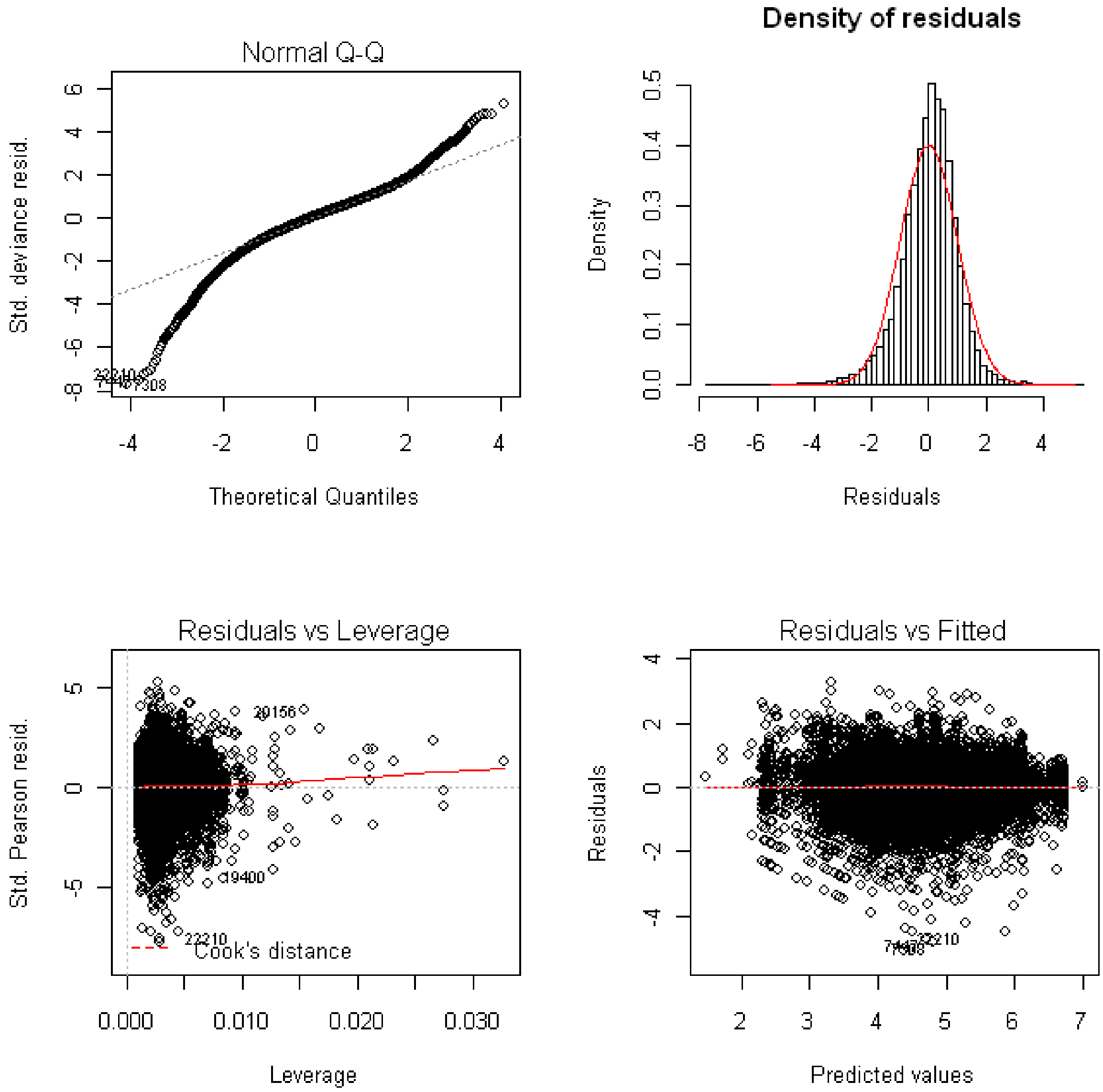


Figure 24: **Analysis of regression model fit for Bay of Plenty longline catches. Top left panel is a Q-Q plot comparing the quantiles of a theoretical normal distribution to that of the residuals from the regression analysis. Top right panel is a histogram of the residuals from the regression analysis. Bottom left panel is a Cook's distance plot investigating the leverage of individual data points. Bottom right panel is a plot of model residuals against the fitted values of the model itself.**

Standardised indices

Although vessel identification number explained the most variation in the data (34 %, Table 7), no one variable stood out as being highly explanatory on the regression step plot (Figure 25). From the influence plots (Appendix 19), the number of hooks had the largest scale of influence.

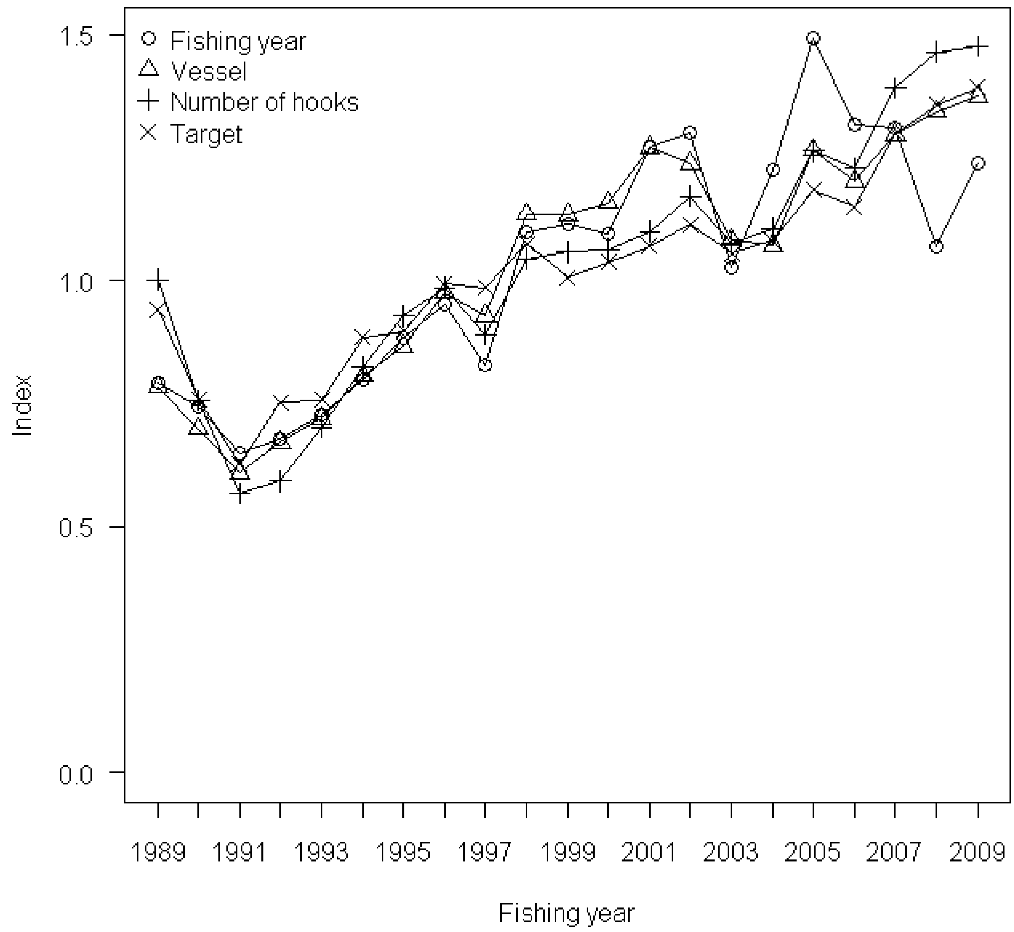


Figure 25: Step plot of the canonical index for Bay of Plenty longline catches. The model starts with just Fishing year and sequentially adds additional variables to the model (symbols lower on the legend).

In terms of the CPUE index itself, it generally increased over the 21 year time series, with the CPUE at the end of the time series being 148 % of that at the beginning of the time series in 1989 (Figure 26). Variation around the index estimates was small; the values of the index and its confidence intervals can be found in Appendix 20.

The final Bay of Plenty index shows a general increase over the 21 year time series, with the CPUE at the end of the time series being 148 % of that at the beginning of the time series in 1989; variation around the index estimates was small (Figure 26; Appendix 20).

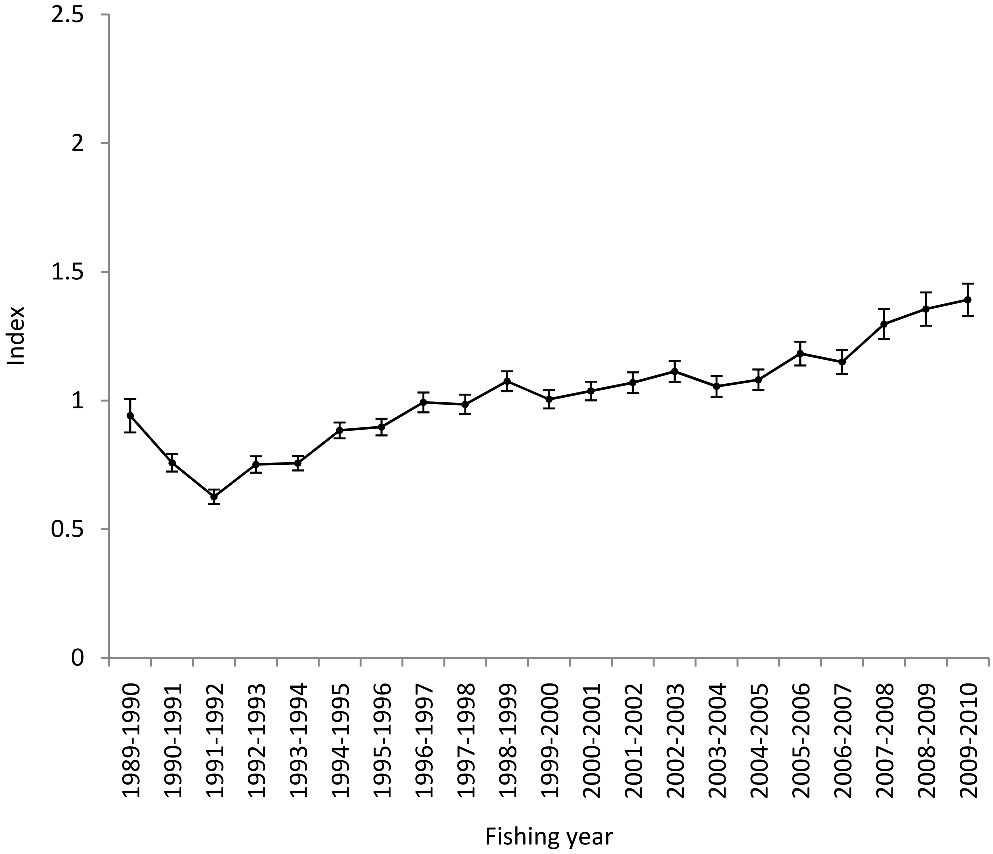


Figure 26: Canonical year indices for Bay of Plenty longline catches, error bars are 95 % confidence intervals.

3.2.2 Bay of Plenty single trawl

3.2.2.1 Data grooming and vessel selection

Frequency distributions of catch and the number of hooks and number of sets were used to determine plausible values for data grooming. Warning flags were set when fishing duration was greater than 8 hours, effort height greater than 15 m, effort width greater than 55 m, effort depth greater than 200m or catch greater than 2000 kg. The data set was then groomed line by line paying particular attention to flagged data entries (Appendix 21).

A common error in the data sets was the transposition of entries for effort height and width. This was often easy to identify and correct. Missing values and typing errors for variables such as effort width and height were also often easy to identify and correct to match the previous pattern for that vessel. Where entries for the catch of snapper or fishing duration were in doubt that whole data line was deleted. After this grooming process was conducted the catch of snapper was plotted by depth (results not presented) and a decision was made to remove all events conducted at a depth greater than 150 m. While snapper do occur below this depth, their abundance is reduced implying that catch rates in deeper water would not be comparable to those from shallower areas. This removed 55 tonnes of snapper catch accounting for 1% of the core fleet's catch.

The final dataset consisted of 33 285 lines of data accounting for 4766 tonnes of snapper catch from 14 vessels. Within the final data set 13% of all events did not capture snapper. Zero versus positive catches were used to derive a second, binomial abundance index.

Data exploration

Log transformed catch data did not appear to have any major departures from that of a normal distribution (Appendix 24). The spread of catch data, across the different levels of the categorical explanatory variables, appeared to be even (Appendix 24), therefore conforming to the assumption of homogeneous variances. A collinearity plot (Appendix 24) found no significant correlations between covariate terms, i.e. no evidence of collinearity.

Model selection and CPUE standardisation

The full stepwise regression analysis of the Bay of Plenty positive trawl catches (i.e. a model that did not force fishing year and allowed interactions) explained 20 % of variation in the data using five model terms (Table 8). Fishing year was not selected in the final model, neither were any fishing year interaction terms.

Table 8: Stepwise regression results for the full model analysis (i.e. fishing year not forced and interactions allowed) of Bay of Plenty positive trawl catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	ns	.
Target	0.0793						
Vessel	0.0648	0.1431					
Statistical area	0.0301	0.1129	0.1654				
Month	0.0119	0.0993	0.1600	0.1856			
Log Depth (3)	0.0344	0.0977	0.1631	0.1833	**0.2040**		
Log Duration (3)	0.0060	0.0828	0.1435	0.1660	0.1861	0.2050	

The constrained log normal stepwise regression analysis of the positive catches from the Bay of Plenty trawl fishery produced an index that explained 21 % of variation in the catch data using six variables (Table 9). Fishing year explained less than 1% of the variation in catch; the covariates that explained the most variation were target (9 %), and vessel (6 %). No effort terms were selected in the final model.

Table 9: Stepwise regression results for log normal analysis of Bay of Plenty trawl catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	6	n.s.
Fishing year	0.0095						
Target		0.0927					
Vessel		0.0704	0.1508				
Statistical area		0.0379	0.1232	0.1723			
Month		0.0207	0.1113	0.1672	0.1920		
Log Depth (3)		0.0425	0.1096	0.1696	0.1696	**0.2093**	
Log Duration (3)		0.0161	0.0960	0.1511	0.1511	0.1728	0.2112

The residuals in the final model were normally distributed across plus or minus two standard deviations (95%) of the range of the data (Figure 27). The plot of Cook's distance scores shows that no observations had undue influence on the model fit (Figure 27). There was no evidence for failure in the homogeneity of variance assumption as seen in the residual plots for each level of the categorical variable and across the range of the continuous variables (Appendix 25). Finally, there was no trend (i.e. significant divergence from the zero line) in the model residuals evident in the temporal order residual plot to suggest that there may have been autocorrelation in the data, i.e. a violation of the independence assumption (Appendix 26). Similarly, there was no evidence for a lack of homogeneity in the residuals plotted in temporal order (Appendix 26).

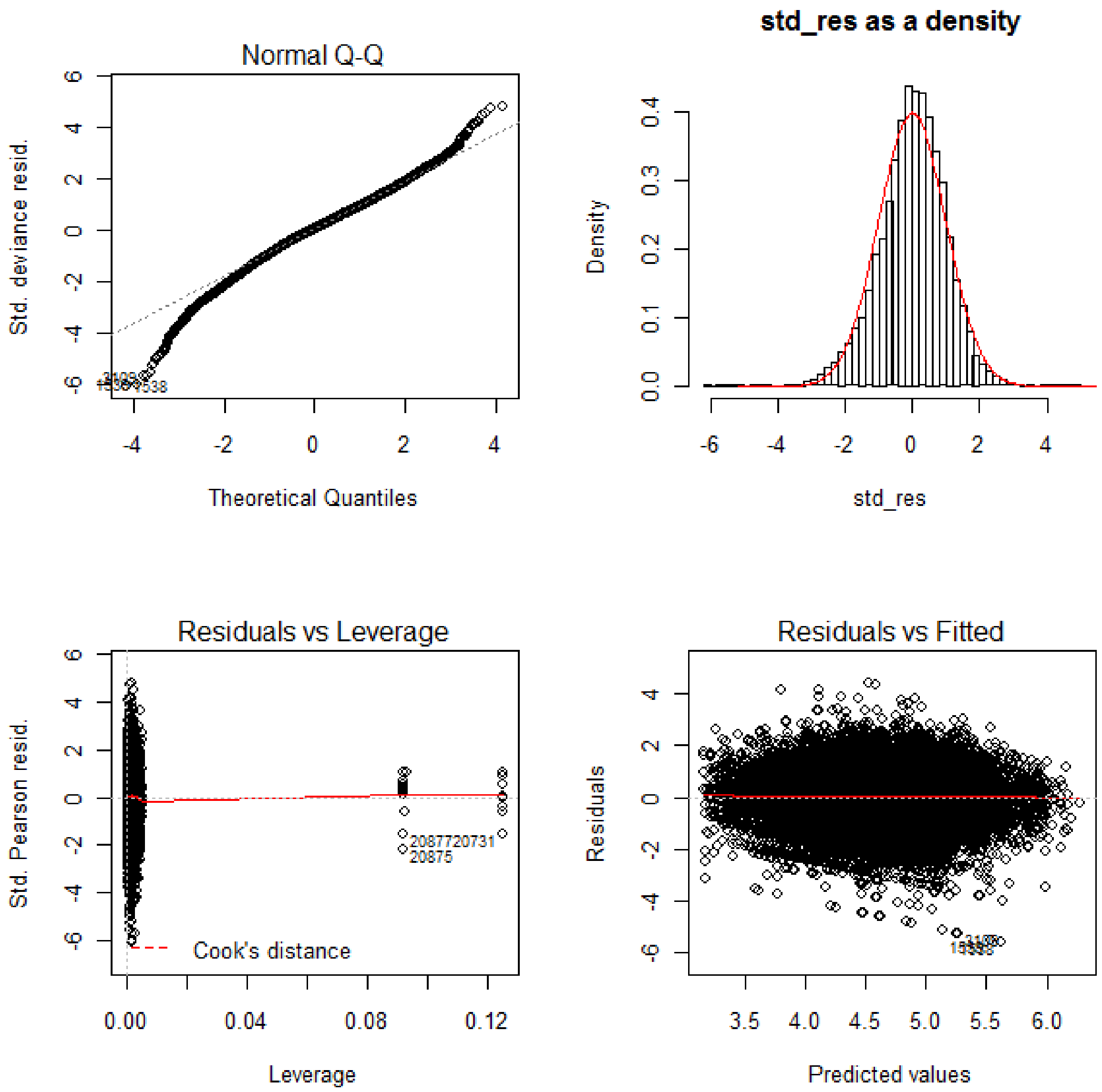


Figure 27: Analysis of the log normal regression model fit for Bay of Plenty trawl catches. Top left panel is a Q-Q plot comparing the quantiles of a theoretical normal distribution to that of the residuals from the regression analysis. Top right panel is a histogram of the residuals from the regression analysis. Bottom left panel is a Cook's distance plot investigating the leverage of individual data points. Bottom right panel is a plot of model residuals against the fitted values of the model itself.

The most influential covariate on the CPUE index was the target species, which only had a noticeable influence on the first year in the series (Figure 28). The addition of the other model terms had little impact on the pattern in the year index (Figure 28). The magnitude of influence the covariate parameters had on the year index is plotted in Appendix 27.

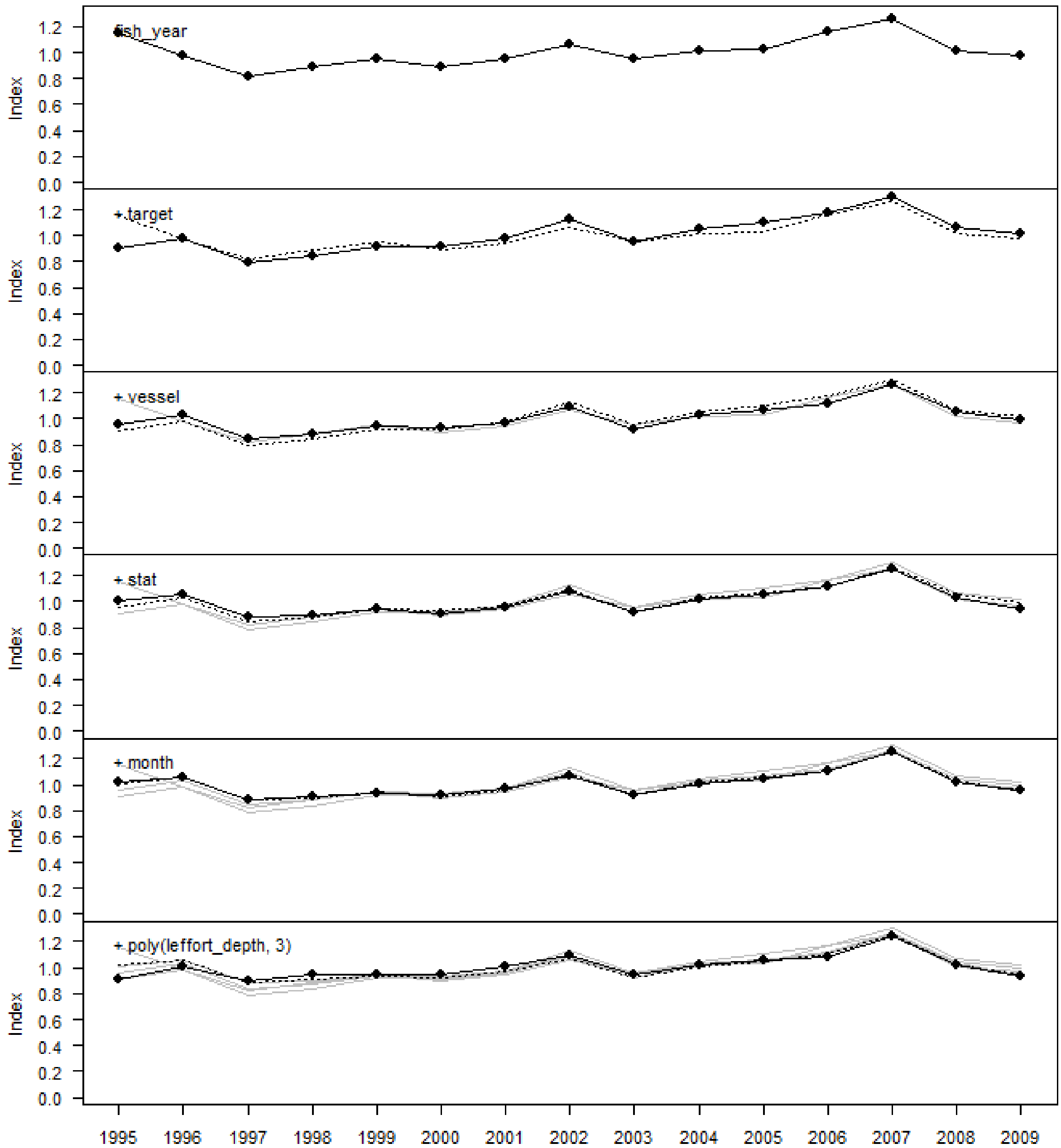


Figure 28: Step plot of the canonical log normal index for Bay of Plenty trawl catches. The model starts with just Fishing year and sequentially adds additional variables to the model.

The full stepwise regression analysis of the Bay of Plenty zero and positive catches (binomial model) explained 26 % of variation in the data using five model terms (Table 10). Unlike the positive catch lognormal model, fishing year was the third selected term and explained 2% additional variation in the model (Table 10). A fishing year/vessel interaction term explained an additional 4.5% of variation (Table 10) and may be cause for caution in the interpretation of the final model index.

Table 10: Stepwise regression results for the binomial full model analysis (i.e. fishing year not forced and interactions allowed) of Bay of Plenty positive trawl catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	5	n.s.
Target	0.1333					
Vessel	0.0327	0.1833				
Fishing year	0.0130	0.1541	0.2066			
fish_year:vessel	-	-	-	0.2522		
fish_year:target	-	-	-	0.2218	**0.2646**	
Month	0.0080	0.1417	0.1928	0.2169	0.26184	0.2737

The constrained binomial stepwise regression analysis (catch = 0 or 1) of the Bay of Plenty trawl fishery produced an index that explained 21 % of variation in the catch data using three variables (Table 11). The variable that explained the most variation was the target of each event (15 %) with the remaining variables explaining 6.5 % between them.

Table 11: Stepwise regression results for binomial analysis of Bay of Plenty trawl catches. Numbers in the top row represent successive iterations of the regression (n.s. represents the final non-significant iteration). Numbers on the diagonal represent final model r^2 values for each variable chosen by the model; numbers below the diagonal represent r^2 values reached for that variable on previous iterations. Value in bold represents the final r^2 reached by the model.

Variable	1	2	3	4	ns
Fishing year	0.0131				
Target		0.1541			
Vessel		0.0484	0.2067		
Month		0.0230	0.1624	**0.2169**	
Log Depth		0.0735	0.1646	0.2122	0.2232

The progressive adding of covariate terms resulted in a steepening of the model index (Figure 29).
A literal interpretation of the index is that the stock increased in size by 900% between 1995 and 2007 then halved in size between 2007 and 2009 (Figure 29).

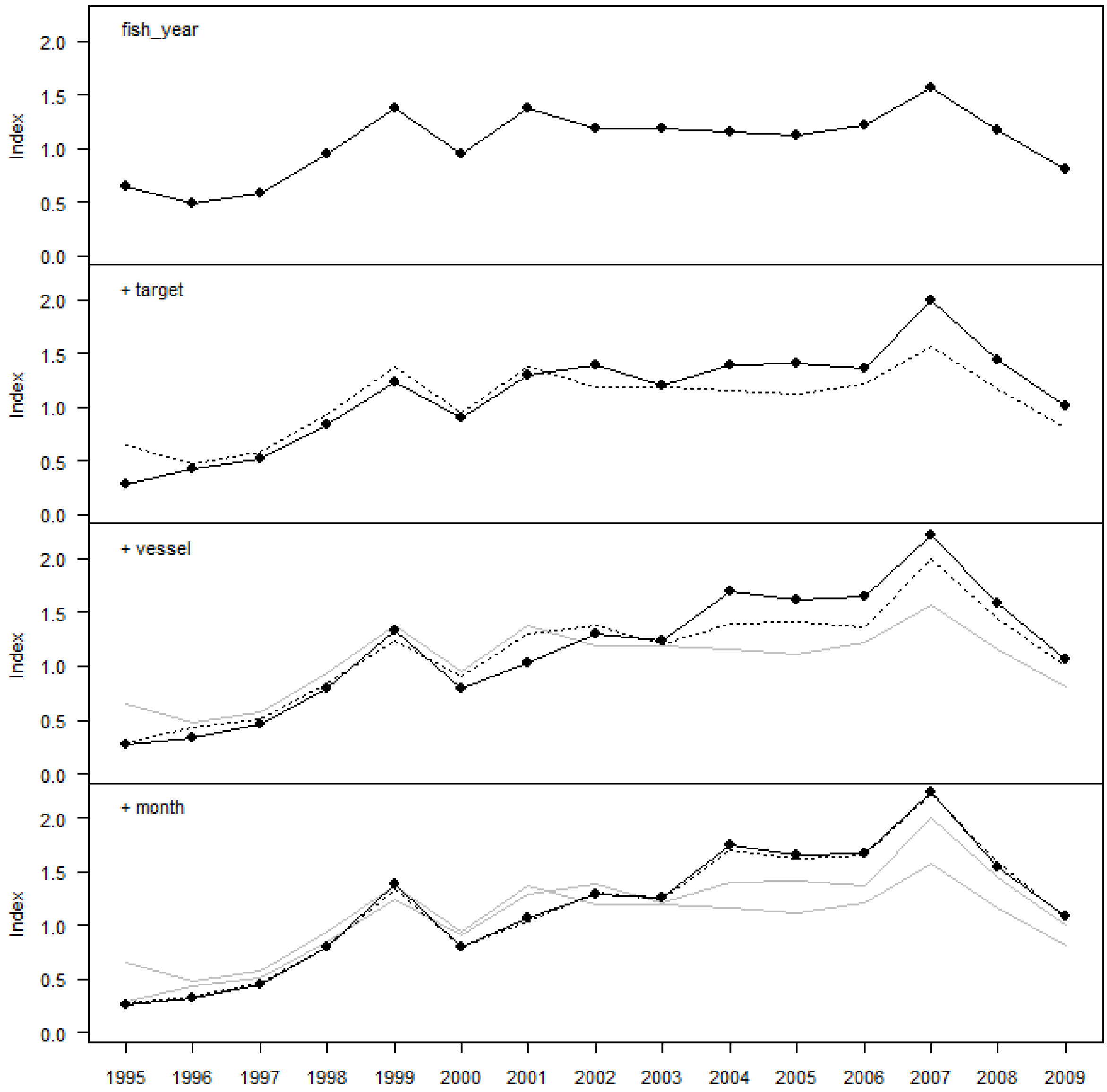


Figure 29: Step plot of the canonical binomial index for Bay of Plenty trawl catches. The model starts with just Fishing year and sequentially adds additional variables to the model.

The final combined trawl index generally reflected the pattern of the log normal and binomial regression analyses (Figure 30;Appendix 28). That was a generally increasing trend, peaking in 2007 before declining.

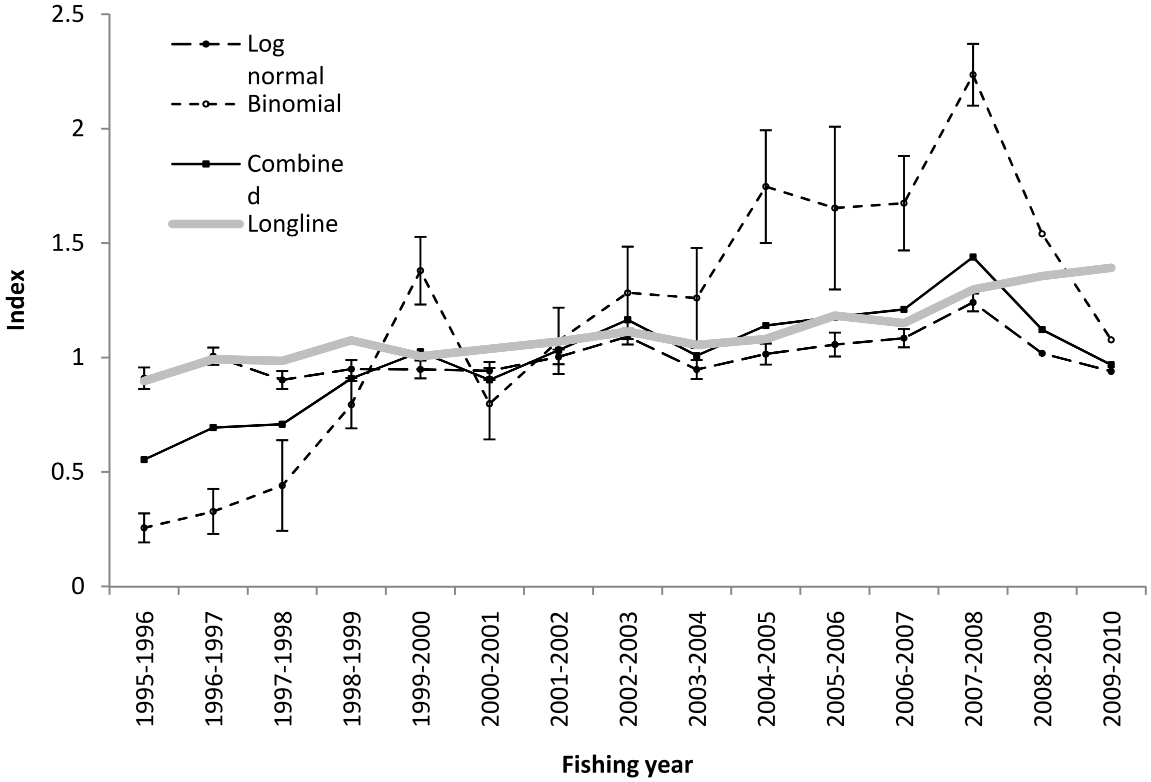


Figure 30: Canonical year indices for Bay of Plenty trawl catches, error bars are 95 % confidence intervals.

In summary; the Hauraki Gulf and Bay of Plenty 20 year standardised CPUE indices show similar increasing trends (Figure 31). In contrast, the East Northland CPUE index exhibits a broad cyclic pattern; the starting and terminating abundance levels being the same (Figure 31).

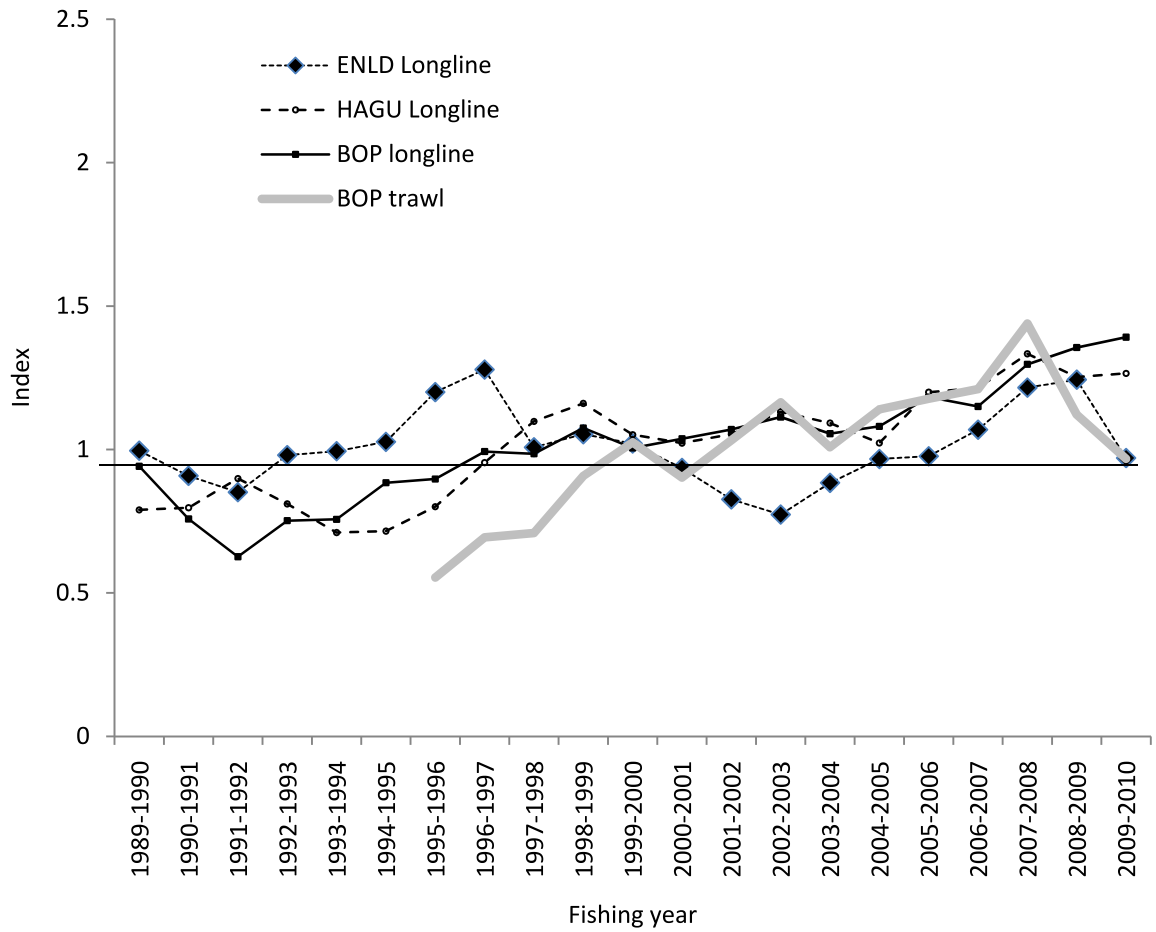


Figure 31: **Comparison of SNA 1 standardised CPUE indices for East Northland, Hauraki Gulf and Bay of Plenty (longline and trawl fisheries).**

4 DISCUSSION

The recent fishery characterisation showed that a high proportion of the annual SNA 1 catch was taken from the Bay of Plenty. Recruited biomass estimates for the Bay of Plenty sub-stock derived from tagging in 1985 and 1994 are only 30% of the Hauraki Gulf biomass and 50% of the East Northland biomass (McKenzie in press). If these biomass ratios are still correct then the catch taken from the Bay of Plenty during the mid-2000s is likely to have been disproportionately high relative to the other SNA 1 sub-stocks and may be cause for concern.

The standardised longline abundance indices derived for the Hauraki Gulf and Bay of Plenty sub-stocks show similar increasing trends and were accepted by the working group as indices of abundance. The Bay of Plenty trawl lognormal index also increased at a similar magnitude to the longline index over most of the series; however a decline in the last two years of the single trawl index was not evident in the longline index. The decreasing trend in the last two years of the trawl index may have been influenced by a change in the effort reporting form types used (which allowed for eight instead of five species to be reported). This change could have resulted in the landed catch being allocated to more tows, thereby introducing a negative bias to the lognormal CPUE and the proportion of tows with zero snapper catch. The working group suggested that the potential for this effect be further investigated when the index is updated for the upcoming SNA1 stock assessment.

In contrast to the other two SNA 1 sub-stocks, the East Northland standardised longline abundance index shows a pronounced cyclical trend over the 21 year series, with the initial and final years in the series at similar relative levels. The WG requested additional sensitivity analysis of the EN index involving the removal of vessels with increasing trends in the number of hooks per set/day. This analysis (not included in this report) showed that the CPUE trend was robust to vessel selection.

Vessel and number of hooks explained most of the variation in each of the three sub-area longline standardisation models. The vessel/fishing year interaction term was significant in all three sub-area models and explained a reasonable amount of the variation in catch in East Northland and the Hauraki Gulf; less so in the Bay of Plenty. Although not fitted in the final models the presence of the vessel /fishing year interaction in the longline data means that the abundance indices vary between vessels weakening our confidence in the final model index magnitude of change and the estimated variation. Evidence for a violation of the homogeneity of variance assumption was found in the temporal order plot of the Hauraki Gulf longline model residuals, which again means that the actual precision on the index is likely to be significantly higher than estimated. The Hauraki Gulf and Bay of Plenty longline indices, however, were accepted by the working group as indices of abundance.

The predominately upward trend in the Bay of Plenty single trawl combined index was largely driven by the binomial index; in contrast the lognormal index was relatively flat. Significant vessel fishing-year and vessel target interactions were present in the Bay of Plenty single trawl binomial data and the same caveats apply as for the longline model. No effort terms (e.g. tow duration, net height, vessel power) were selected by either the final lognormal or the binomial Bay of Plenty single trawl models; the variation in the data were mostly explained by the target and vessel covariates. Due to the presence of these interactions and the poor corroboration with the lognormal index, both the binomial and combined BOP single trawl indices were rejected by the working group. The working group further recommended that the trawl binomial index should not be updated, or used in a future stock assessment.

In summary; the presence of interactions with Fishing year in some of the analyses; a lack of homogeneity in the Hauraki Gulf longline data in combination with the general recognition that changes in CPUE over time may not necessarily be driven by changes in fish abundance (e.g. see Rose

& Kulka 1999 and Maunder & Punt 2004 for a discussion of these issues), suggests that the indices presented should be interpreted with caution. With these considerations in mind the working group was satisfied that the East Northland, Hauraki Gulf and Bay of Plenty longline, and Bay of Plenty lognormal indices did reflect abundance, whereas the Bay of Plenty binomial and combined single trawl indices did not.

5 ACKNOWLEDGMENTS

We thank the Ministry of Fisheries for funding and the Data Management Team for supplying catch information. We would also like to thank the Northern Inshore Working Group, especially Marc Griffiths and Paul Starr, for constructive feedback which greatly improved this report, and Marianne Vignaux for final editing of the report.

6 REFERENCES

Bentley, N.; Kendrick, T.H.; Starr, P.J.; Breen, P.A. (in press) Influence plots and metrics: tools for better understanding fisheries catch per unit effort standardisations. ICES Journal Marine Science.

Cook, R.D.; Weisberg, S. (1982). Residuals and influence in regression. Chapman and Hall, London.

Francis, R.I.C.C. (1999). The impact of correlations in standardised CPUE indices. New Zealand Fisheries Assessment Research Document 99/42. (Unpublished report held by NIWA library, Wellington). 30 p.

Gilbert, D.J.; McKenzie, J.R.; Field, K.D.; Davies, N.M. (2000). Assessment of the SNA 1 stocks for the 1999–2000 fishing year. New Zealand Fisheries Assessment Report 2000/38, 52 p.

Hartill, B.; Bian, R.; Armiger, H.; Vaughan, M.; Rush, N. (2007). Recreational marine harvest estimates of snapper, kahawai, and kingfish in QMA 1 in 2004–05. New Zealand Fisheries Assessment Report 2007/26. 44p.

Kendrick, T.H.; Bentley, N. (2011). Fishery characterisations and catch-per-unit-effort indices for three sub-stocks of red gurnard in GUR 1, 1989–90 to 2008–09. New Zealand Fisheries Assessment Report 2011/4. 81p

McKenzie, J.R. (2008) Catch per unit effort analysis of SNA 1 longline (1990–2005) and Bay of Plenty single trawl (1996–2005) fisheries. New Zealand Fisheries Assessment Report 2008/36. 22 p.

McKenzie, J.R.; Gilbert, D.; Bian, R. (2011). A simulation analysis to derive cost-optimised mark-recapture designs for estimating the biomass of east coast North Island snapper (Pagrus auratus) sub-stocks (SNA 1). New Zealand Fisheries Assessment Report 2011/39. 170p

McKenzie, J.R. (in pub). An evaluation of fully age-structured spatially disaggregated stock assessment models for the SNA 1 QMA. Draft New Zealand Fisheries Assessment Report 2011/ 81 p.

Maunder, M.N.; Punt, A.E. (2004). Standardizing catch and effort data: a review of recent approaches. *Fisheries Research 70:* 141–151.

Paul, L. J. (1976) A study on age, growth, and population structure of the snapper, *Chrysophrys auratus* (Forster), in the Hauraki Gulf, New Zealand Fisheries Research Division Bulletin, Fisheries Research Division, N.Z. Ministry of Agriculture & Fisheries, No 13 84 p.

Rose, G.A.; Kulka, D.W. (1999). Hyperaggregation of fish and fisheries: How catch-per-unit-effort increased as the northern cod (Gadus horhua) declined. *Canadian Journal of Fisheries and Aquatic Sciences, 56:* 118–127.

Vignaux, M. (1994). Catch per unit effort (CPUE) analysis of west coast South Island and Cook Strait spawning hoki fisheries, 1987–93. New Zealand Fisheries Assessment Research Document 94/11, 29 p. (Unpublished report held by NIWA library, Wellington).

Walsh, C.; Buckthought, D.; Armiger, H.; Spong, K.; Vaughan, M.; Smith, M.; Kohn, Y. (2009). Age composition of commercial snapper landings in SNA 1, 2007–08. New Zealand Fisheries Assessment Report 2009/46 40 p.

Zuur, A.F.; Ieno, E.H.; Elphick, C.S. (2009). A protocol for data exploration to avoid common statistical problems Methods in Ecology and Evolution 2010, 1, 3–14

7 APPENDICES

Appendix 1: Data extracted for SNA 1 fishery characterisation

Details from all trips landing SNA 1 between

1/10/2000 and 30/9/2010

This request generated two component data tables

- Effort table data and estimated catch information
- Landed green-weight catch information

The extract criteria were as follows:

1. Identify all trips where SNA1 quota was landed.
2. Provide ALL effort data, at tow or set level, in relation to ALL trips identified in (1) regardless of snapper being recorded in the tow/set or not.

Specific fields requested:

Effort and Estimated Catch table

vessel_id
event_key
dcf_key
Trip start date
Trip end date
trip_key
Fishing Date
start_latitude
start_longitude
primary_method
start_stats_area_code
target_species
Total catch (kgs)
species_code
effort_depth
Species Est catch (kgs)
Form_type

Landed catch table

Fishstock [should be all SNA 1 but please output it anyway]
dcf_key
trip_key
Point of Landing
Landing date
Trip start date
Trip end date
vessel_id
Greenweight (kgs)
Container type
Destination Type
Form_type

Appendix 2: Ministry Catch Effort data generic errors and ambiguities.

The Ministry catch effort data as a generality can be categorised into:

1. Landed catch information;
2. At-sea or effort information.

For analytical purposes there is usually a requirement to link the two data sources together, but because these data are often collected on more than one type of form the information is often “decoupled” and difficult to link; as a result a large number of errors and ambiguities inherent in the Ministry catch-effort data are a result of “orphaned” landing and effort data.

The quality of fisheries data varies greatly between fisheries and those undertaking analysis need to be aware of the specific data quality issues for each respective fishery. It is often the case that data errors cannot be corrected or inferred and the analyst must make a judgement call (often subjective) on what to exclude from the analysis. Analyses using fisheries catch effort information should ideally include a summary of the data errors and a description on how they were dealt with.

The following are some of the more common causes of erroneous catch and effort data:

Double recording of landed catch weights

Double recording of landed catch can come about as a result of fish being transferred to a location whereby it becomes part of the “Catch” of another trip. Double recording of landed weights is usually not as a result of erroneous recording, but typically comes about as a legitimate artefact of the catch reporting process. A typical instance where catch totals are reported twice is when a catch from one vessel is transferred to another, say at sea. The transference at sea by the first vessel constitutes as legitimate landing event for which the first vessel must complete a landing form. When the second vessel lands, say to shore, it is required to report both its own catch and the catch from the other vessel. Unfortunately there is no requirement for the second vessel to report the two catch totals separately.

Fishers are required to record the destination of all landed catch using a range of single letter codes thus it is possible to rationalise some of this double counting at least in the landed catch information. There are four codes indicating that a catch has been transferred to a non-terminating destination, i.e. the catch will have to be recorded again in a subsequent landing event. These codes are:
P - transferred to a holding receptacle in water (e.g. a lobster holding pot);

Q - transferred to a holding receptacle on land (e.g. a wharf chiller);
R - retained on board;
T - transferred to another vessel.

When summarising landed catch information it is reasonable to ignore any catch associated with these four landing codes as it will appear again as part of another landing (i.e. trip). The problem comes if there is a need to link effort to landed catch. In the above example the effort recorded by the second vessel for the "trip" will not represent all the catch it records as landed on that trip.

There are legitimate instances where fishers land catch in relation to a trip for which there has been no fishing effort; an example of this being where a vessel has received catch from another vessel at sea but has not undertaken any fishing of its own.

The difficultly for the Ministry, and for those analysing Ministry data, is determining whether the absence of corresponding landing (trip) and effort data is legitimate, or that the fisher has failed to provide the effort data. This is an example where an understanding of the fishery is required to determine if the ratio of missing effort to landed data is reasonable.

Missing or misaligned effort information resulting from trip date misspecification

Because the effort and landed catch data from a specific fishing trip is often recorded on separate forms, there is often no formal link between these two sets of information at the time of landing. The formal link, in the form of a trip key, is assigned by the Ministry data collating process at a later time on the basis on the fisher's reported trip start and end dates and the landing date. Effort and landed catch data can therefore become decoupled if a fisher makes an error in recording the trip dates.

Discrepancies between estimated and landed catch weights

For most trips a fisher is required to report catch information twice: first as an estimated catch on the effort reporting forms; second as a landed green weight. The estimated weights are the fisher's best guess at the weight; there is no legally binding requirement that these values are accurate. In contrast there are strong legal requirements for the declared landed weights to be accurate. Consequently estimated catch totals often disagree with the landed catch weights. However, the catch estimates are usually believed to be reasonable in a relative sense, e.g. twice as much of species A was caught in the first tow than the second. The usual practice in most analytical situations is to prorate the effort estimated catches by the actual landed greenweight totals (assuming there are no reporting errors when linking the effort with the landed catch and that the landed catch is reasonable). However, under scenarios where the landed catch is non-terminating the effort and landed catch information will be matched to the effort such that incorrect scaling will result.

"Top five/eight" missing catch effort issue

Prior to 2006 most Ministry effort reporting forms only required estimated weights for the top five species caught in a tow or set. This meant, although all species caught should appear in the landed catch reporting forms, the catch of some species may go unrecorded in effort forms if they were not in the top five species. Reporting forms introduced after 2006 allow reporting of up to eight species making it more likely that effort will be recorded for most species of significance. Inaccurate prorating of the landed catch weights will occur if the "top five/eight" issue is common for the species of interest. In the worst cases the estimated catches will be a very small proportion of the landed catch weight total, making it invalid to simply prorate on the basis of estimated catch alone. In these instances the solution is to use some form of lumping or assigning criteria other than estimated catch (e.g. total amount of fishing effort or total catch of all species in a tow or set). The use of lumping criteria (sometimes referred to as "rolling up") can have a major influence on the interpretation of the results and if inappropriate may significantly bias the resultant analysis. It is very important that criteria for rolling up effort data is clearly described and justified.

Misreporting and general data quality issues

Inaccurate or incomplete reporting is generic issue. Missing data or data outside a normal range are often easy to identify and allow for. Typically missing data is either imputed from the other data provided or the record is simply deleted. The problem comes when the data is plausible but inaccurate. In most instances these types of errors remain unidentified in the dataset.

Data quality of the smaller owner-operator type fishers tends to be more erroneous than data from than the large company fishers. As a generality the effort data has more inherent errors than the landed catch data. Although there are strong legal requirements for fishers to furnish accurate effort information, enforcement is difficult and expensive and seemingly a low priority to the Ministry of Fisheries enforcement wing; fishers tend not to be called to task and there are many examples where a fisher's blatant misreporting has been allowed to continue unchallenged over a number of years.

Appendix 3: Data extracted for SNA 1 longline CPUE analysis.

Details from all **longline** (BLL, SLL) trips landing SNA1 quota between

1/10/1989 and 30/9/2010

This request generated two data extract tables:
- Longline effort data and estimated catch totals from CELR and LTCER reporting forms
- Associated landed catch greenweight totals

The extract criteria were as follows:
1. All Longline (BLL, SLL) trips where SNA1 quota was landed.

2. ALL effort data, at finest reporting scale in relation to ALL trips identified in (1) regardless of snapper being recorded in event or not.

Specific fields requested:

Effort and Estimated Catch table
vessel_id
event_key
dcf_key
Trip start date
Trip end date
trip_key
primary_method
Fishing Date
start_latitude
start_longitude
start_stats_area_code
target_species
effort_num

total_hook_num
effort_width
effort_length
Total catch (kgs)
species_code
Species Est catch (kgs)
Form_type

Landed catch table

Fishstock [should be all SNA 1 but please output it anyway]
dcf_key
trip_key
Point of Landing
Landing date
Trip start date
Trip end date
vessel_id
Greenweight (kgs)
Container type
Destination Type
Form_type

Appendix 4: Data extracted for SNA 1 Bay of Plenty single trawl CPUE analysis.

Details from all **single trawl** (BT) trips landing SNA1 quota between

1/10/1989 and 30/9/2010

This request generated three data tables:

- single trawl effort data and estimated catch totals CELR, TCEPR, TCE
- Associated landed catch greenweight totals
- vessel specific details

The extract process was as follows:

1. Indentify all single trawl (BT) trips where SNA1 quota was landed.
   and/or
   All BT trips where there is at least one event record targeting SNA TRE TAR JDO GUR or BAR in SNA 1 (stat areas 001 – 010)

2. ALL effort data, at finest reporting scale in relation to ALL trips and events indentified in (1) regardless of snapper being recorded in event or not.

Effort and Estimated Catch table
vessel_key
event_key
dcf_key
Trip start date
Trip end date
trip_key
primary_method

Fishing Date
start_latitude
start_longitude
start_stats_area_code
target_species
fishing_duration
effort_height
effort_num
effort_total_num
effort_width
effort_depth
Total catch (kgs)
species_code
Species Est catch (kgs)
Form_type

Landed catch table

Fishstock [should be all SNA 1 but please output it anyway]
dcf_key
trip_key
Point of Landing
Landing date
Trip start date
Trip end date
vessel_key
Greenweight (kgs)
Container type
Destination Type
Form_type

Vessel Details
vessel_key
overall_length_metres
draught_metres
beam_metres
built_year
engine_kilowatts
history_start_datetime
history_end_datetime

Appendix 5: SNA 1 annual commercial catch (t) by method and area (BLL = bottom longline; BPT = bottom pair trawl; BT = bottom trawl; DS = Danish seine; SN =setnet).

	ENLD						HAGU						BOP					
FYear	BLL	BPT	BT	DS	SN	other	BLL	BPT	BT	DS	SN	other	BLL	BPT	BT	DS	SN	other
2000–01	797	47	183	19	21	9	1 121	12	507	249	101	5	392	50	630	200	17	2
2001–02	650	123	202	33	17	5	1 081	3	529	239	122	8	364	29	592	319	14	4
2002–03	486	179	172	10	16	4	1 056	8	520	274	61	3	347	76	746	428	19	5
2003–04	587	124	203	27	12	9	828	14	569	277	109	6	286	110	706	584	12	9
2004–05	482	96	341	118	20	11	742	13	540	251	70	1	336	123	877	526	5	1
2005–06	541	99	456	97	25	10	698	0	622	235	61	2	332	44	773	478	11	2
2006–07	601	120	377	127	32	6	631	12	720	417	44	4	287	9	598	377	5	6
2007–08	556	163	235	131	22	6	688	3	741	459	68	5	260	11	697	408	9	5
2008–09	549	139	254	105	19	6	779	24	899	451	65	3	175	9	572	407	6	3
2009–10	534	96	241	164	16	16	753	16	674	440	33	9	331	13	587	498	8	6

Appendix 6: Annual SNA 1 longline catch (t) by statistical reporting area.

		ENLD				HAGU			BOP	
FYear	001	002	003	004	005	006	007	008	009	010
2000–01	28	432	331	6	432	264	425	239	107	46
2001–02	12	356	277	5	346	342	392	186	128	50
2002–03	17	199	269	2	321	363	372	222	77	48
2003–04	32	257	294	5	297	277	255	175	79	32
2004–05	0	242	239	1	246	236	260	173	109	53
2005–06	0	227	301	13	252	246	199	149	111	72
2006–07	0	313	276	12	256	160	215	142	36	109
2007–08	0	335	217	4	302	171	216	163	73	25
2008–09	0	354	185	11	264	224	291	148	27	0
2009–10	1	280	250	3	218	273	262	296	35	0

Appendix 7: Annual SNA 1 longline catch (t) by target species.

		ENLD			HAGU			BOP	
FYear	SNA	GUR	other	SNA	GUR	other	SNA	GUR	other
2000–01	790	1	6	1 120	0	1	387	5	0
2001–02	647	1	2	1 080	0	0	361	4	0
2002–03	485	0	0	1 055	1	0	346	1	0
2003–04	580	7	1	828	1	0	282	3	1
2004–05	466	13	2	740	2	0	333	3	0
2005–06	529	10	1	698	0	0	331	1	0
2006–07	594	5	1	630	1	0	284	3	0
2007–08	546	1	8	688	0	0	250	9	1
2008–09	542	1	7	778	0	0	174	1	0
2009–10	527	1	6	753	0	0	318	8	5

Appendix 8: Annual SNA 1 bottom trawl catch (t) by statistical reporting area.

		ENLD				HAGU			BOP	
FYear	001	002	003	004	005	006	007	008	009	010
2000–01	1	30	150	2	280	224	3	173	195	262
2001–02	0	49	152	1	253	276	0	112	189	291
2002–03	0	64	107	1	335	182	3	143	265	338
2003–04	0	57	140	6	316	253	0	110	237	359
2004–05	0	60	279	1	315	219	5	182	330	365
2005–06	12	132	297	14	348	271	4	237	268	269
2006–07	0	140	228	8	401	319	0	169	196	233
2007–08	0	83	148	4	418	323	0	186	239	272
2008–09	0	82	167	6	439	459	0	133	184	254
2009–10	0	47	190	4	301	373	1	147	164	276

Appendix 9: Annual SNA 1 bottom trawl catch (t) by target species.

sub-stock	FYear	SNA	BAR	GUR	JDO	TAR	TRE	other
ENLD	2000–01	60	11	3	63	31	14	0
	2001–02	83	8	8	43	24	37	0
	2002–03	73	4	5	28	18	42	1
	2003–04	99	5	2	43	21	33	1
	2004–05	205	1	7	60	30	37	0
	2005–06	242	11	14	51	67	69	1
	2006–07	169	2	17	58	44	87	0
	2007–08	73	2	4	55	45	56	0
	2008–09	126	3	1	37	47	41	0
	2009–10	111	1	9	44	47	29	0
HAGU	2000–01	326	1	46	119	2	5	9
	2001–02	384	1	42	94	2	6	0
	2002–03	341	0	36	132	1	9	0
	2003–04	359	0	47	155	1	6	1
	2004–05	393	0	56	83	0	7	
	2005–06	462	1	42	106	0	10	0
	2006–07	430	0	62	219	0	9	1
	2007–08	506	1	31	173	3	25	0
	2008–09	646	1	24	206	3	18	2
	2009–10	483	0	40	141	1	8	0
BOP	2000–01	224	30	42	22	39	258	15
	2001–02	253	18	45	37	59	178	3
	2002–03	260	30	48	35	140	228	5
	2003–04	294	12	43	20	137	198	2
	2004–05	429	2	51	63	105	222	5
	2005–06	300	7	94	51	108	207	6
	2006–07	336	5	31	55	44	120	6
	2007–08	387	8	10	52	78	161	1
	2008–09	292	0	19	43	63	152	3
	2009–10	300	2	19	40	57	168	1

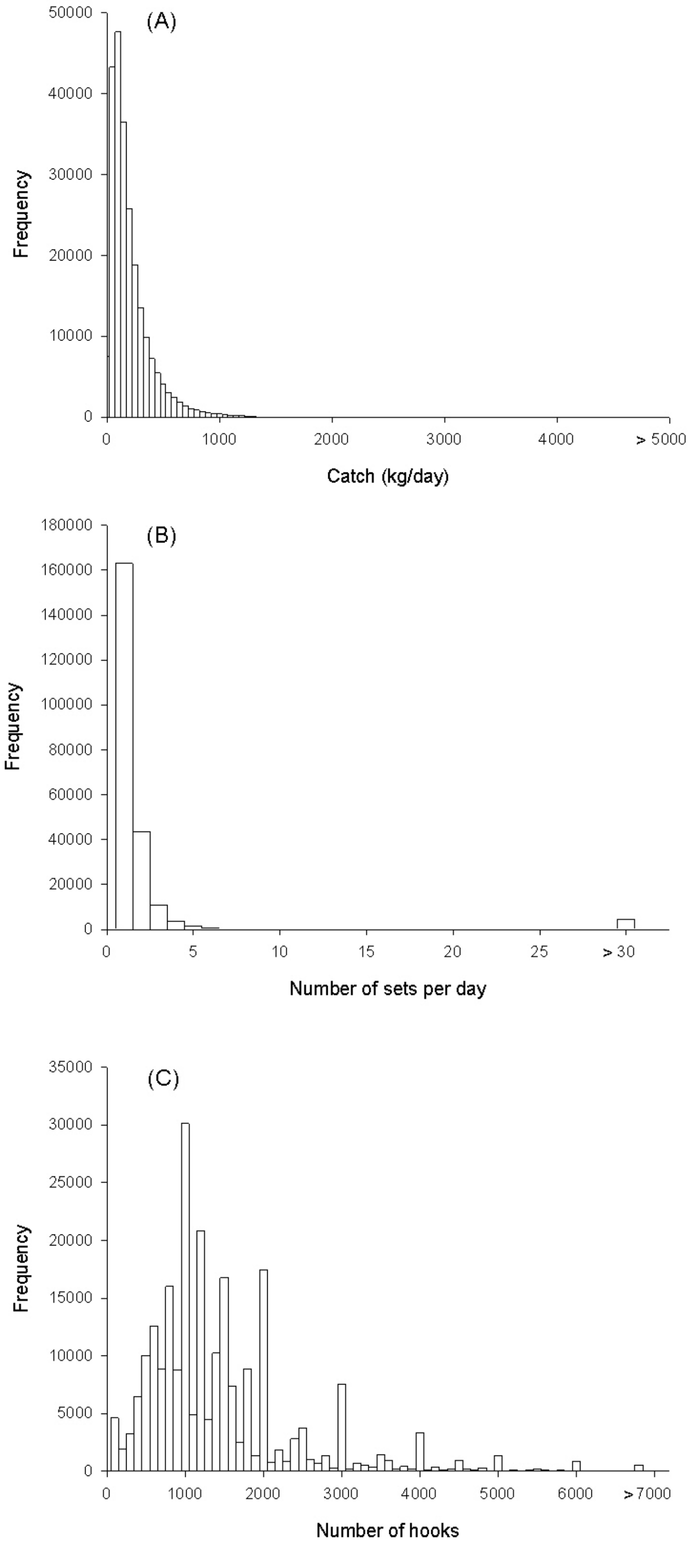
Appendix 10: Annual SNA 1 Danish seine catch (t) by statistical reporting area.

		ENLD				HAGU			BOP	
FYear	001	002	003	004	005	006	007	008	009	010
2000–01	0	0	19	0	17	226	6	63	98	38
2001–02	0	8	24	0	20	210	10	78	168	73
2002–03	0	2	7	0	25	249	1	99	223	106
2003–04	0	24	3	0	6	270	1	140	326	117
2004–05	0	40	78	0	40	210	1	108	292	127
2005–06	0	24	72	1	36	198	1	109	235	134
2006–07	0	1	126	1	66	349	3	51	153	173
2007–08	0	22	108	0	48	410	0	137	214	57
2008–09	1	24	80	0	90	360	1	151	213	43
2009–10	0	44	119	1	106	327	7	142	283	73

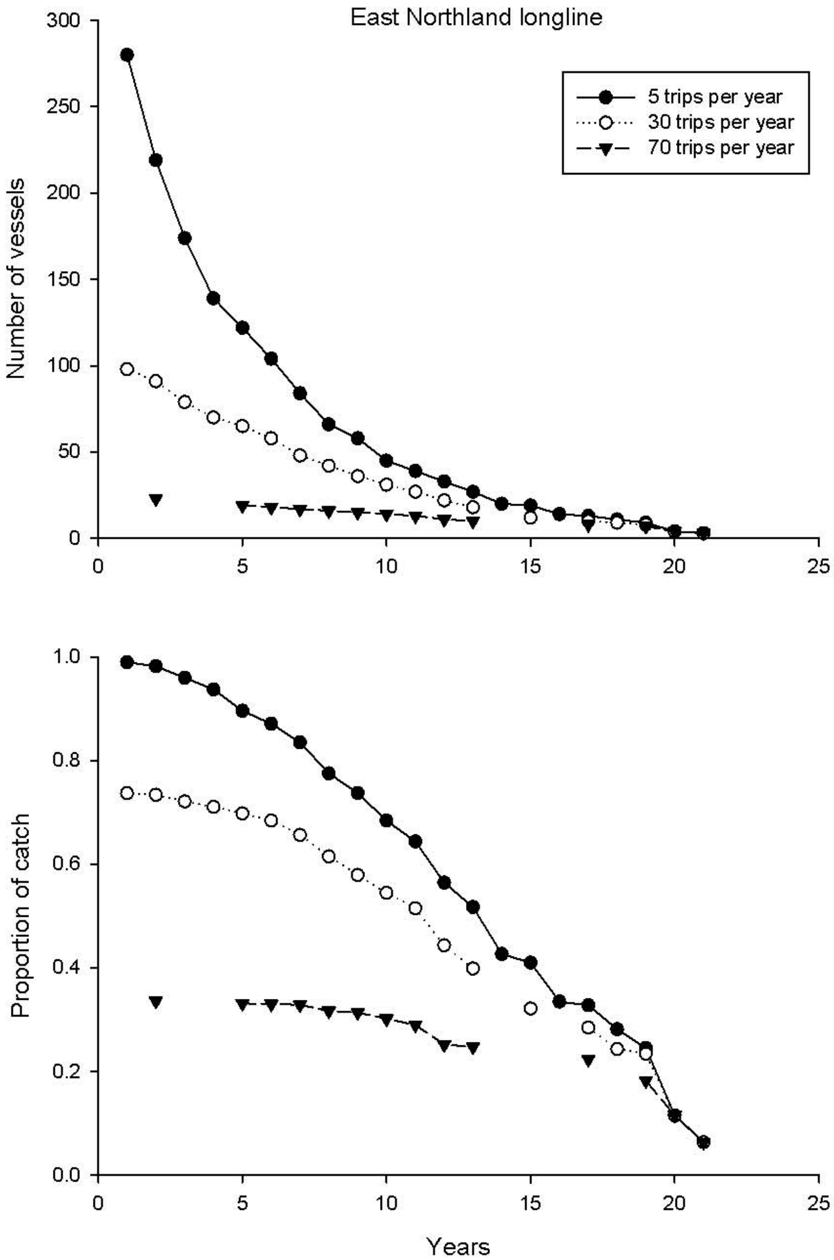
Appendix 11: Annual SNA 1 Danish seine catch (t) by target species.

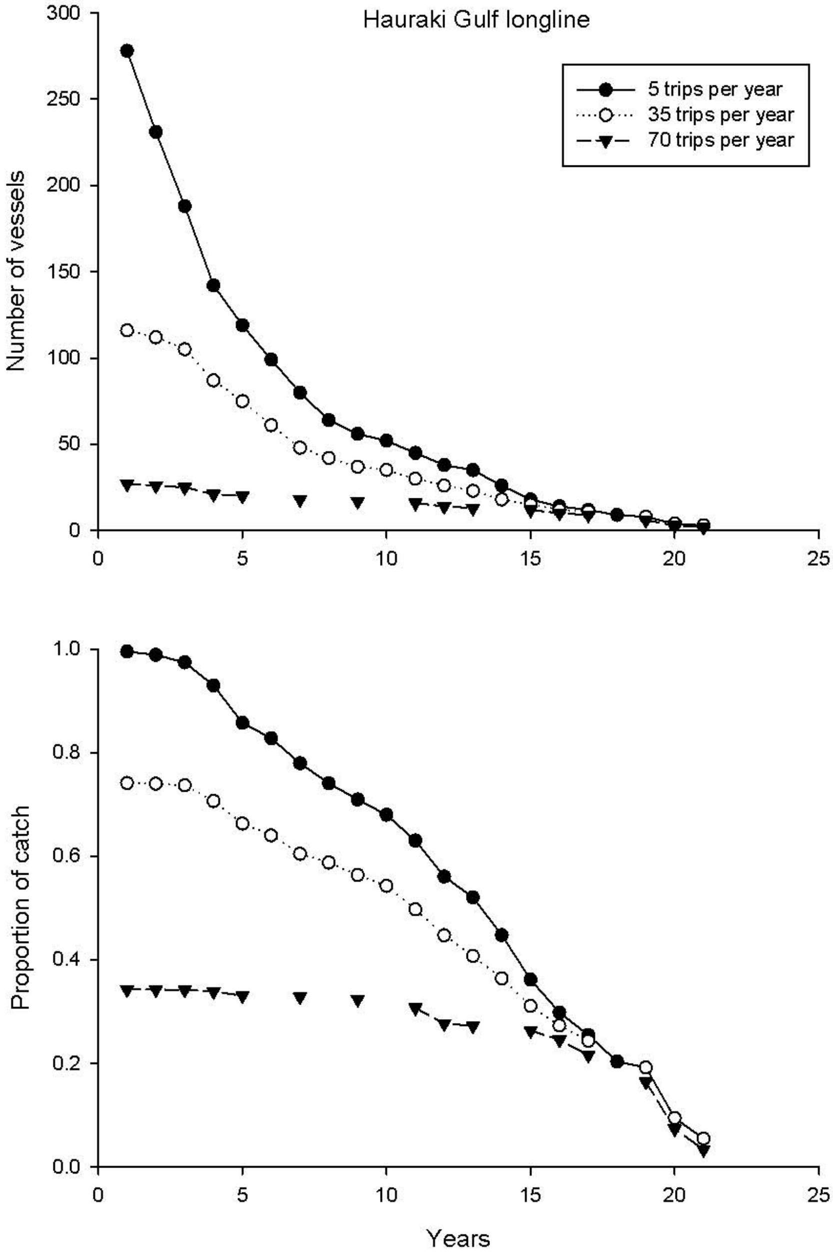
sub-stock	FYear	SNA	GUR	JDO	TAR	TRE	other
ENLD	2000–01	3	15	1	0	0	0
	2001–02	15	17	1	0	0	0
	2002–03	5	3	0	1	0	0
	2003–04	17	5	1	4	0	0
	2004–05	53	64	2	0	0	0
	2005–06	38	58	0	1	0	0
	2006–07	109	17	0	1	0	0
	2007–08	113	13	1	3	0	0
	2008–09	82	15	1	8	0	0
	2009–10	66	54	2	33	10	0
HAGU	2000–01	192	6	50	0	0	0
	2001–02	183	6	50	0	0	0
	2002–03	257	3	15	0	0	0
	2003–04	245	1	28	0	0	2
	2004–05	213	5	33	0	0	0
	2005–06	156	1	78	0	0	0
	2006–07	183	15	219	0	0	0
	2007–08	342	8	109	0	0	0
	2008–09	302	21	121	5	0	1
	2009–10	291	32	109	0	0	7
BOP	2000–01	136	51	1	5	7	0
	2001–02	174	123	1	7	13	0
	2002–03	289	131	3	3	1	0
	2003–04	510	60	5	5	4	0
	2004–05	483	40	0	3	0	0
	2005–06	386	71	3	3	8	7
	2006–07	338	31	6	2	0	0
	2007–08	380	14	11	1	2	0
	2008–09	352	51	0	4	0	0
	2009–10	333	144	1	18	2	1

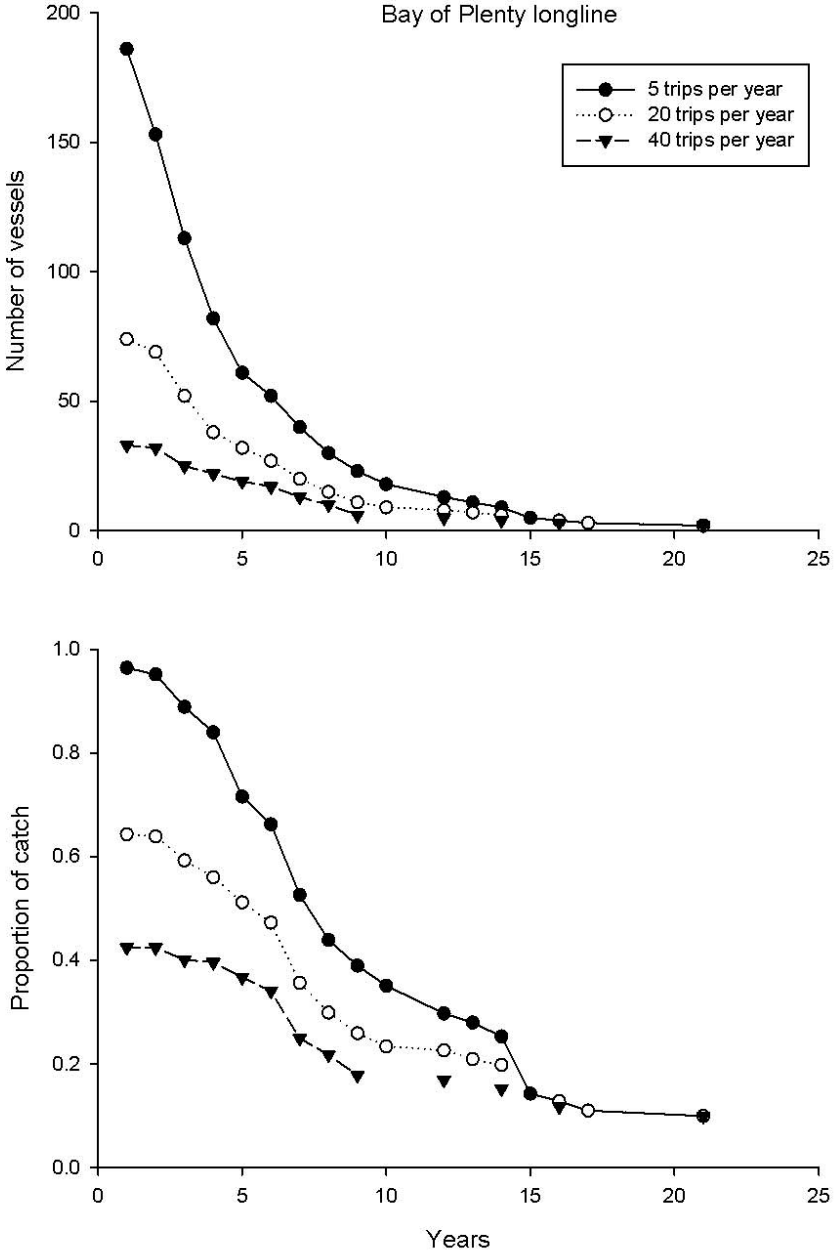
Appendix 12: Frequency distributions for catch (A), number of sets per day (B) and number of hooks set (C) from the snapper bottom longline data set.



Appendix 13: Graphics used to aid the vessel selection process. For each sub-stock/method combination, the number of vessels and the proportion of catch accounted for are plotted against the number of years these vessels contributed to the fishery in that sub-stock. Each of these graphs has three series spanning the range of the number of trips per year vessels undertook.

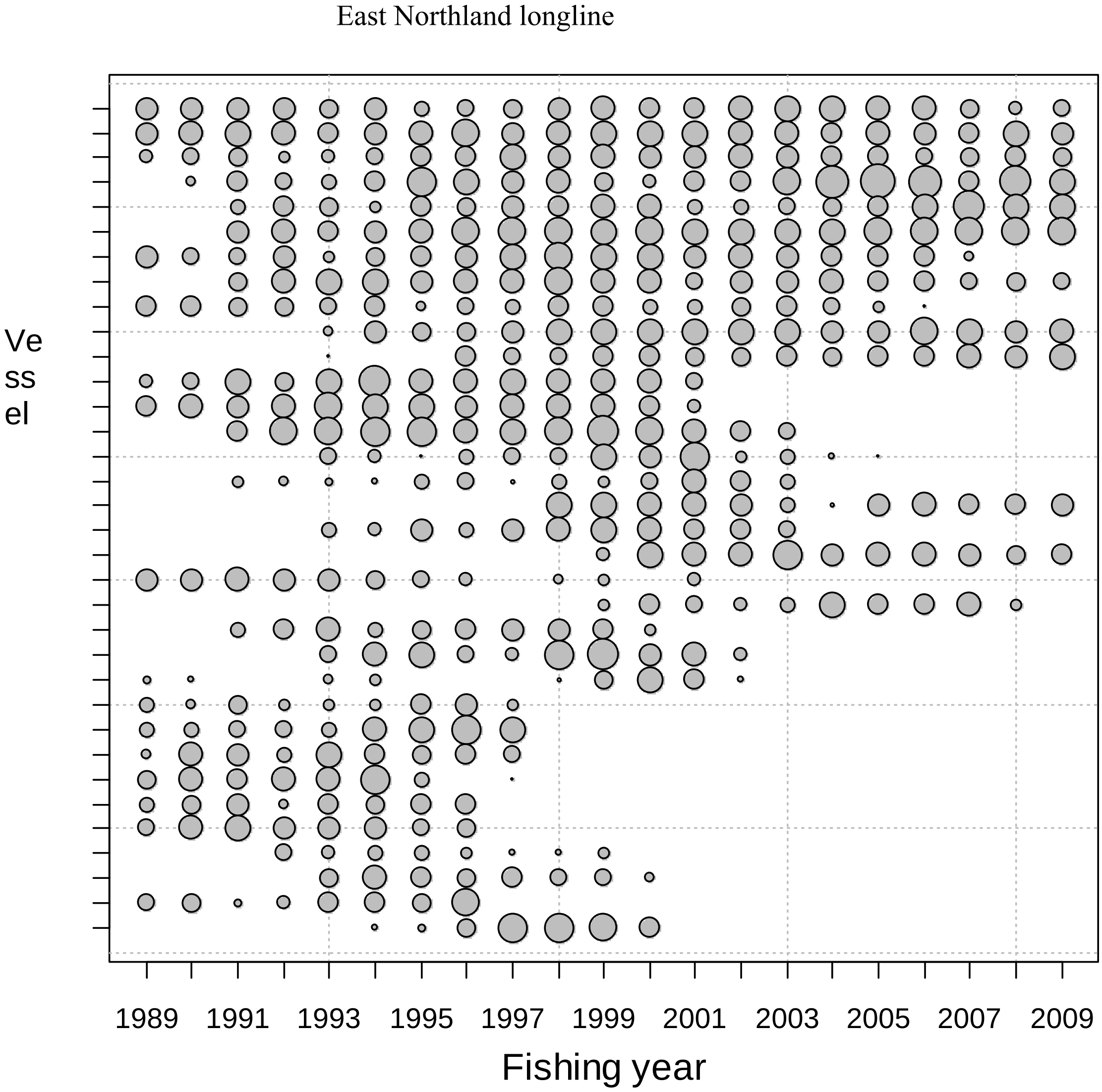



Hauraki Gulf longline
5 trips per year
35 trips per year
70 trips per year
Number of vessels
Proportion of catch
Years
0
5
10
15
20
25
300
250
200
150
100
50
1.0
0.8
0.6
0.4
0.2
0.0

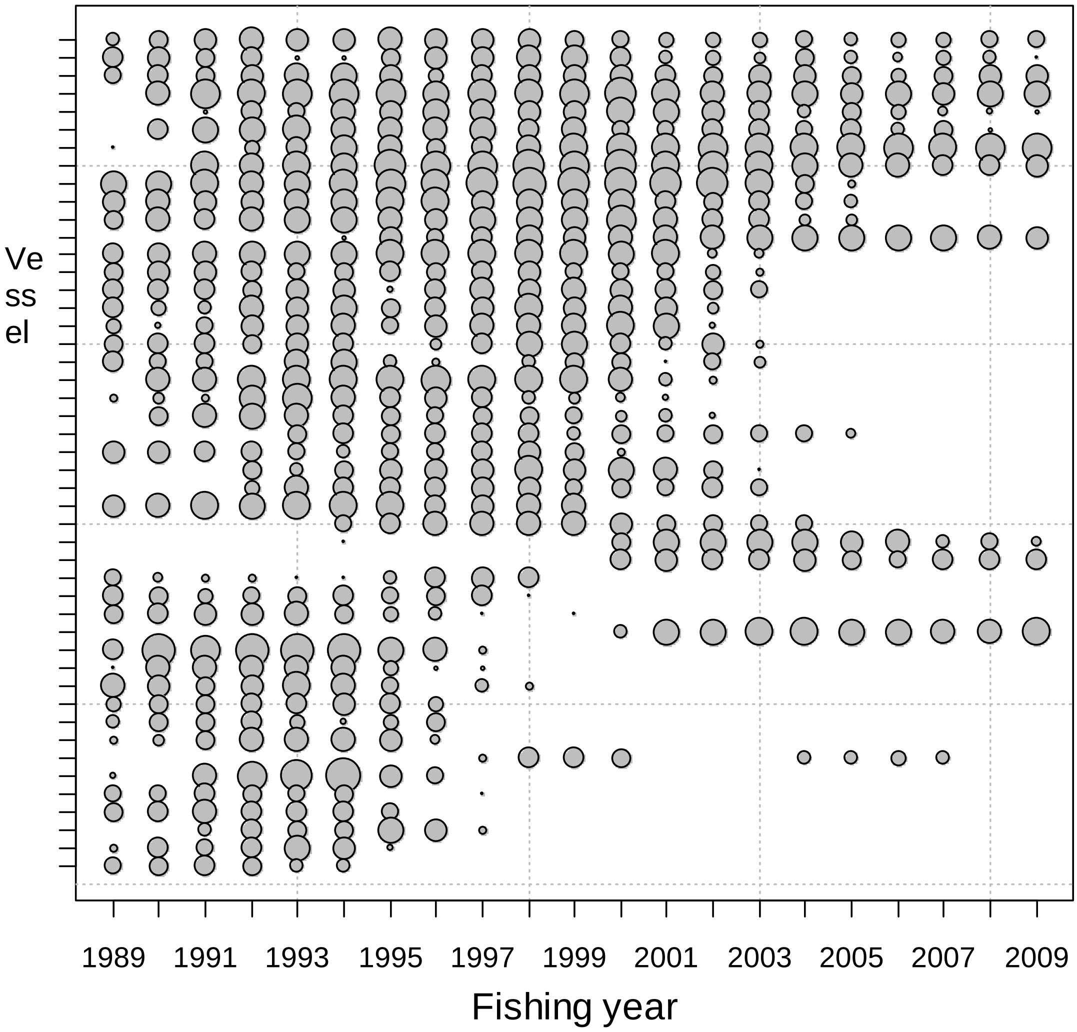


Bay of Plenty longline
5 trips per year
20 trips per year
40 trips per year
Number of vessels
Proportion of catch
Years

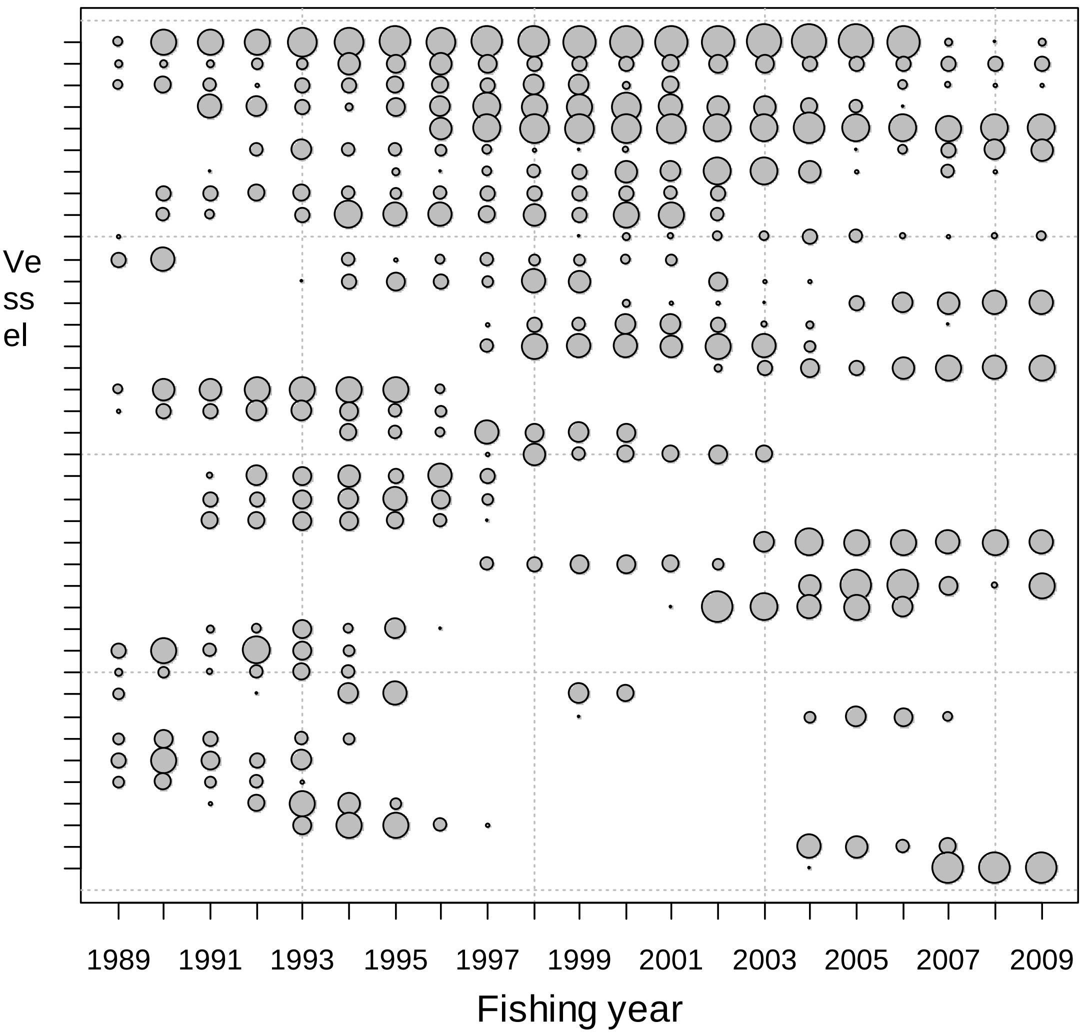
Appendix 14: Data coverage of core vessels selected for each sub-stock/method combination. Circles are proportional to the sum of catch for each vessel in that year.



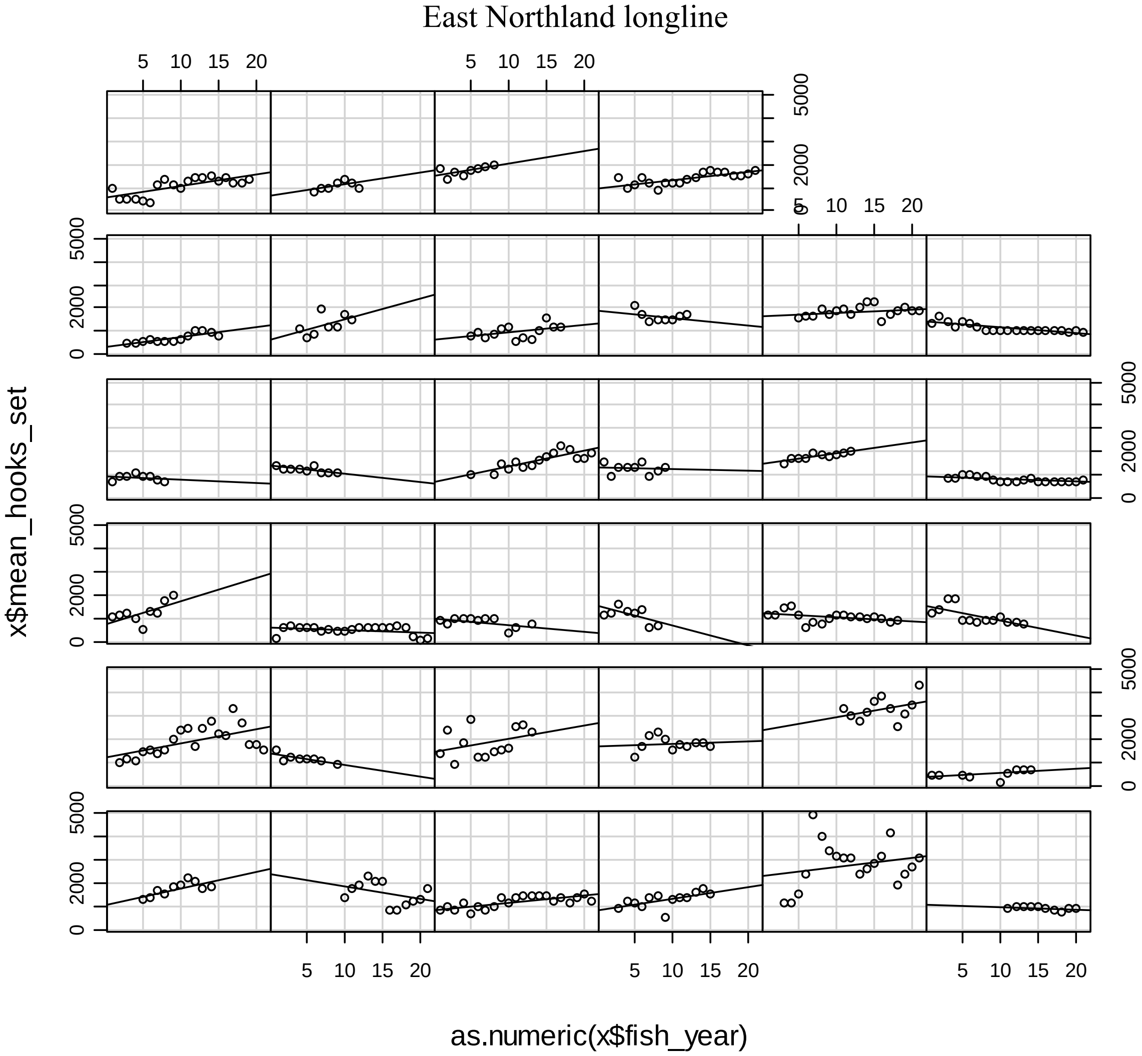
Hauraki Gulf longline

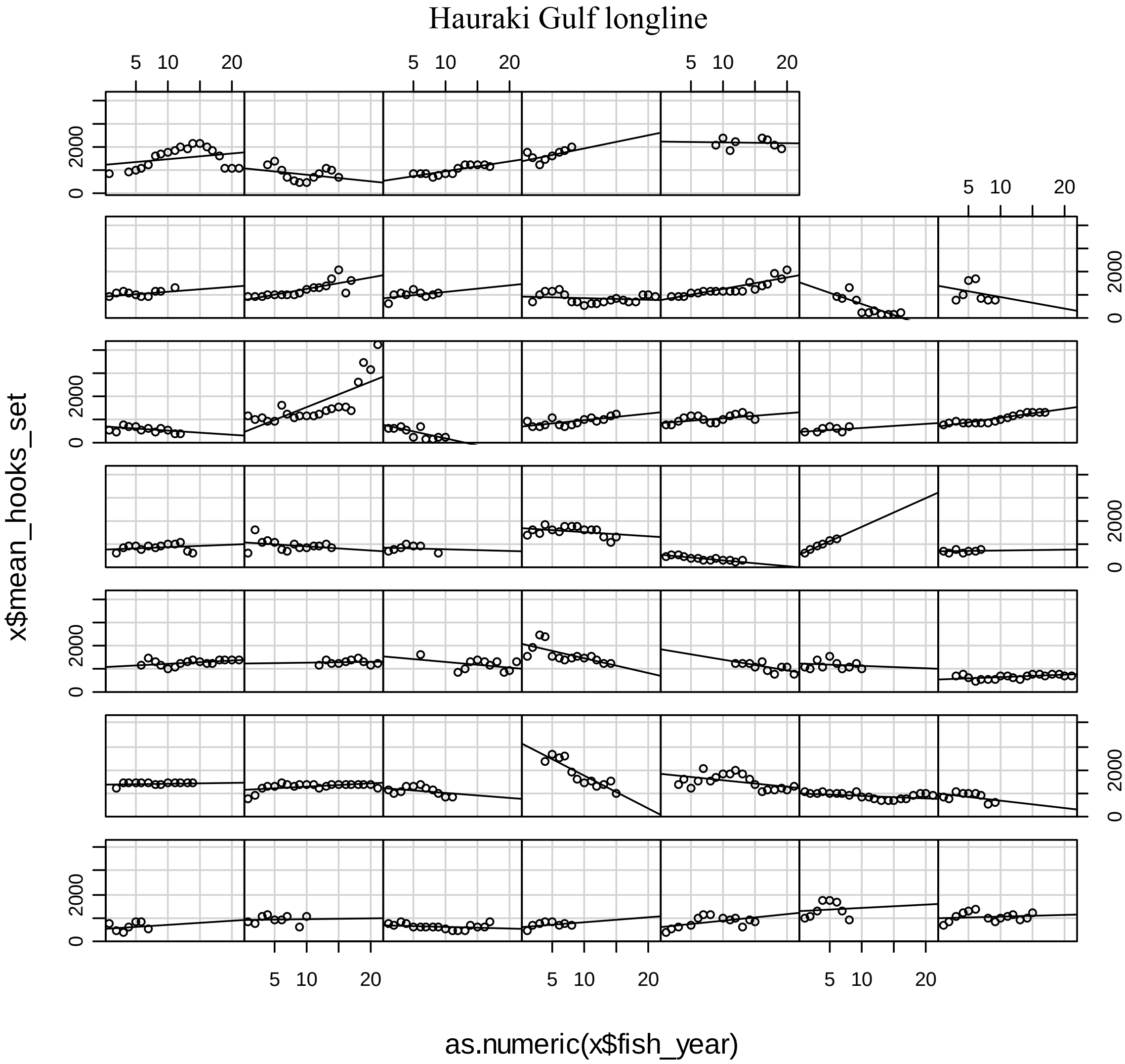


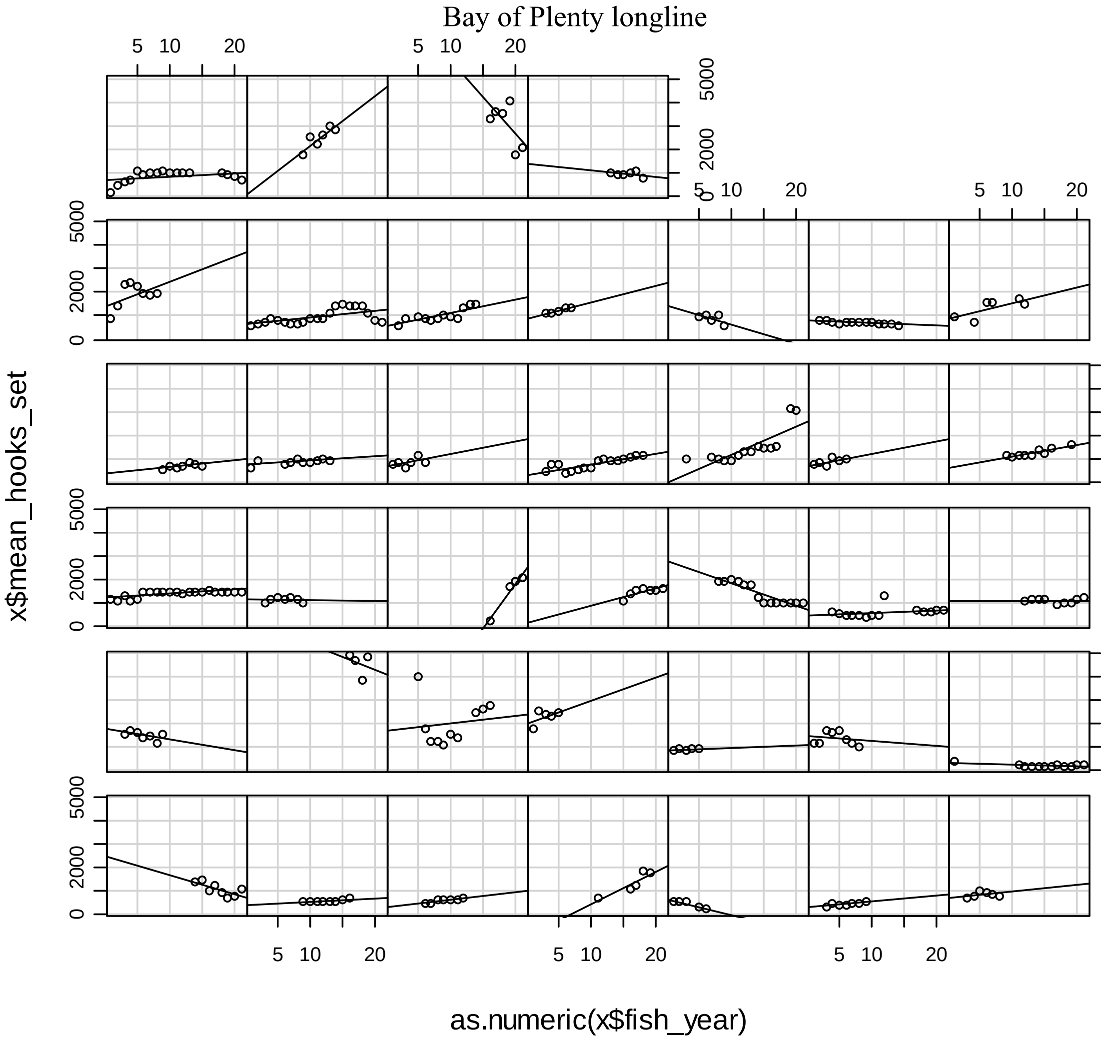
Bay of Plenty longline



Appendix 15: Individual vessel raw effort plots (mean hooks/per set) by fishing year (as numeric 1-21) used to identify discontinuities in effort reporting between form types (last three fishing years in series: 19, 20, 21).



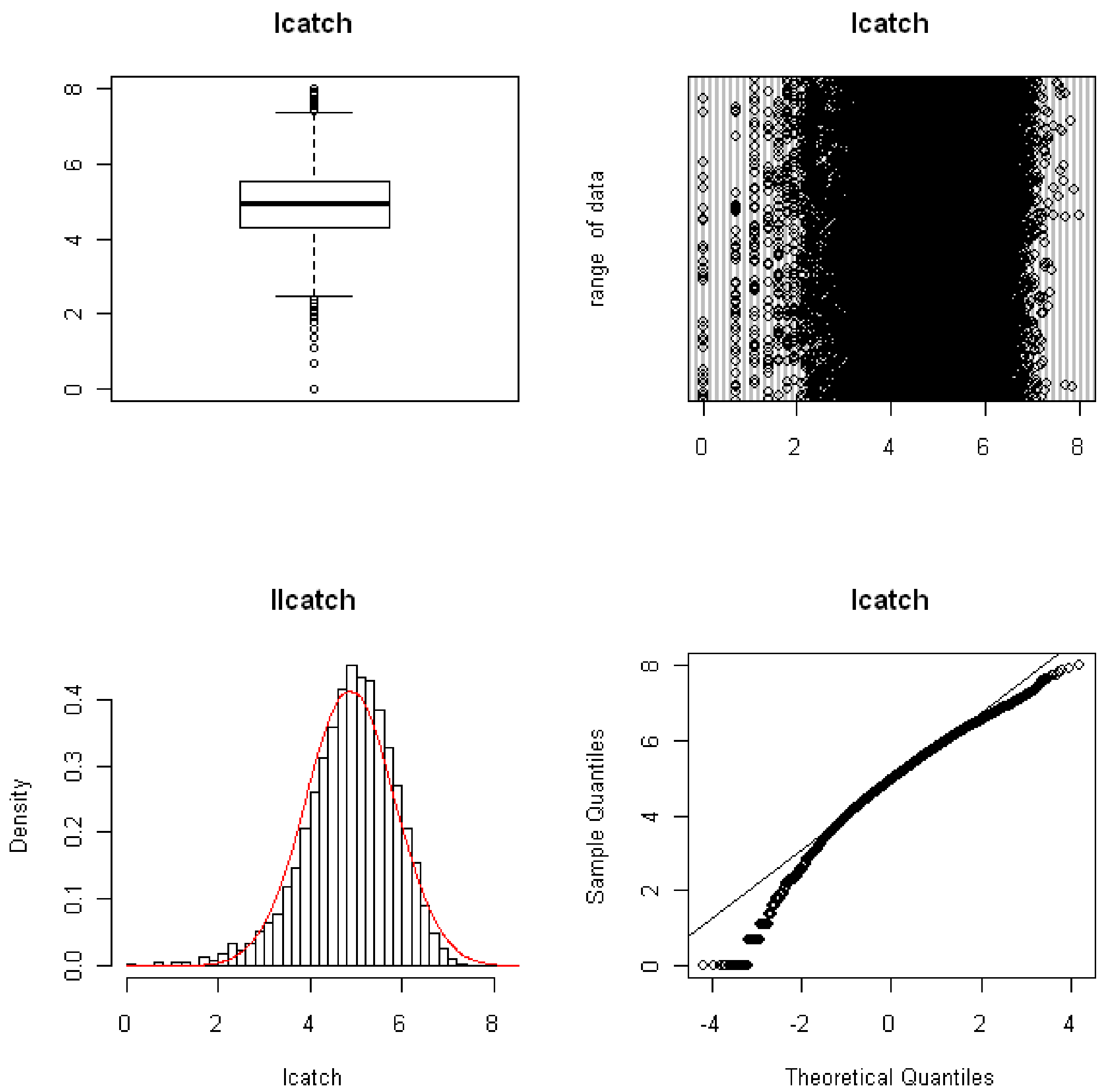

Hauraki Gulf longline
x$mean_hooks_set
as.numeric(x$fish_year)
5
10
20
0
2000

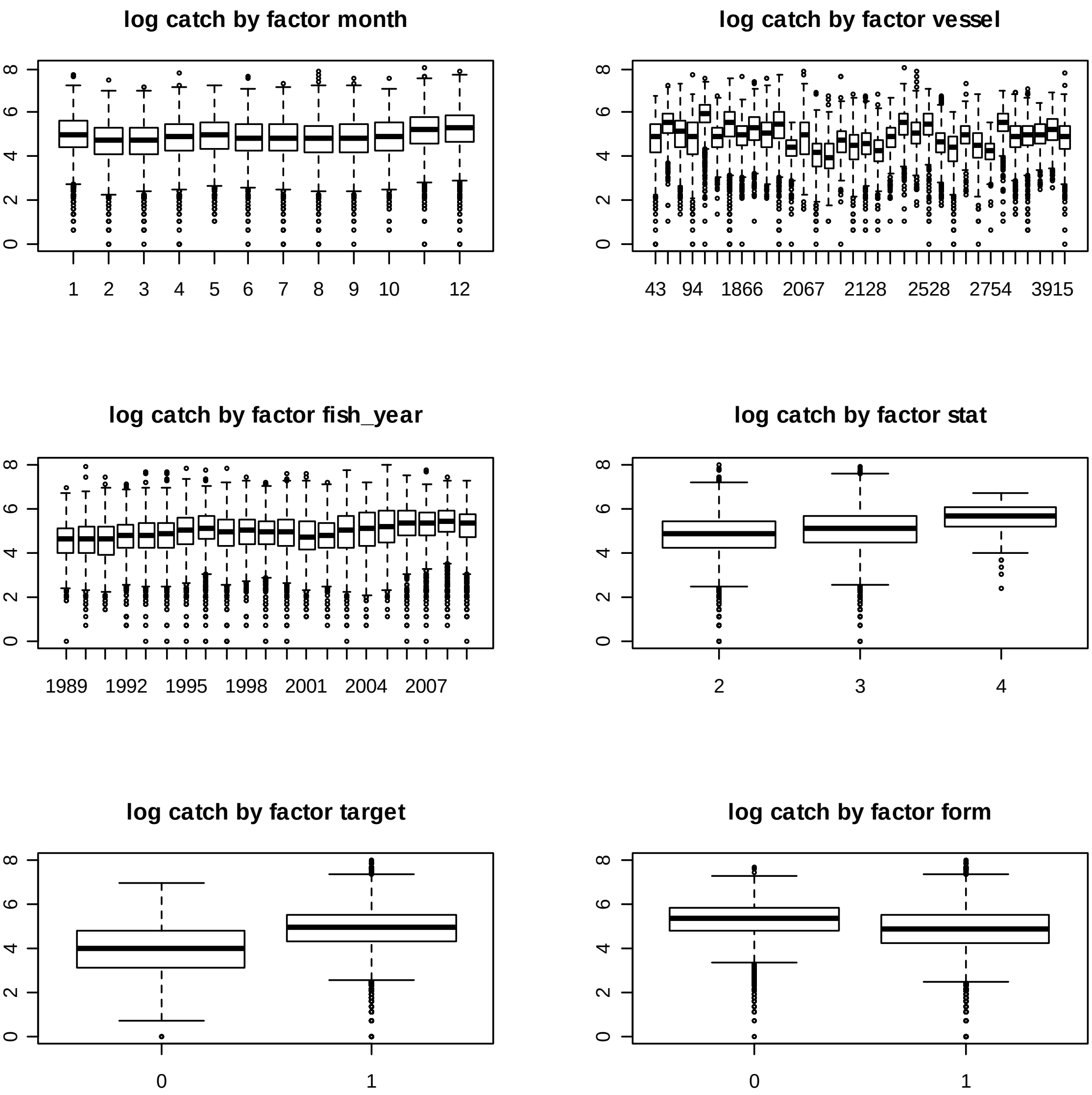


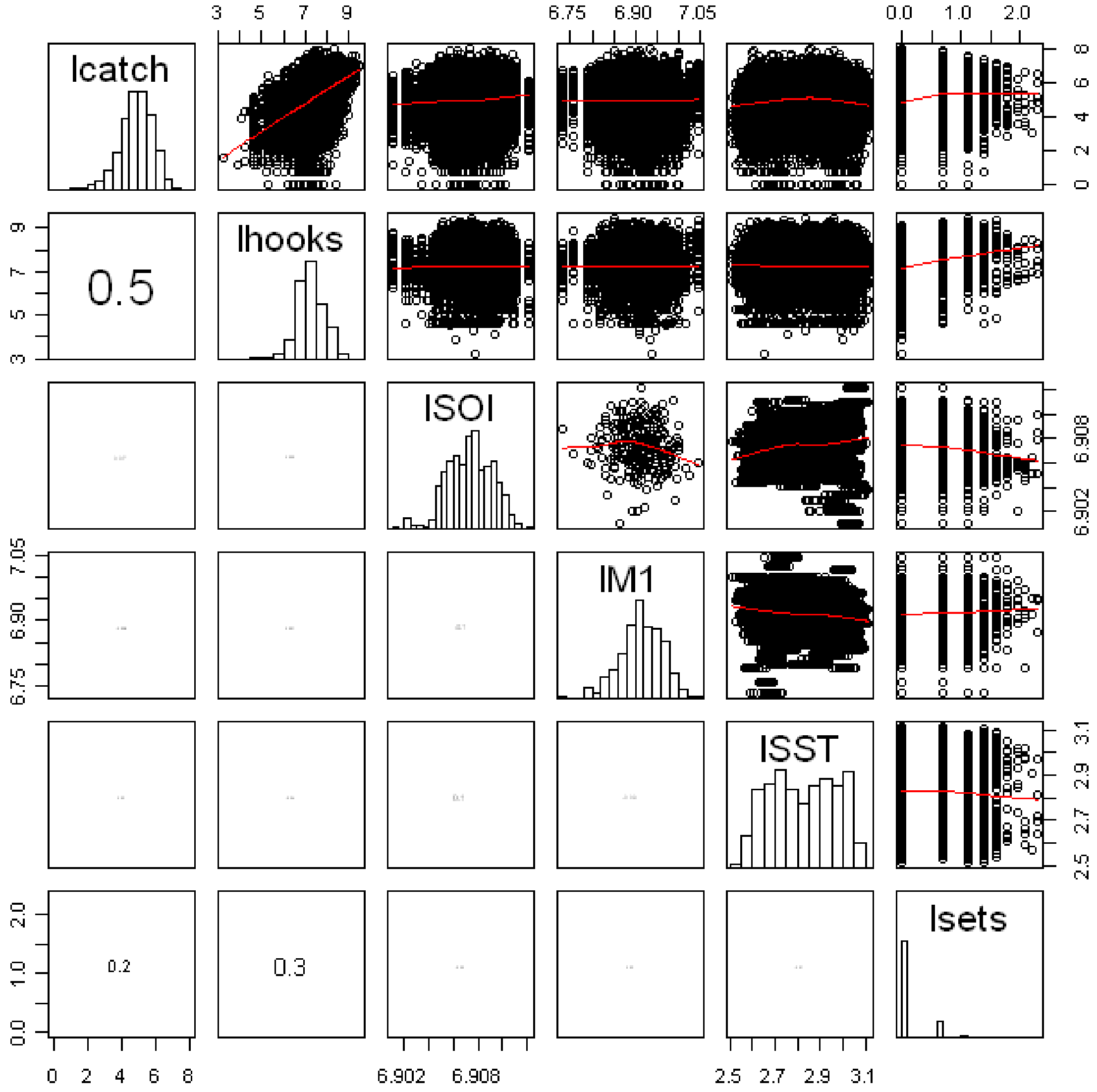
Appendix 16: Raw response variable (log catch) plots investigating normality, homogeneity of variance, collinearity.

The first series of four plots for each area are all relevant to the assumption of normality. Three of these plots illustrate the distribution of the catch data, which should be similar to that of a normal distribution. The plot in the bottom right hand corner is a Q-Q plot, which plots the quantiles of the catch data against those of a theoretical normal distribution (if the catch data were exactly normal the data series would lie along the y = x line indicated on the graph). The second series of plots for each area are all relevant to the assumption of homogeneous variances. In these plots catch is illustrated across the different levels of categorical explanatory variables. A similar spread of catch across all levels would indicate homogeneous variances. The third series of plots for each area illustrate the presence of any correlations between explanatory variables offered to the model. The red trend lines and the size of the correlation coefficients opposite the relevant graphics indicate the strength of any correlations that may be present.

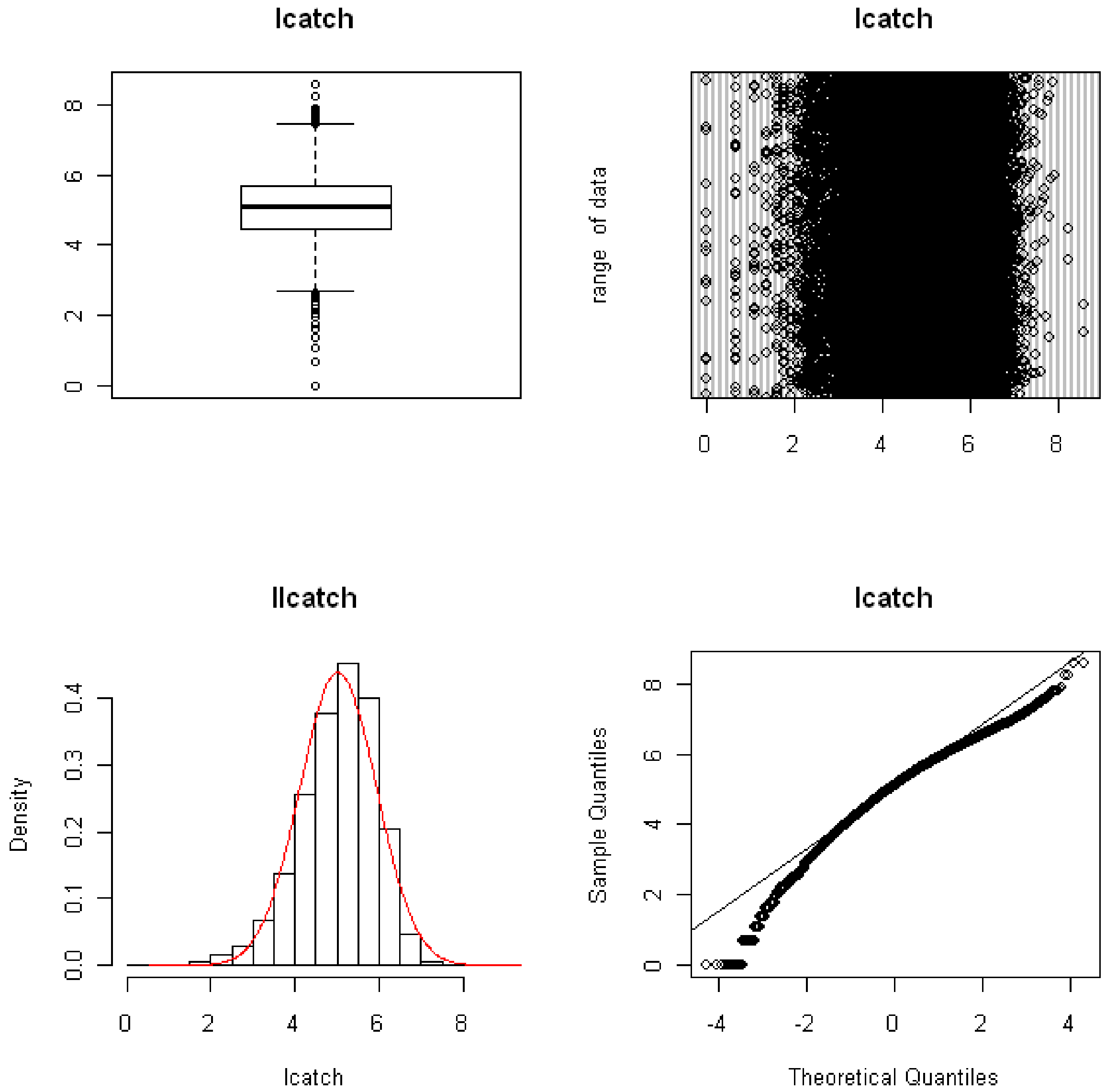
East Northland longline

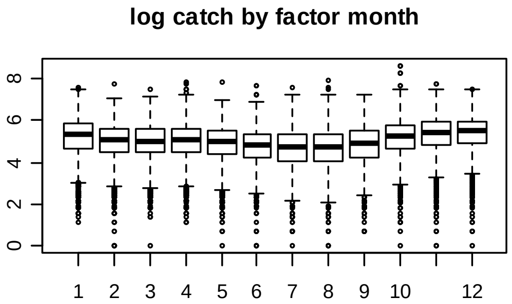


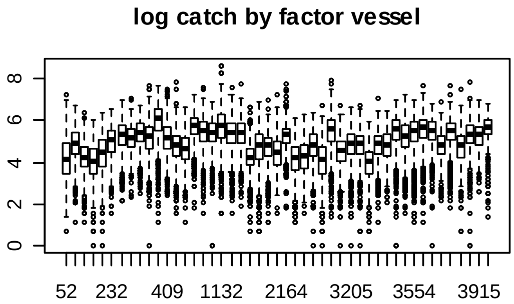

log catch by factor month
0 2 4 6 8
1 2 3 4 5 6 7 8 9 10 12
log catch by factor vessel
0 2 4 6 8
43 94 1866 2067 2128 2528 2754 3915
log catch by factor fish_year
0 2 4 6 8
1989 1992 1995 1998 2001 2004 2007
log catch by factor stat
0 2 4 6 8
2 3 4
log catch by factor target
0 2 4 6 8
0 1
log catch by factor form
0 2 4 6 8
0 1

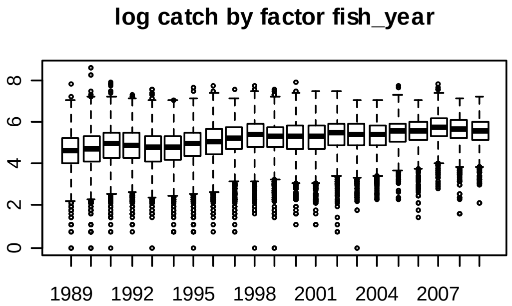

lcatch
lhooks
lSOI
lM1
lSST
lsets
0.5
0.2
0.3

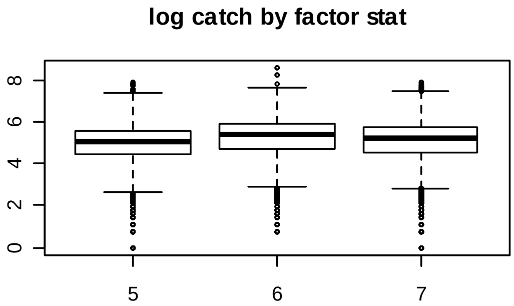
Hauraki Gulf longline

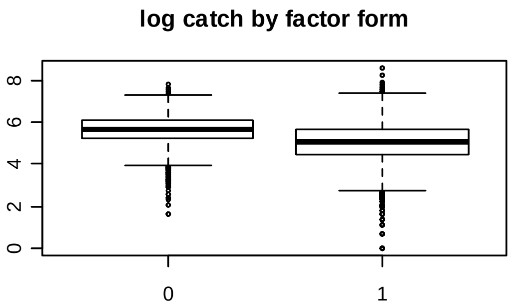


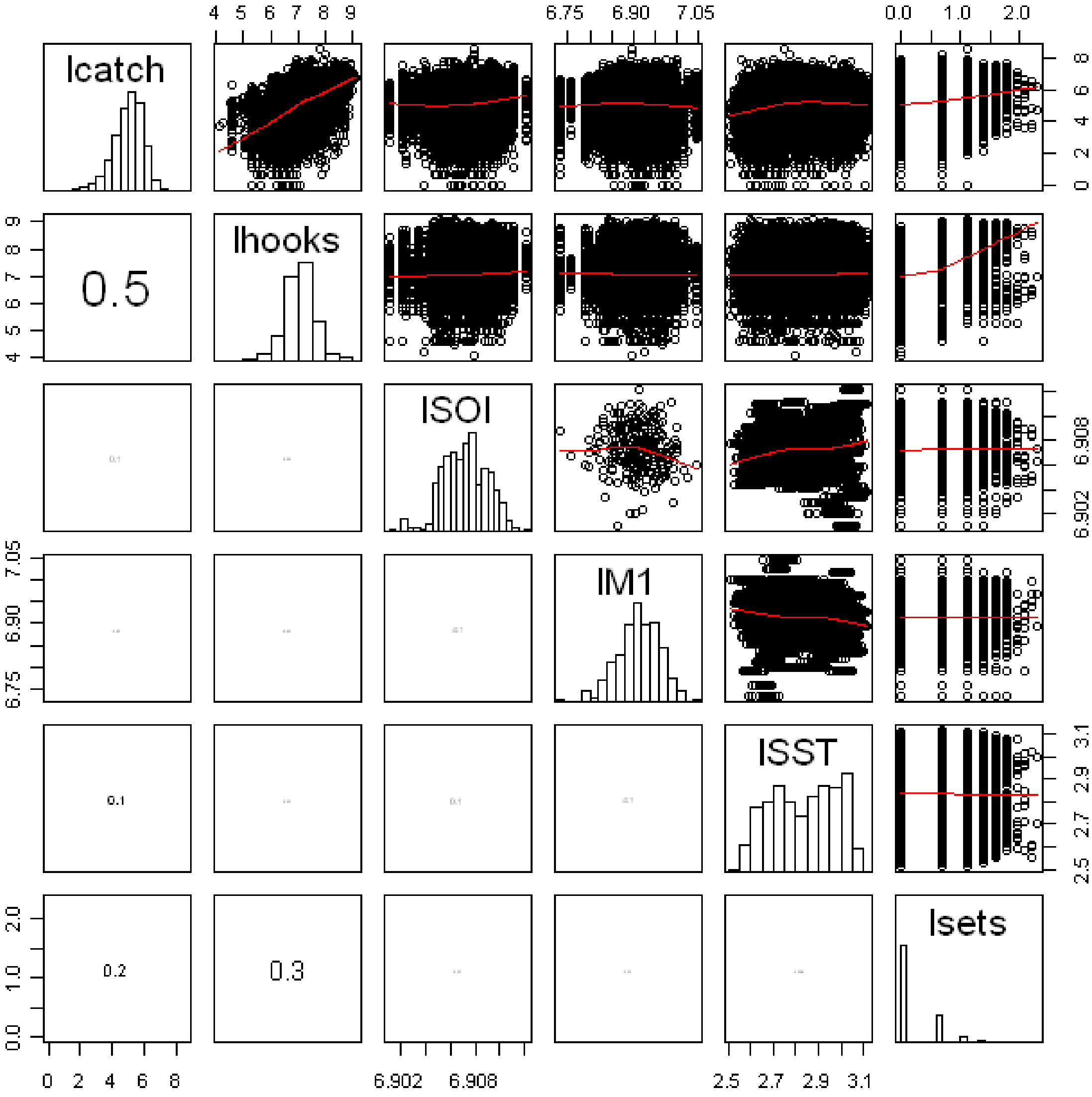

log catch by factor month
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 12


log catch by factor vessel
0
2
4
6
8
52 232 409 1132 2164 3205 3554 3915

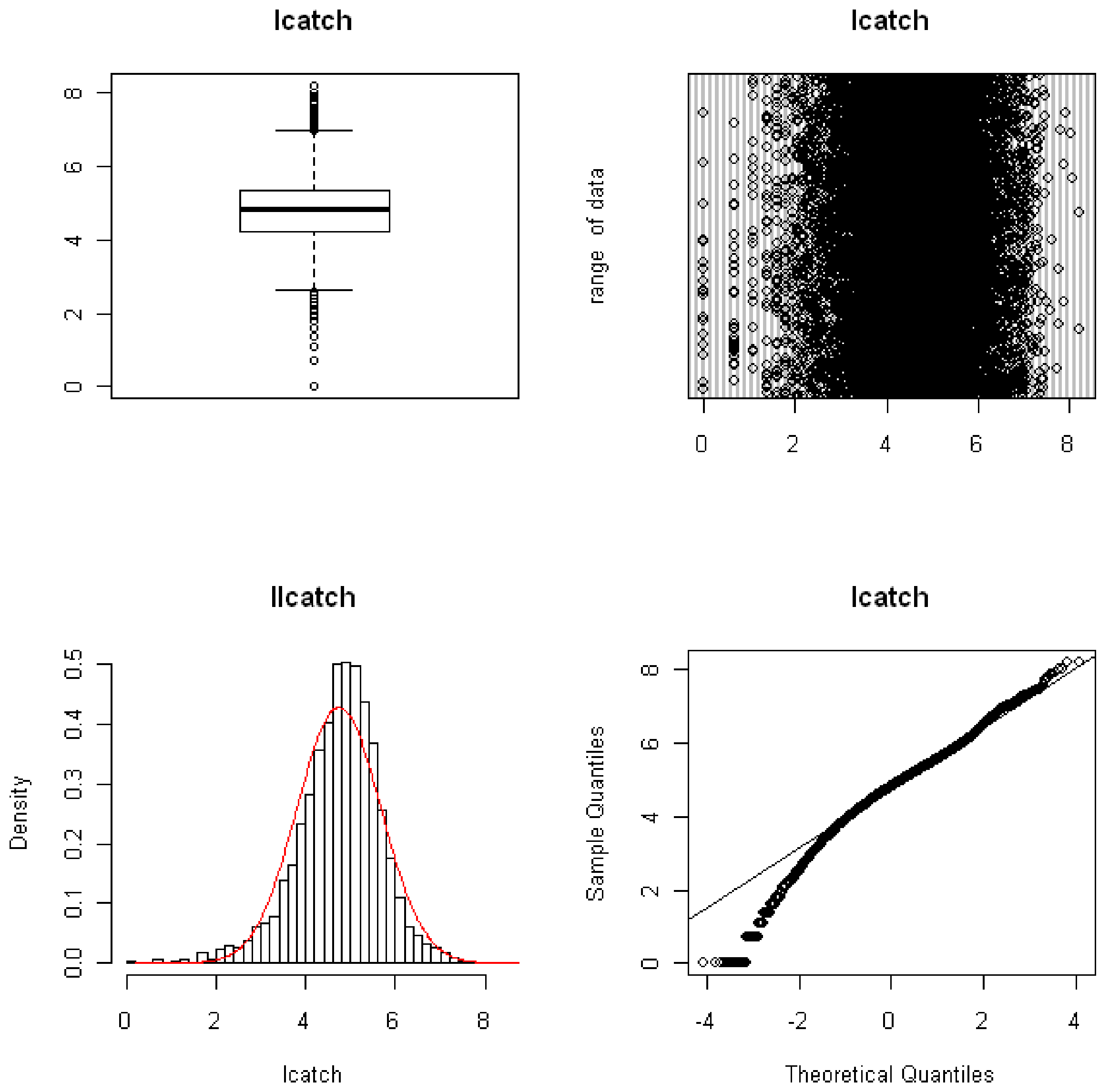

log catch by factor fish_year
0
2
4
6
8
1989 1992 1995 1998 2001 2004 2007

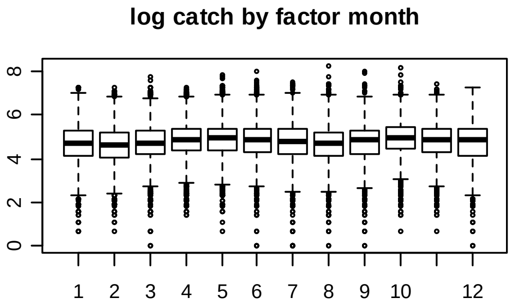

log catch by factor stat
0
2
4
6
8
5 6 7

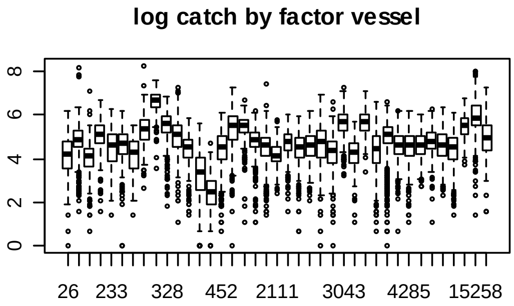

log catch by factor form
0
2
4
6
8
0 1

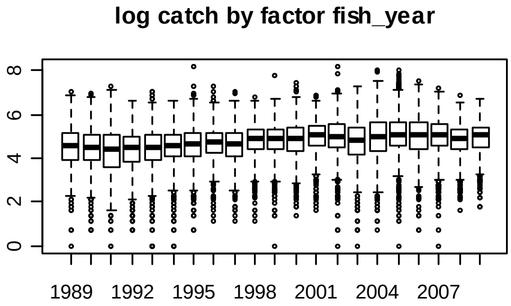

lcatch
lhooks
lSOI
lM1
lSST
lsets
0.5
0.1
0.2
0.3
4 5 6 7 8 9
6.75 6.90 7.05
0.0 1.0 2.0
0 2 4 6 8
6.902 6.908
2.5 2.7 2.9 3.1

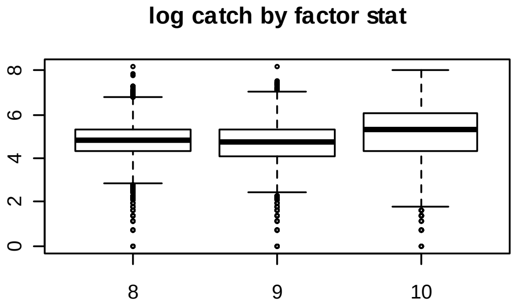
Bay of Plenty longline

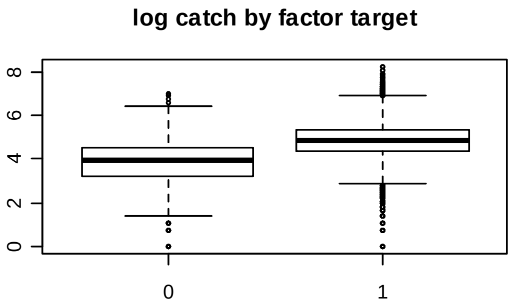


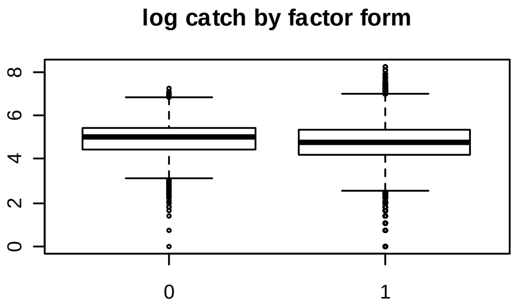

log catch by factor month
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 12

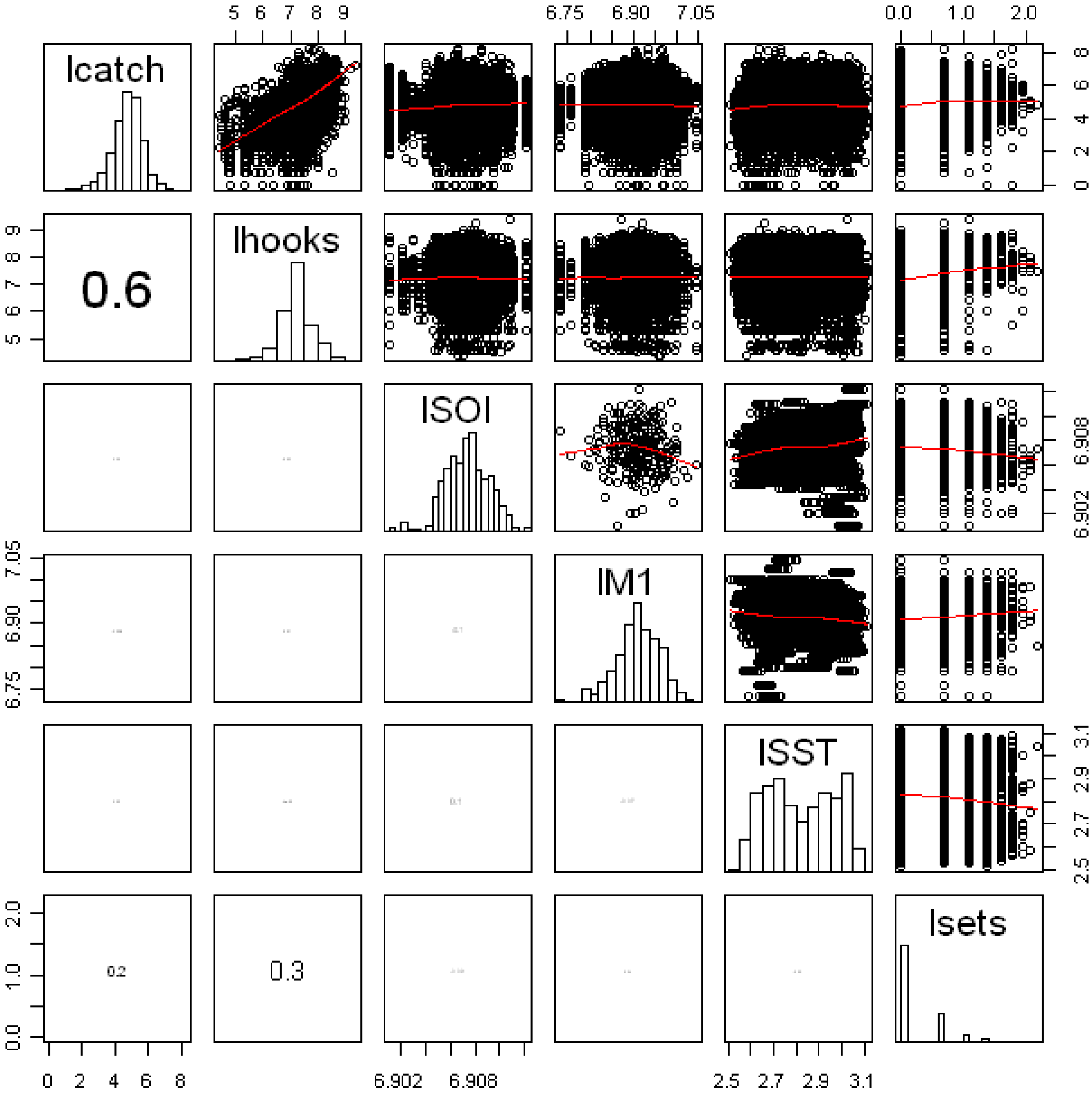

log catch by factor vessel
0
2
4
6
8
26 233 328 452 2111 3043 4285 15258


log catch by factor fish_year
0
2
4
6
8
1989 1992 1995 1998 2001 2004 2007


log catch by factor stat
0
2
4
6
8
8 9 10


log catch by factor target
0
2
4
6
8
0 1

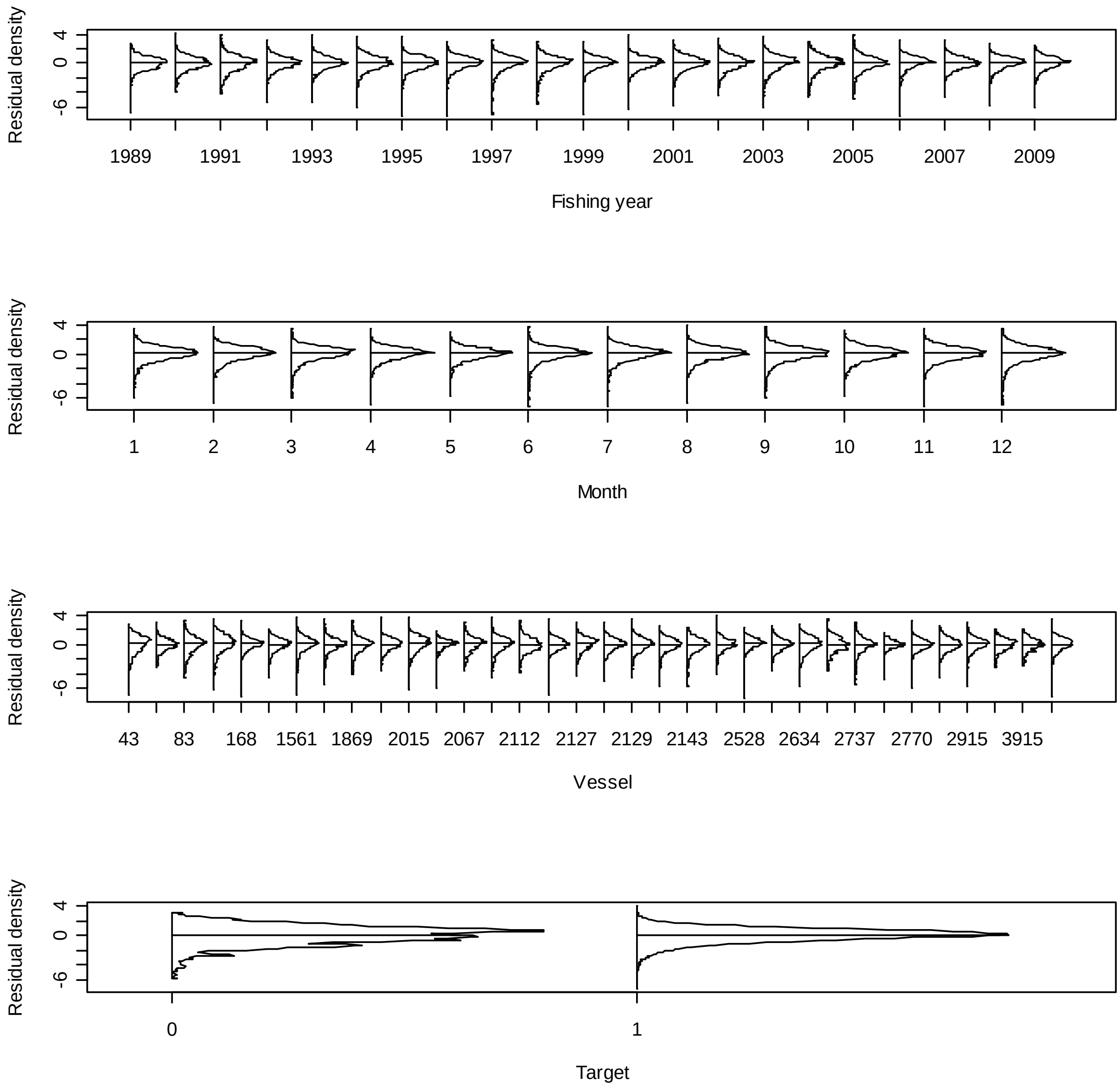

log catch by factor form
0
2
4
6
8
0 1

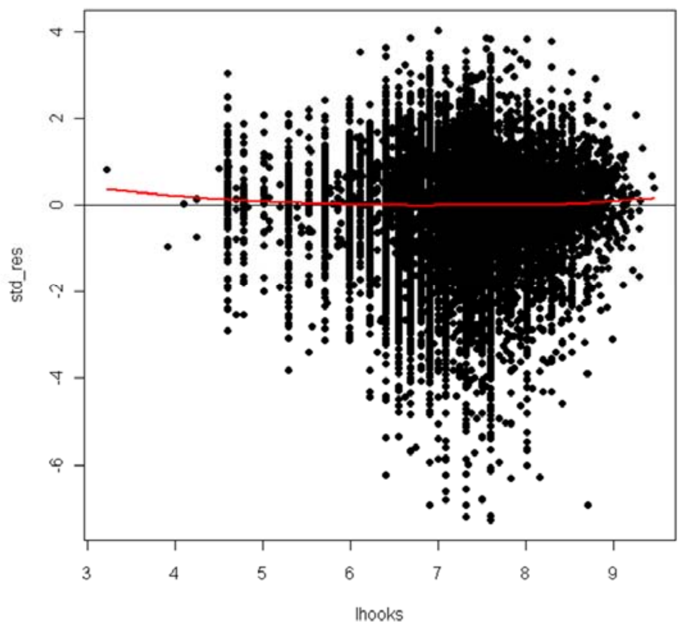

lcatch
lhooks
lSOI
lM1
lSST
lsets
0.6
0.2
0.3

Appendix 17: Residual plots from the final regression model again testing for violations in the homogeneity of variance assumption from the fitted model.

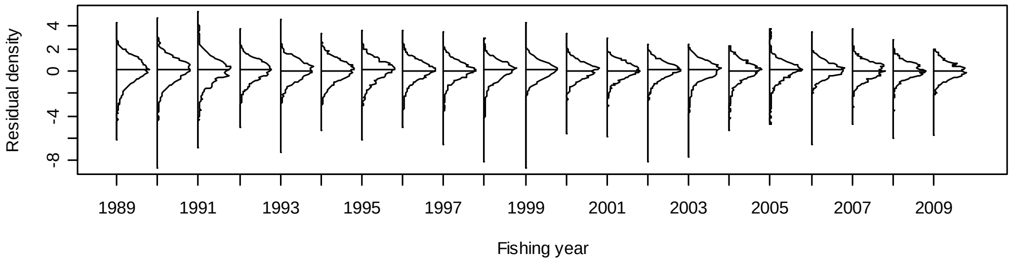
Probability density distributions of model residuals across the different levels of each categorical variable selected by the model (first series of graphs for each area) and the distribution of model residuals in the date order that fishing events occurred in (second graph for each area).

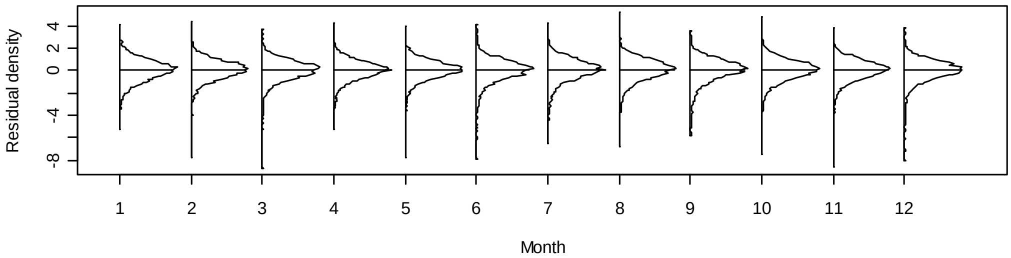
East Northland longline

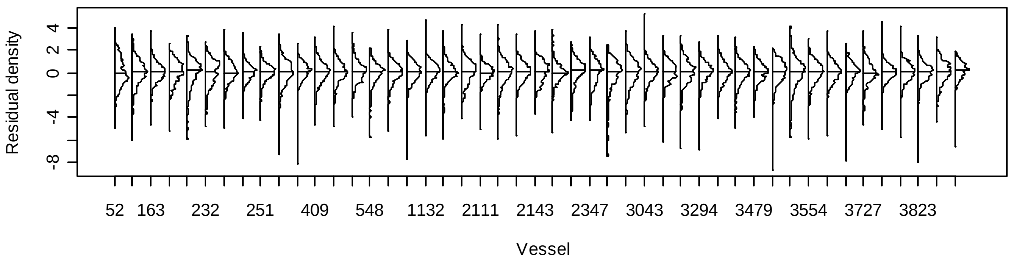


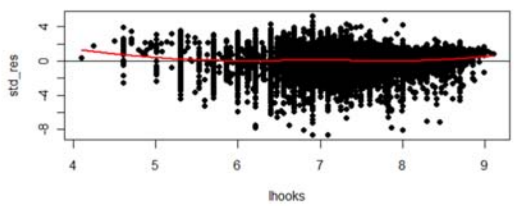

std_res
lhooks

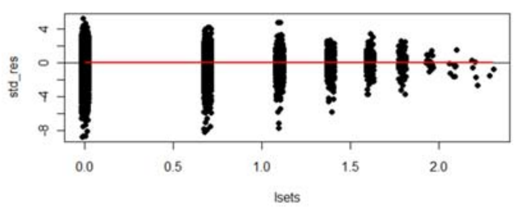
Hauraki Gulf longline



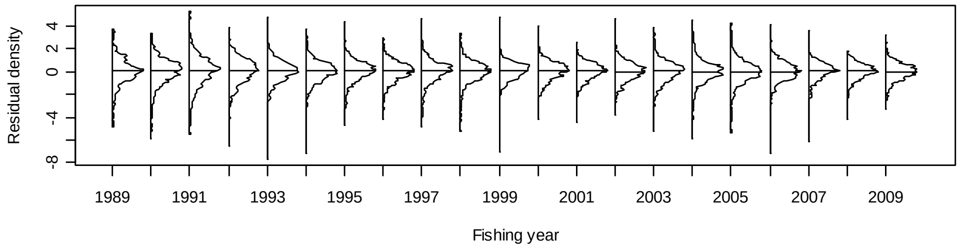


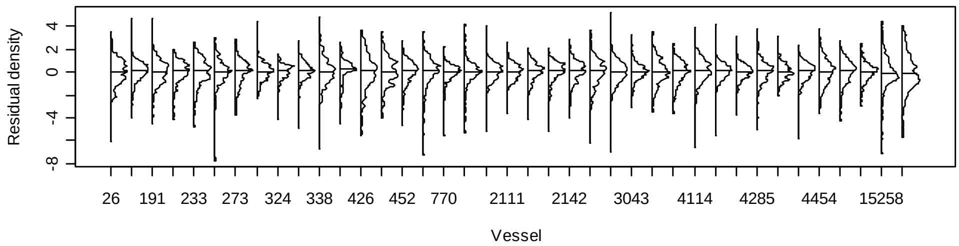


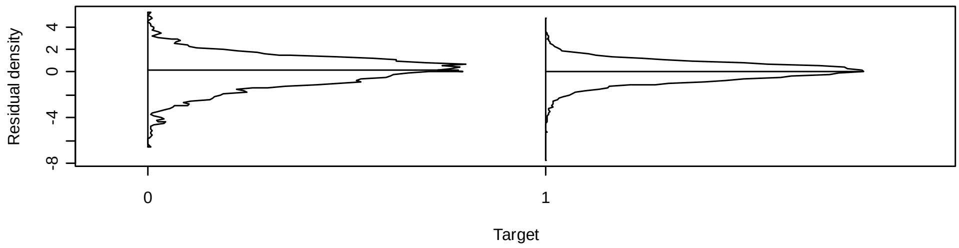


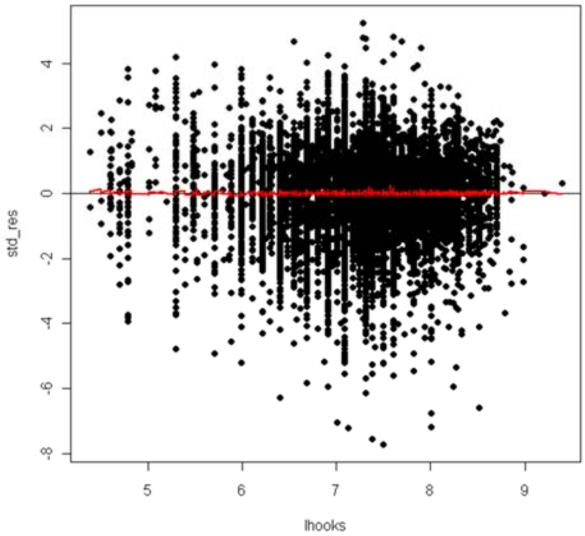


Bay of Plenty longline



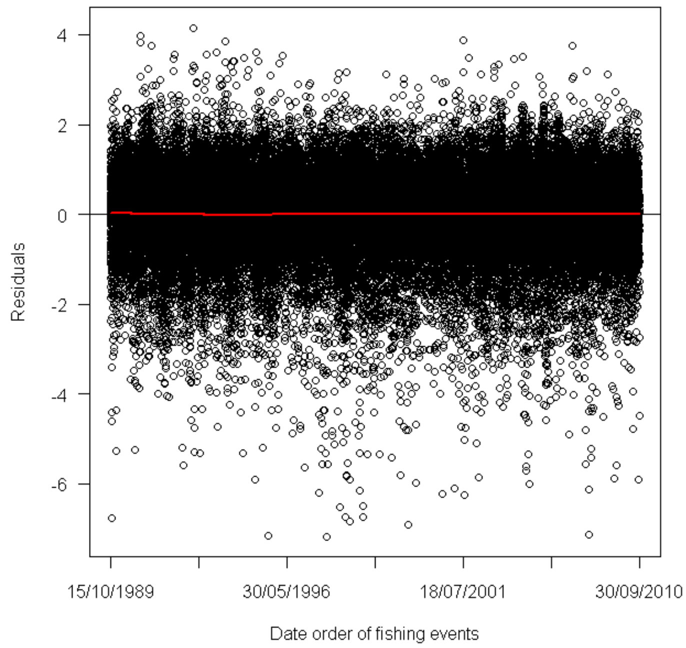




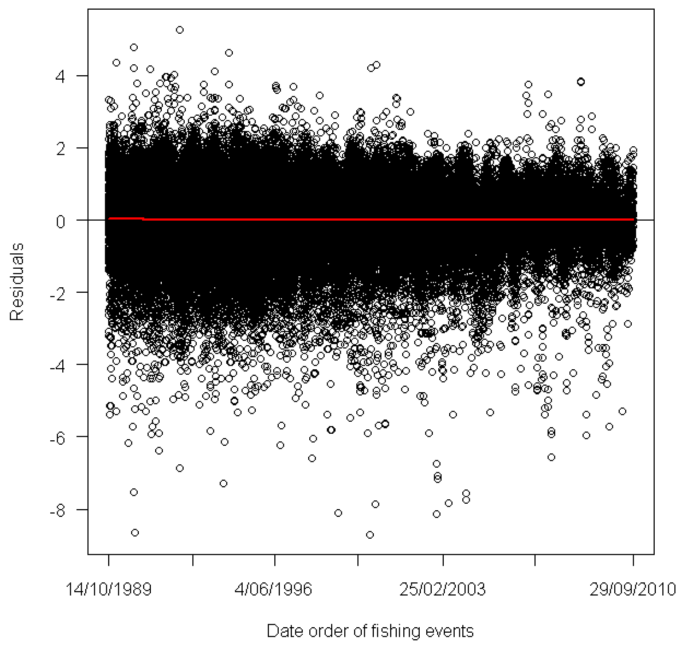


Appendix 18: Residual plots from the final regression model testing for violations in the independence of the response data observations.

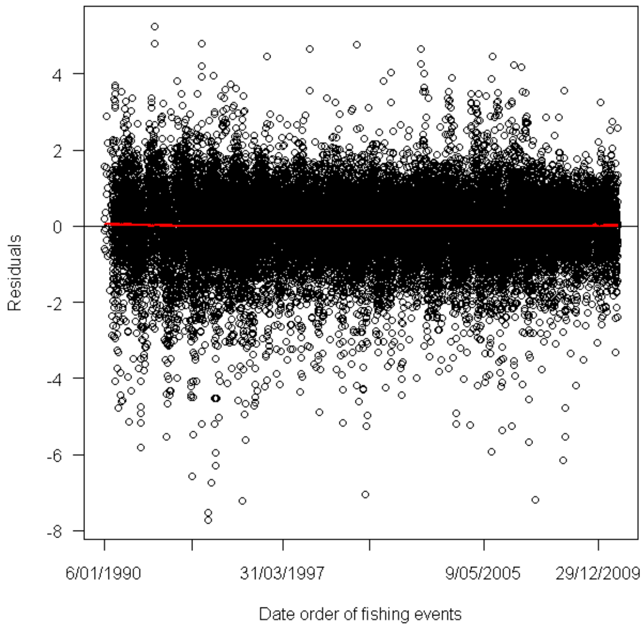
East Northland longline



Hauraki Gulf longline



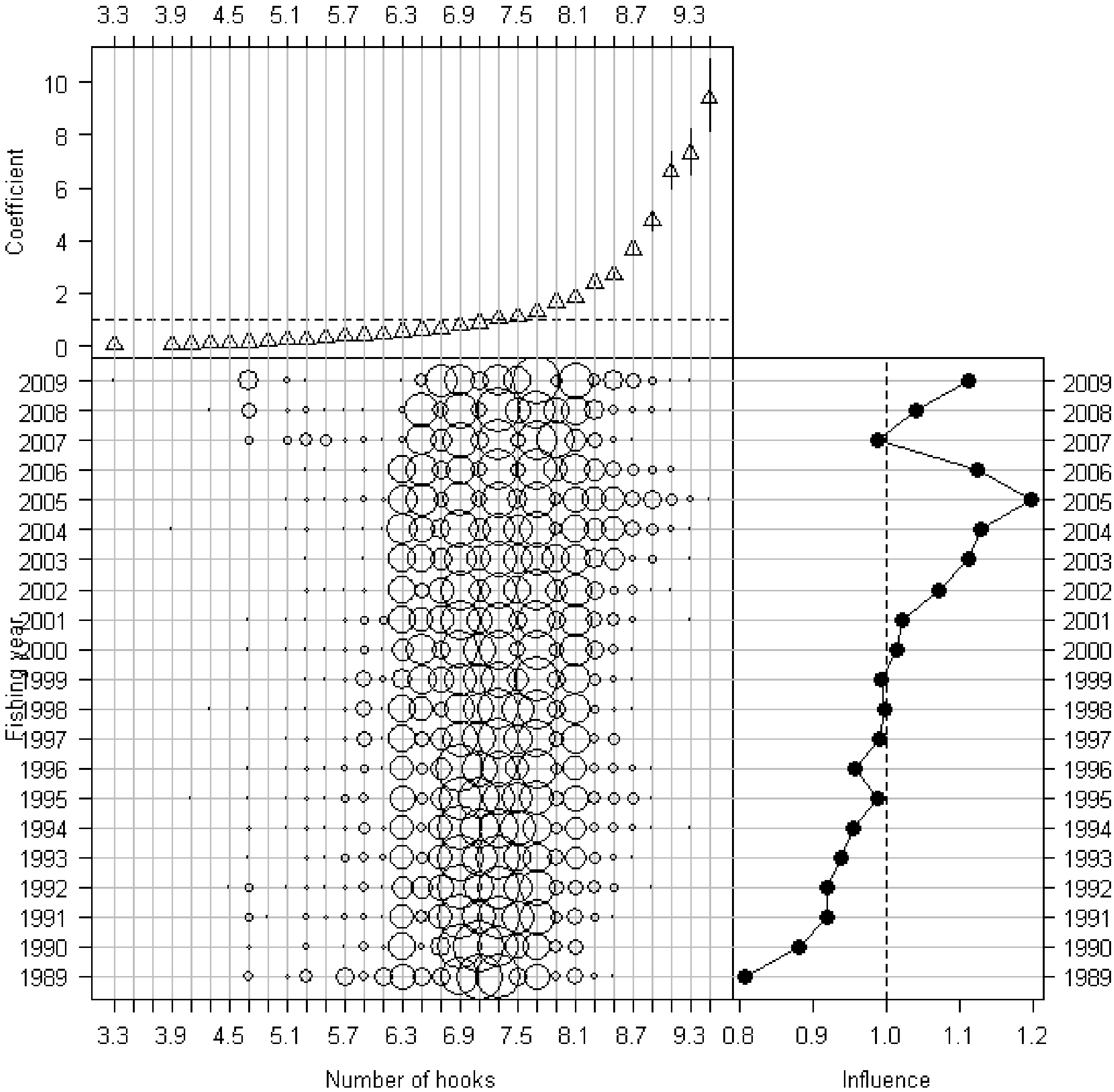
Bay of Plenty Longline

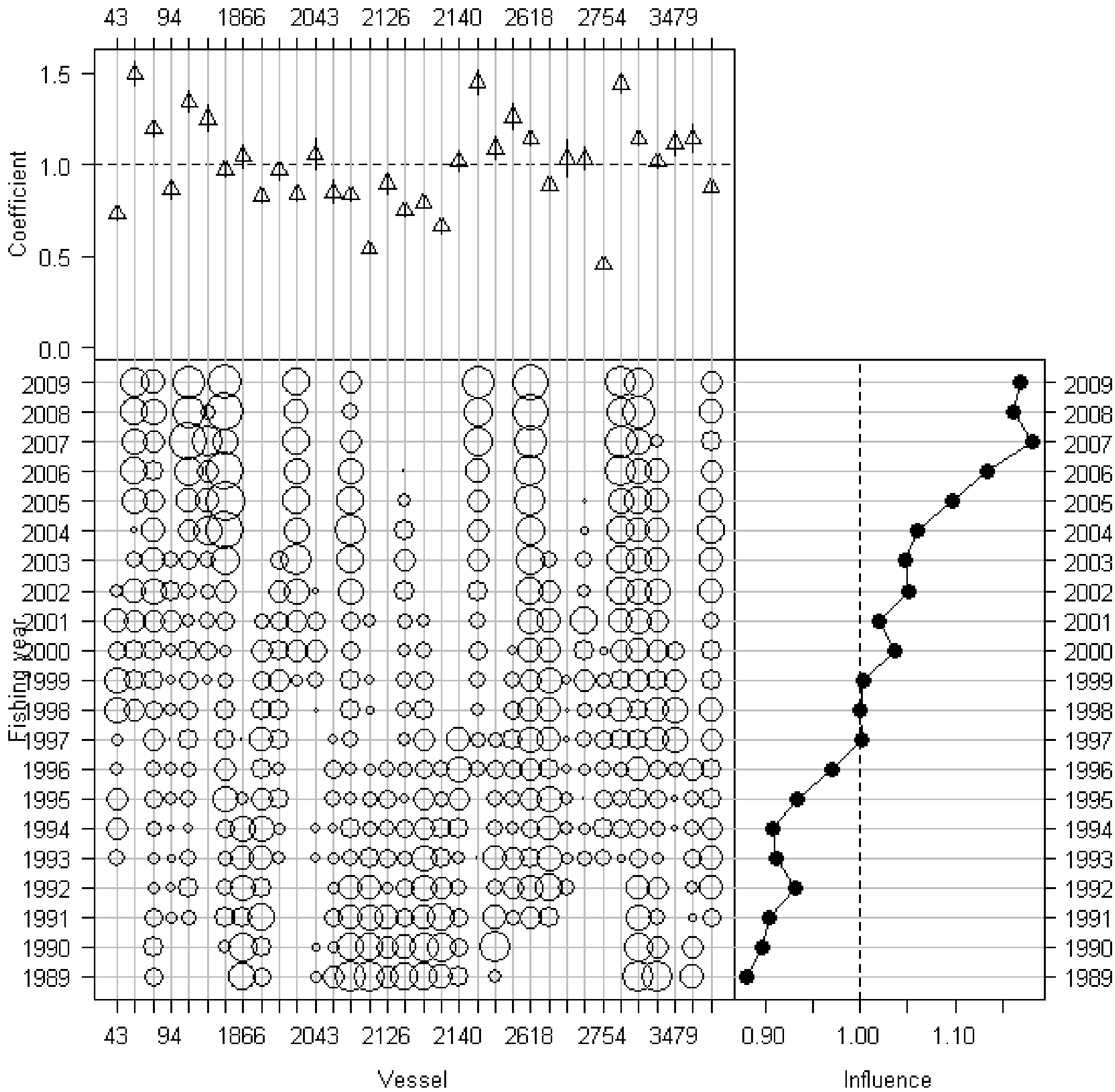


Appendix 19: Plots illustrating the influence of all significant variables from regression analyses for each sub-stock/method combination (after Bentley et al. in press).

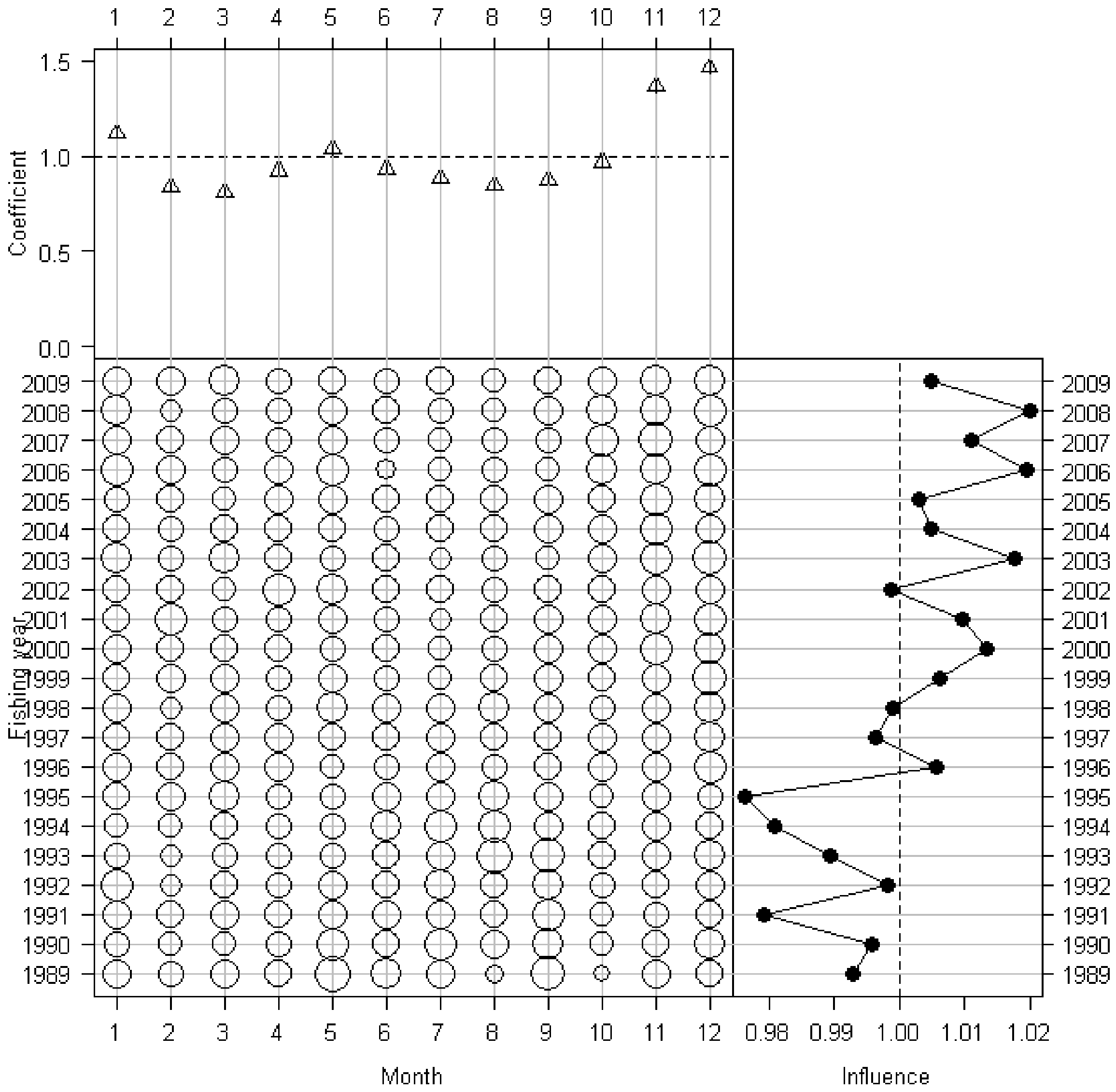
The top panel of each plot represents the relationship between the variable of interest and catch. The bottom left panel represents the distribution of the number of events for the variable of interest in each year of the time series (where target is listed as either 1 = "SNA" or 0 = "other"). The bottom right panel represents the proportional influence (with no influence represented at 1) a variable had on the CPUE index for each year of the time series. Continuous variables are plotted on a log scale.

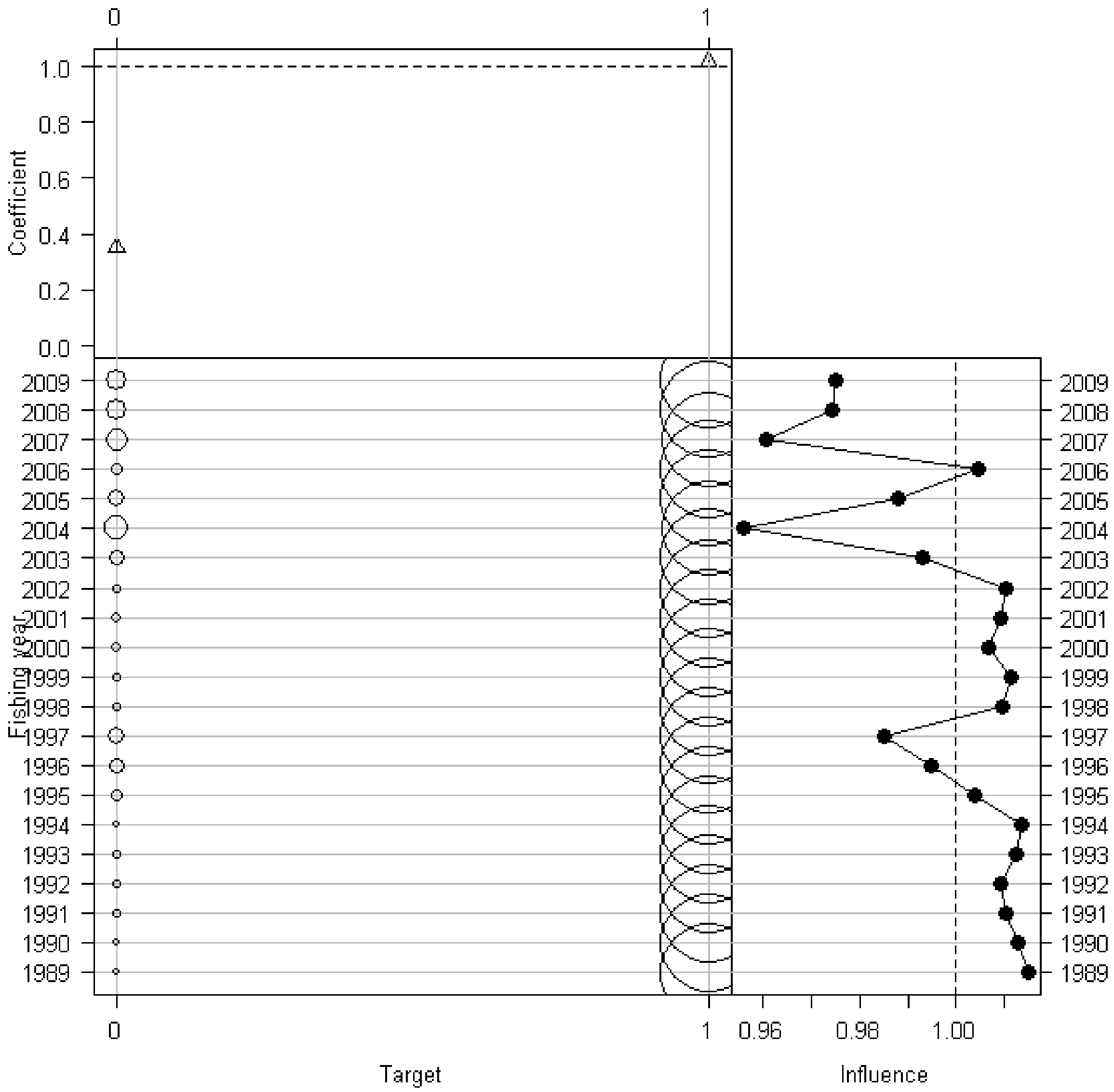
East Northland longline



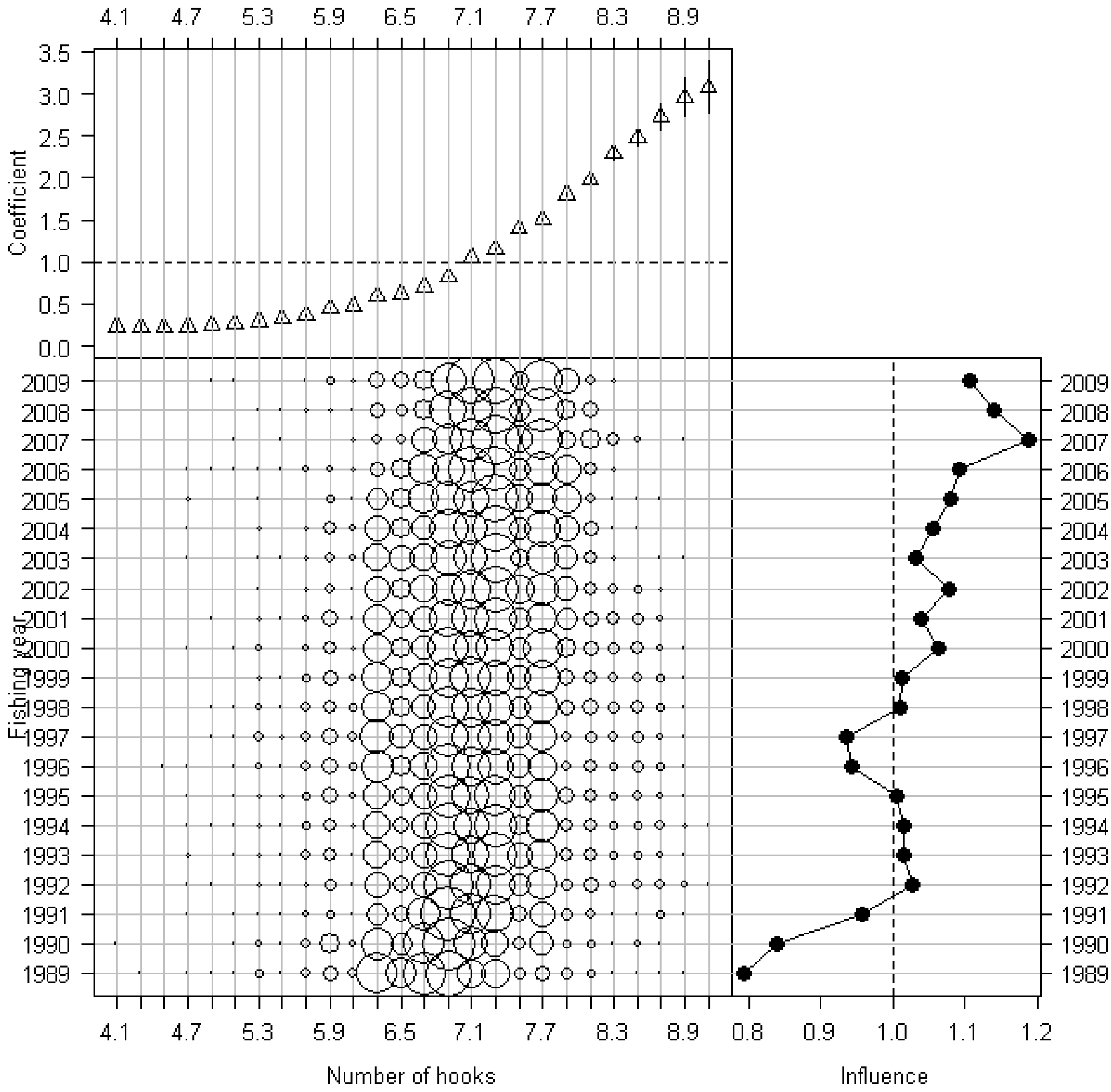


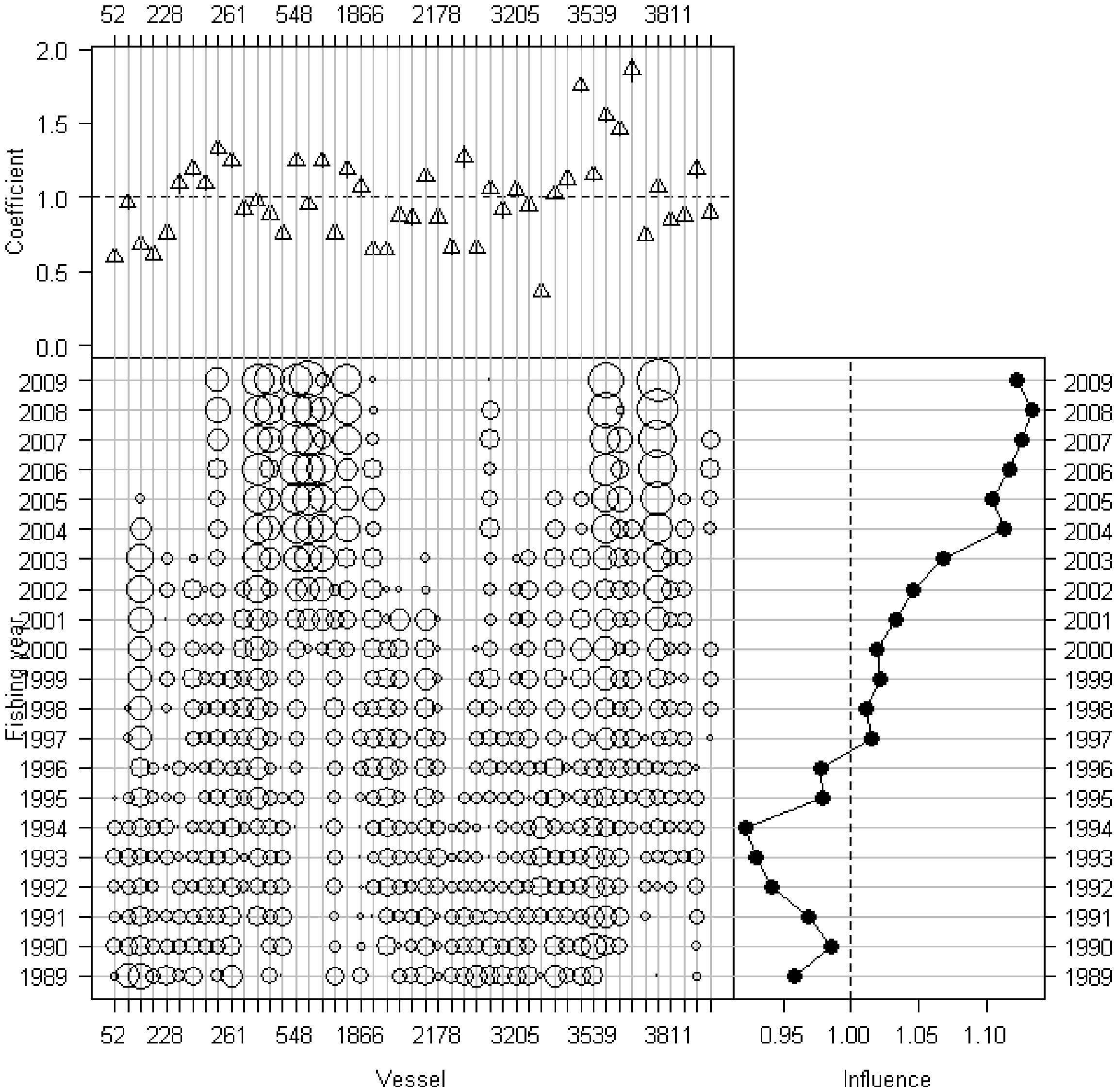
Coefficient
Fishing year
Vessel
Influence
43
94
1866
2043
2126
2140
2618
2754
3479
0.90
1.00
1.10


Coefficient
Fishing year
Month
Influence
1 2 3 4 5 6 7 8 9 10 11 12
0.0 0.5 1.0 1.5
2009 2008 2007 2006 2005 2004 2003 2002 2001 2000 1999 1998 1997 1996 1995 1994 1993 1992 1991 1990 1989
0.98 0.99 1.00 1.01 1.02

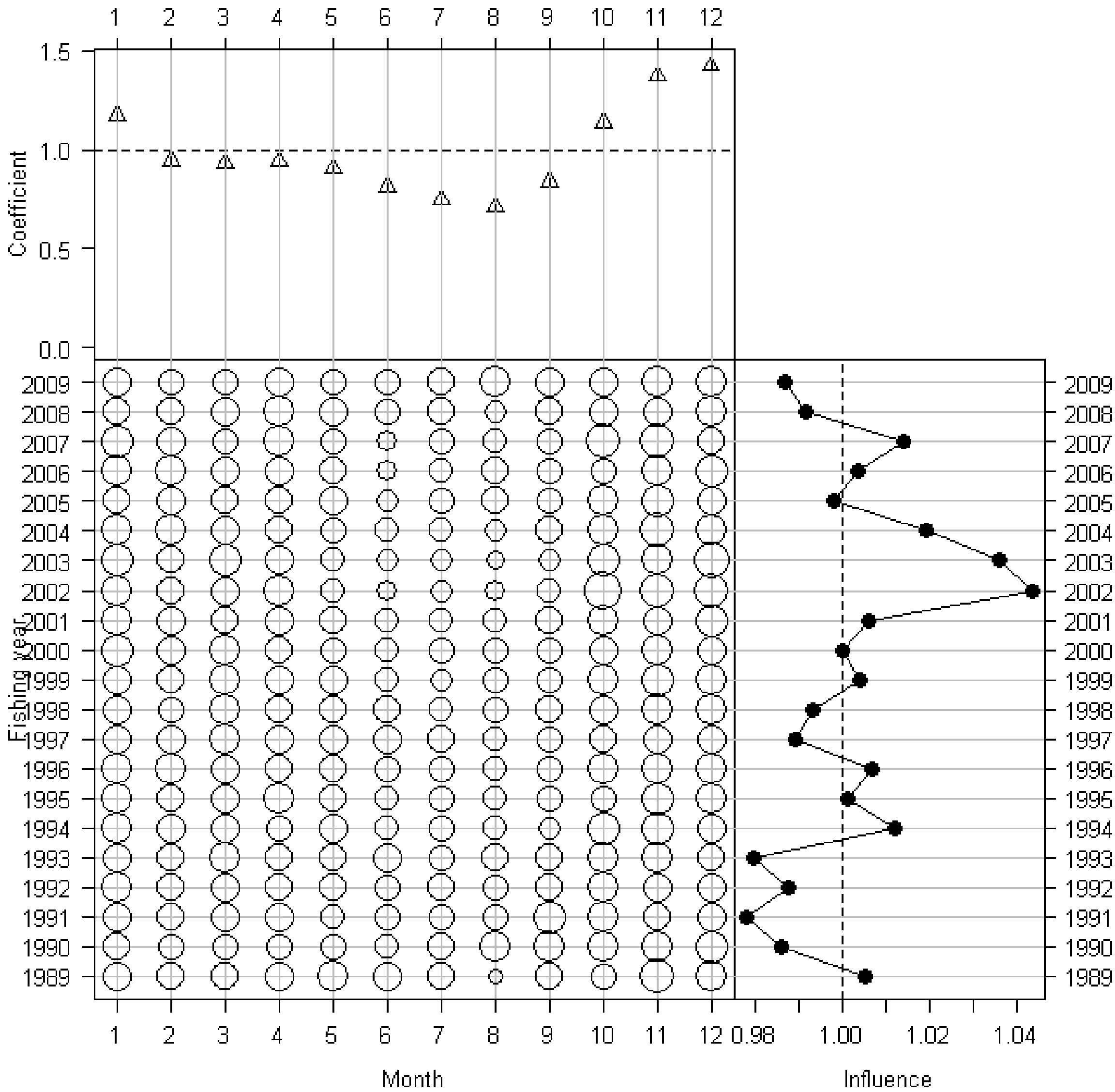

0
1
1.0
0.8
0.6
0.4
0.2
0.0
Coefficient
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
Fishing year
0
1
0.96
0.98
1.00
Target
Influence

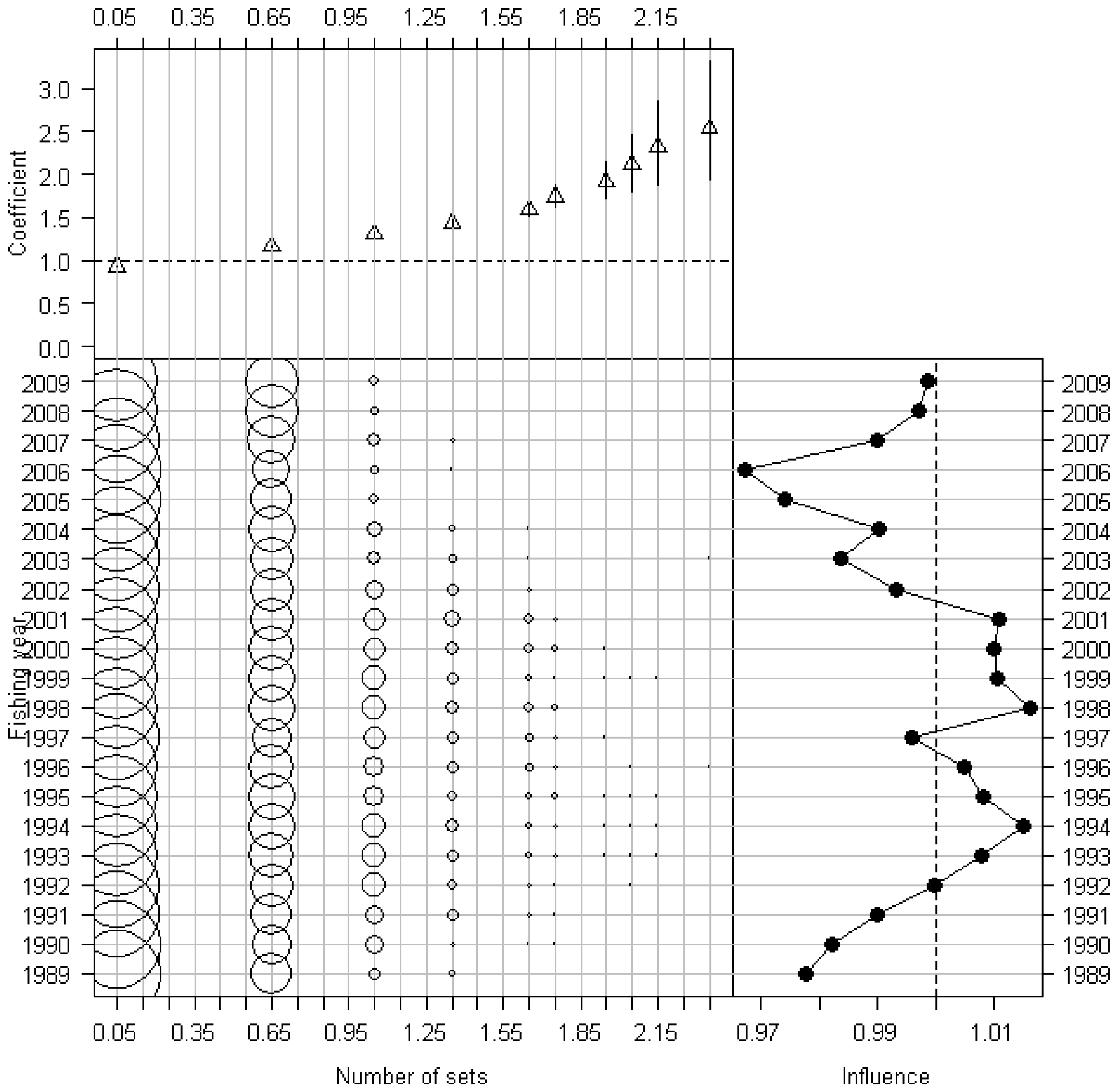
Hauraki Gulf longline

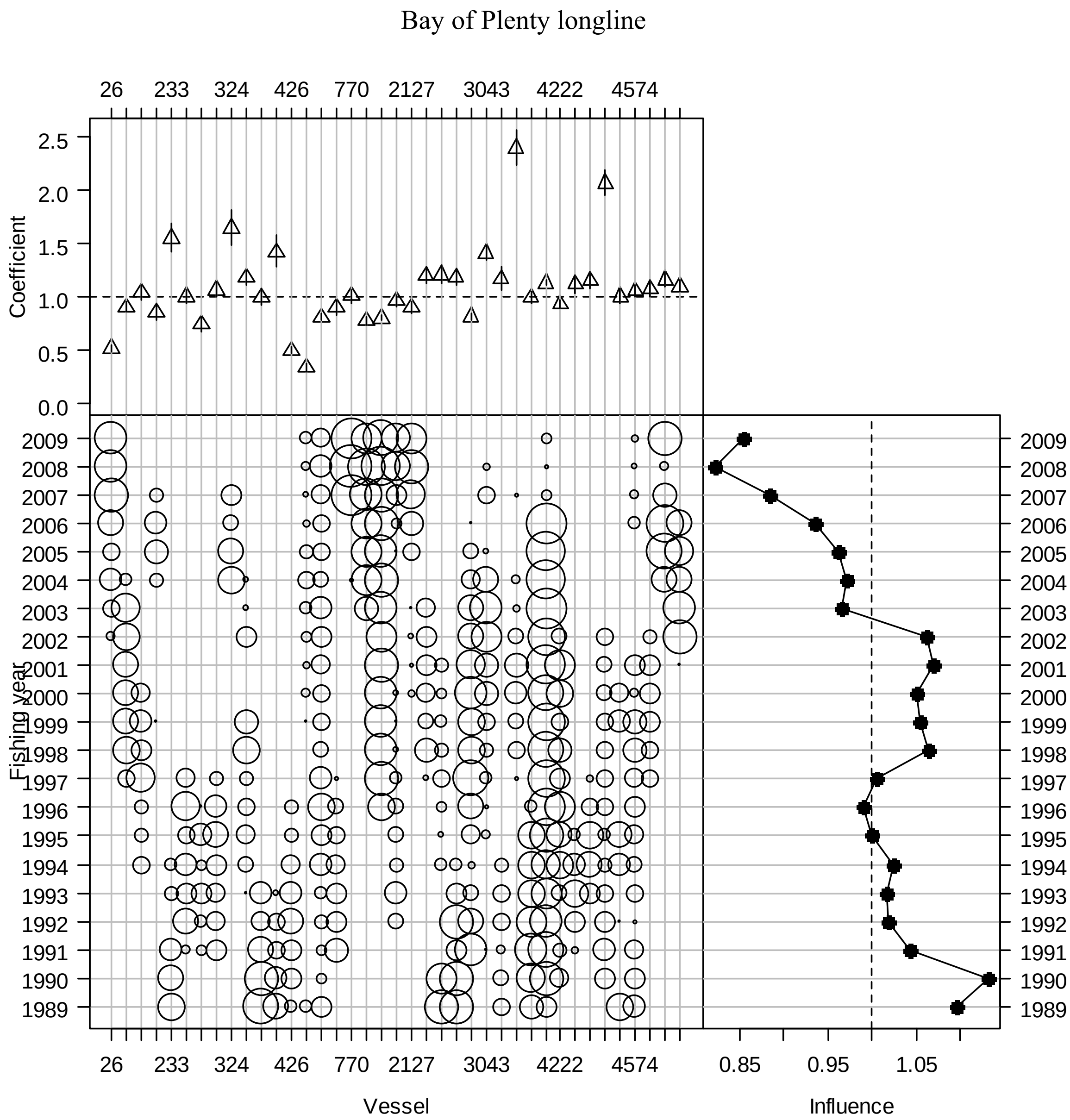




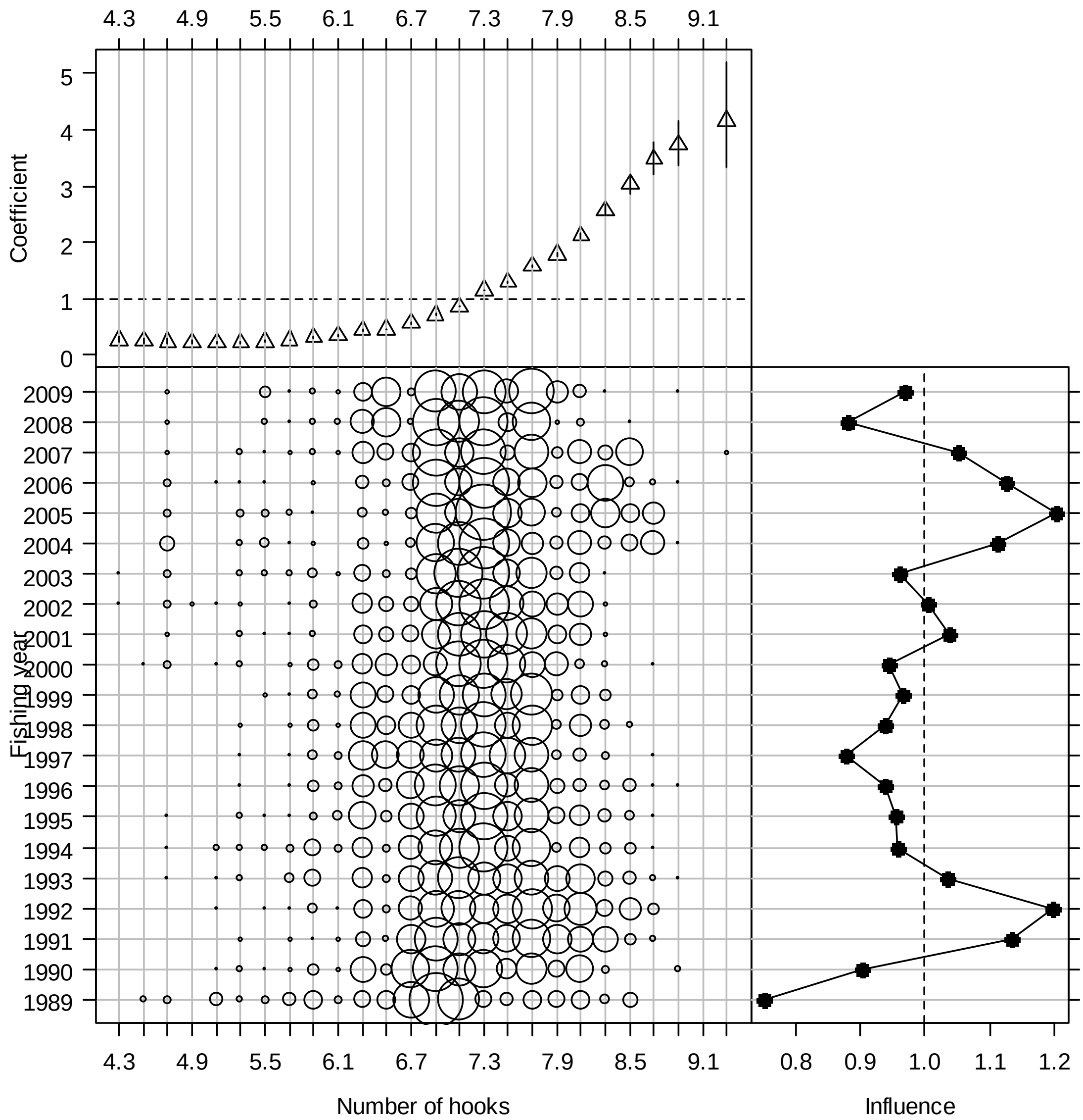
52
228
261
548
1866
2178
3205
3539
3811
Coefficient
2.0
1.5
1.0
0.5
0.0
Fishing year
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
Vessel
0.95
1.00
1.05
1.10
Influence

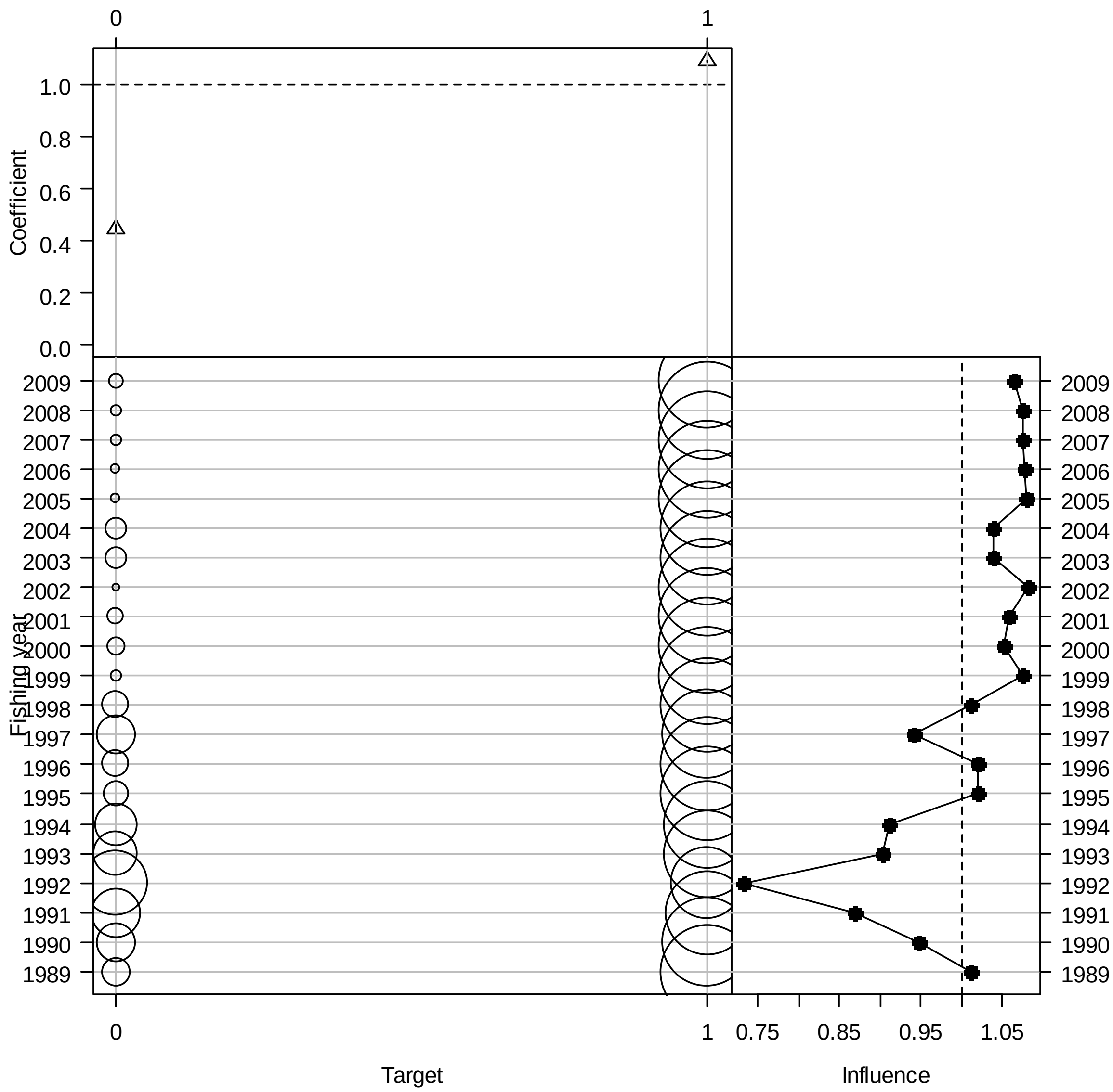

Coefficient
Fishing year
Month
Influence


Coefficient
Fishing year
Number of sets
Influence
0.05
0.35
0.65
0.95
1.25
1.55
1.85
2.15
0.97
0.99
1.01



Bay of Plenty longline
Coefficient
2.5
2.0
1.5
1.0
0.5
0.0
26
233
324
426
770
2127
3043
4222
4574
Fishing year
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
Vessel
0.85
0.95
1.05
Influence


4.3
4.9
5.5
6.1
6.7
7.3
7.9
8.5
9.1
Coefficient
0
1
2
3
4
5
Fishing year
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
0.8
0.9
1.0
1.1
1.2
Number of hooks
Influence

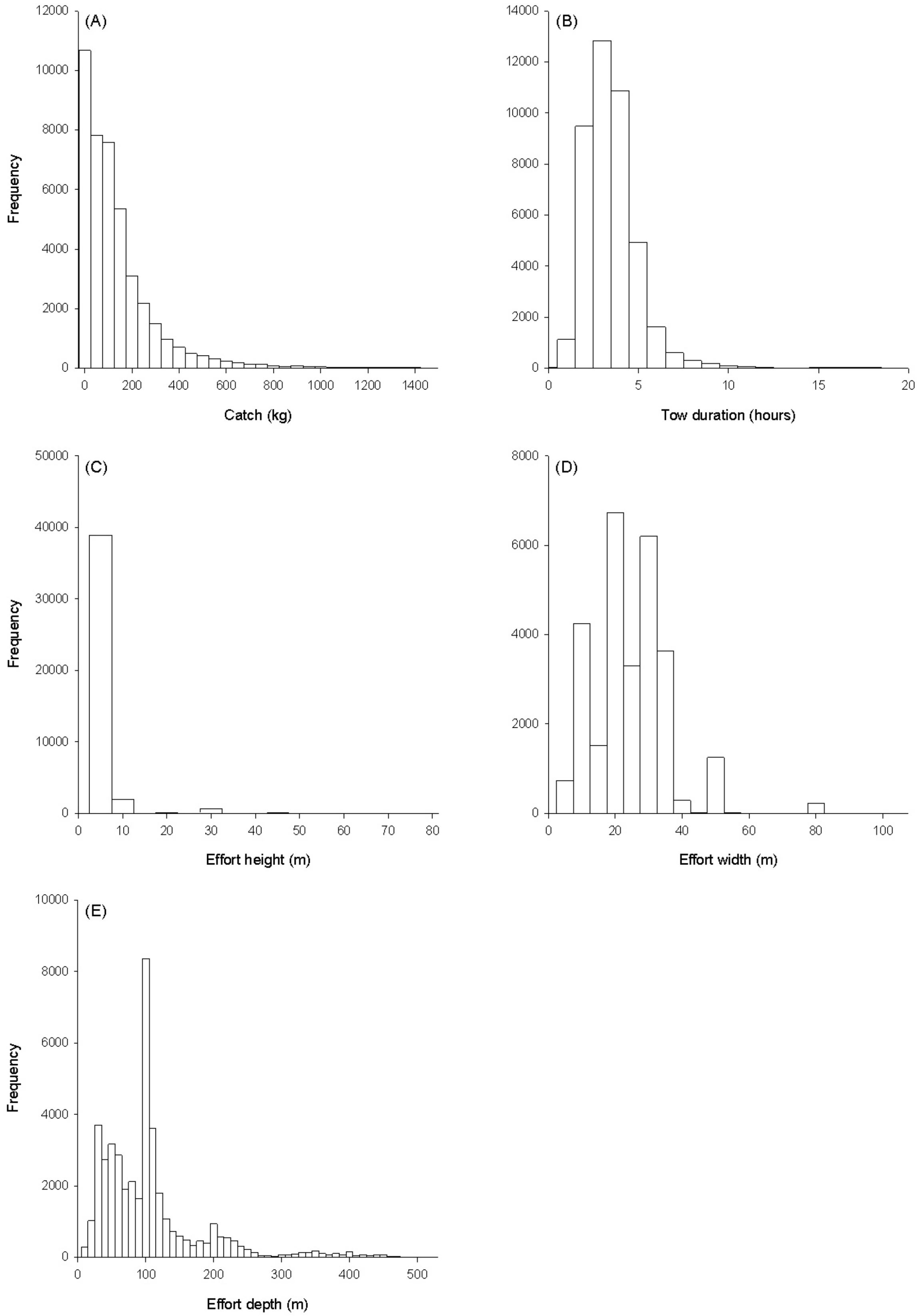


0
1
Coefficient
1.0
0.8
0.6
0.4
0.2
0.0
Fishing year
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
0
1
Target
0.75
0.85
0.95
1.05
Influence

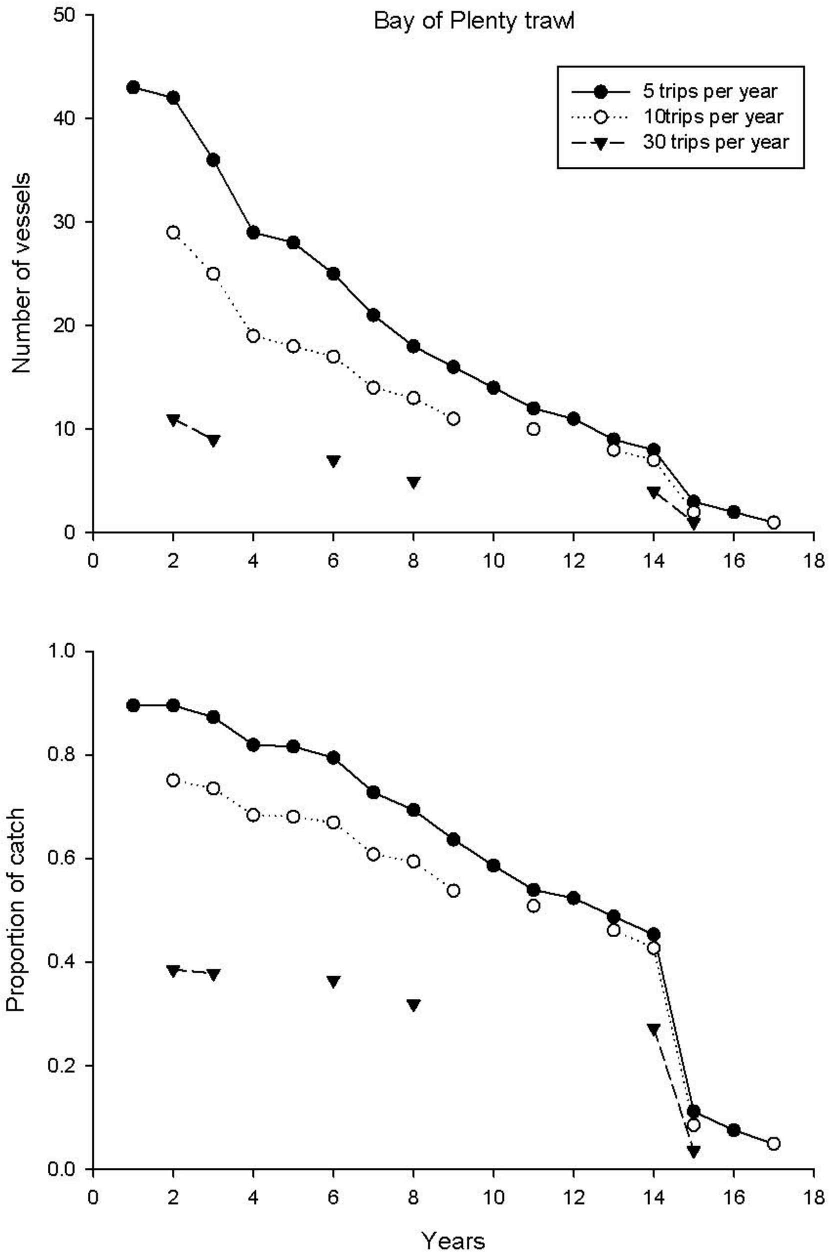
Appendix 20: Annual canonical SNA 1 sub-stock longline CPUE indices (c.v.'s in brackets).

Fishing year	Longline East Northland	Longline Hauraki Gulf	Longline Bay of Plenty
1989–1990	0.9956 (0.02)	0.7895 (0.01)	0.9415 (0.03)
1990–1991	0.9081 (0.02)	0.7969 (0.01)	0.7581 (0.02)
1991–1992	0.8509 (0.02)	0.8988 (0.01)	0.6262 (0.02)
1992–1993	0.9802 (0.02)	0.8103 (0.01)	0.7519 (0.02)
1993–1994	0.9941 (0.02)	0.7104 (0.01)	0.7566 (0.02)
1994–1995	1.0266 (0.02)	0.7155 (0.01)	0.8842 (0.02)
1995–1996	1.2002 (0.02)	0.8006 (0.01)	0.8971 (0.02)
1996–1997	1.2788 (0.02)	0.9542 (0.01)	0.9928 (0.02)
1997–1998	1.0071 (0.02)	1.0981 (0.01)	0.9851 (0.02)
1998–1999	1.0533 (0.02)	1.1606 (0.01)	1.0751 (0.02)
1999–2000	1.0206 (0.01)	1.0517 (0.01)	1.0052 (0.02)
2000–2001	0.9355 (0.02)	1.022 (0.01)	1.0373 (0.02)
2001–2002	0.8261 (0.02)	1.0532 (0.01)	1.0698 (0.02)
2002–2003	0.7729 (0.02)	1.1318 (0.01)	1.1131 (0.02)
2003–2004	0.8836 (0.02)	1.0922 (0.01)	1.0553 (0.02)
2004–2005	0.9668 (0.02)	1.0227 (0.02)	1.0807 (0.02)
2005–2006	0.9763 (0.02)	1.1997 (0.02)	1.1827 (0.02)
2006–2007	1.0691 (0.02)	1.2144 (0.02)	1.1511 (0.02)
2007–2008	1.2156 (0.02)	1.3335 (0.02)	1.2971 (0.02)
2008–2009	1.2434 (0.02)	1.2527 (0.02)	1.3555 (0.02)
2009–2010	0.9703 (0.02)	1.2652 (0.02)	1.3916 (0.02)

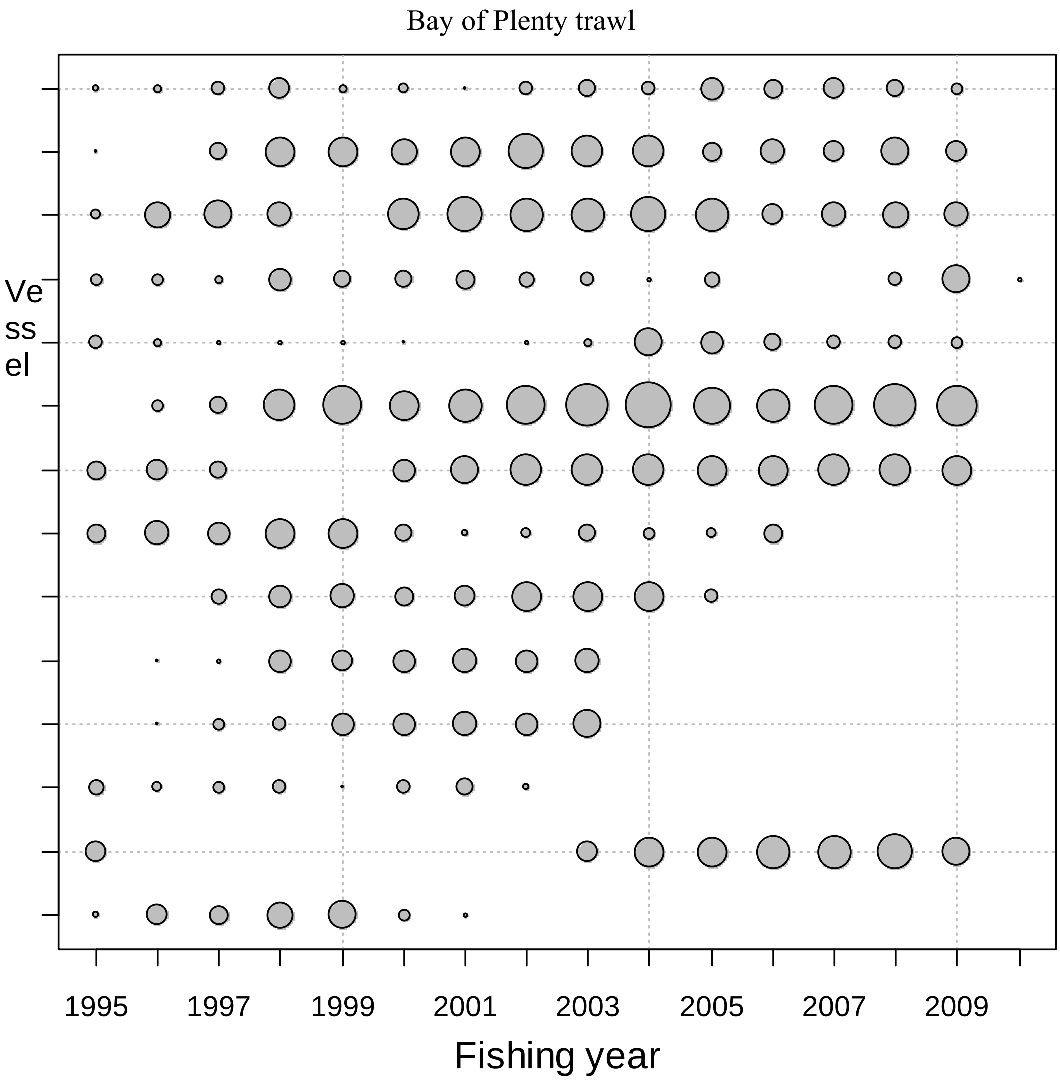
Appendix 21: Frequency distributions for catch (A), tow duration (B) headline height (C) gear wing-spread (D) gear depth (E) from the Bay of Plenty TCEPR and TCE data set.



Appendix 22: Graphics used to aid the vessel selection process. For each sub-stock/method combination, the number of vessels and the proportion of catch accounted for are plotted against the number of years those vessels contributed to the fishery in that sub-stock. Each of these graphs has three series spanning the range of the number of trips per year vessels undertook.

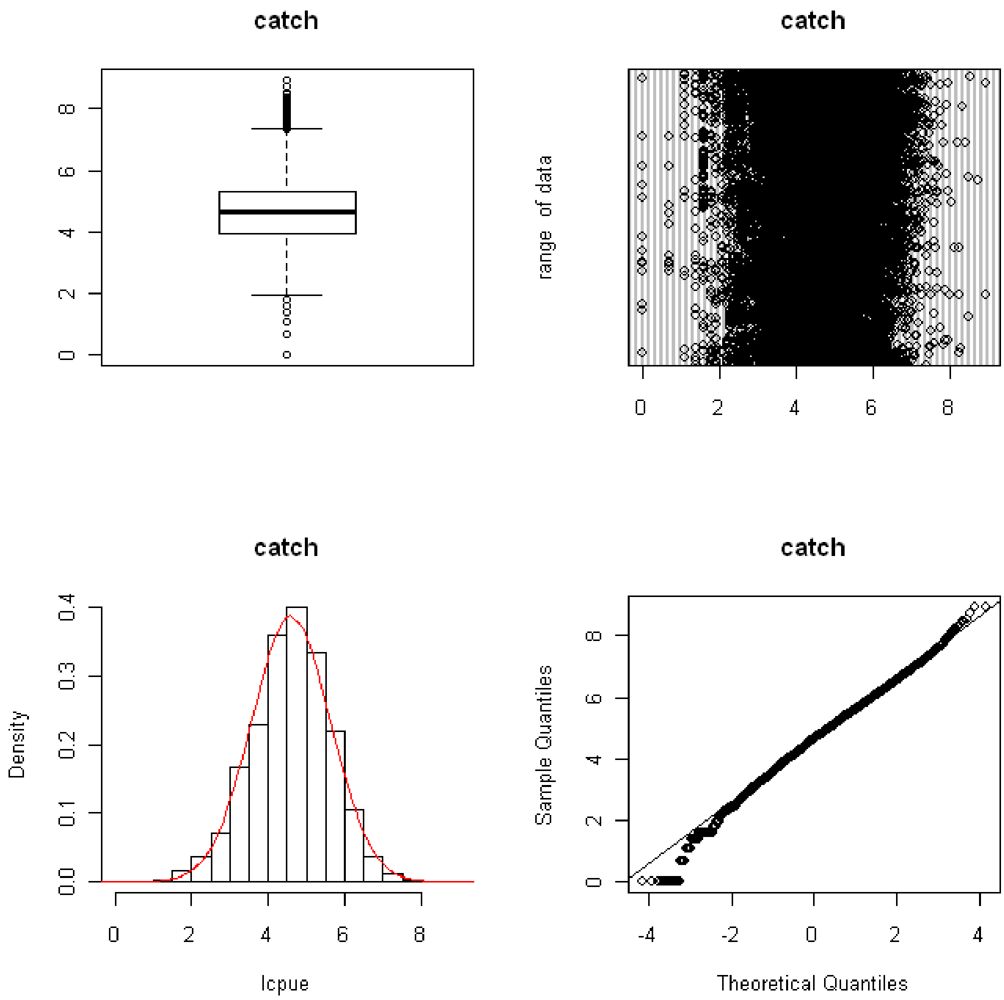


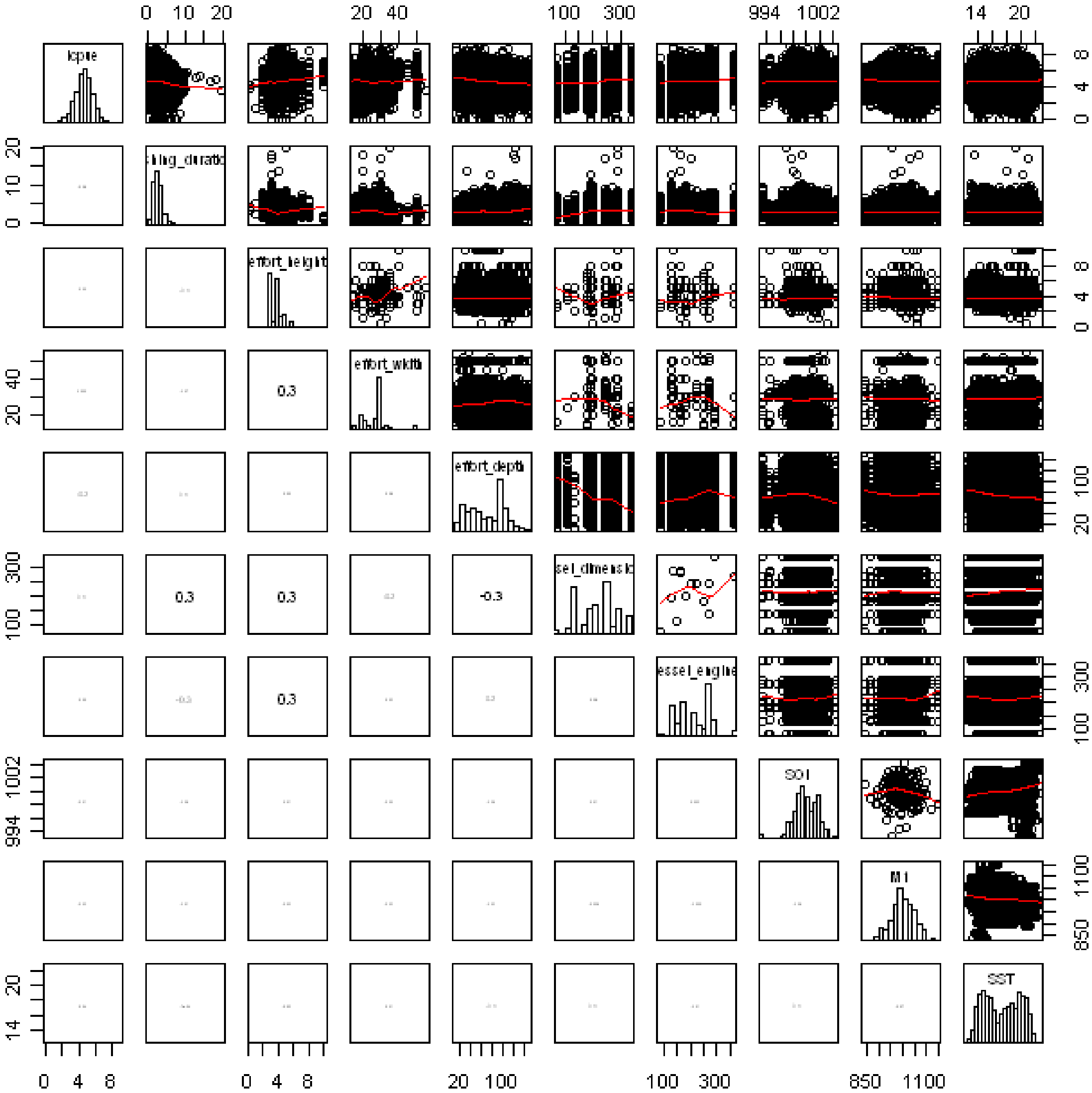
Appendix 23: Data coverage of core vessels selected for each sub-stock/method combination. Circles are proportional to the sum of catch for each vessel in that year.



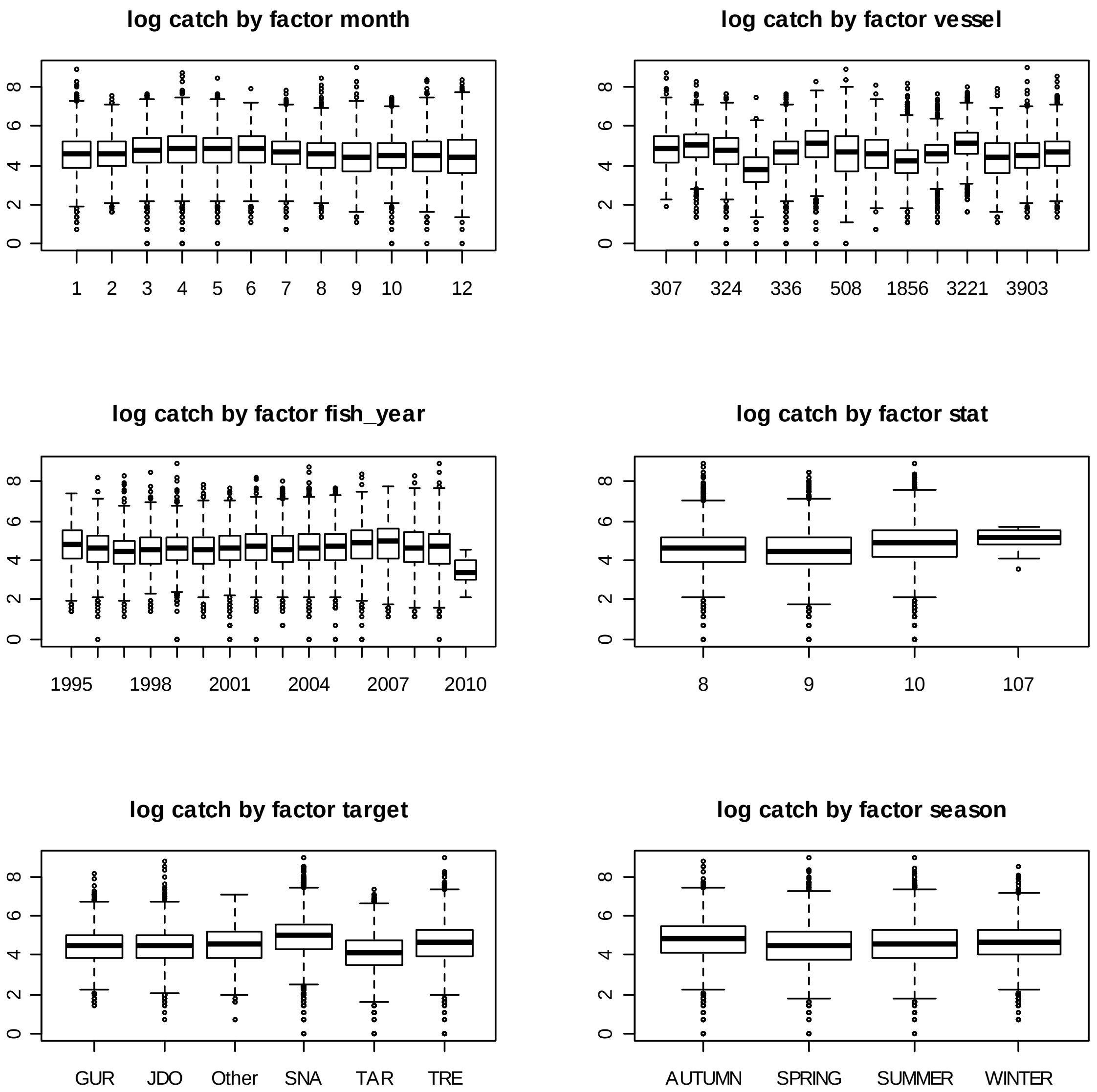
Appendix 24: Bay of Plenty single trawl raw response variable (log catch) plots investigating normality, homogeneity of variance, collinearity.

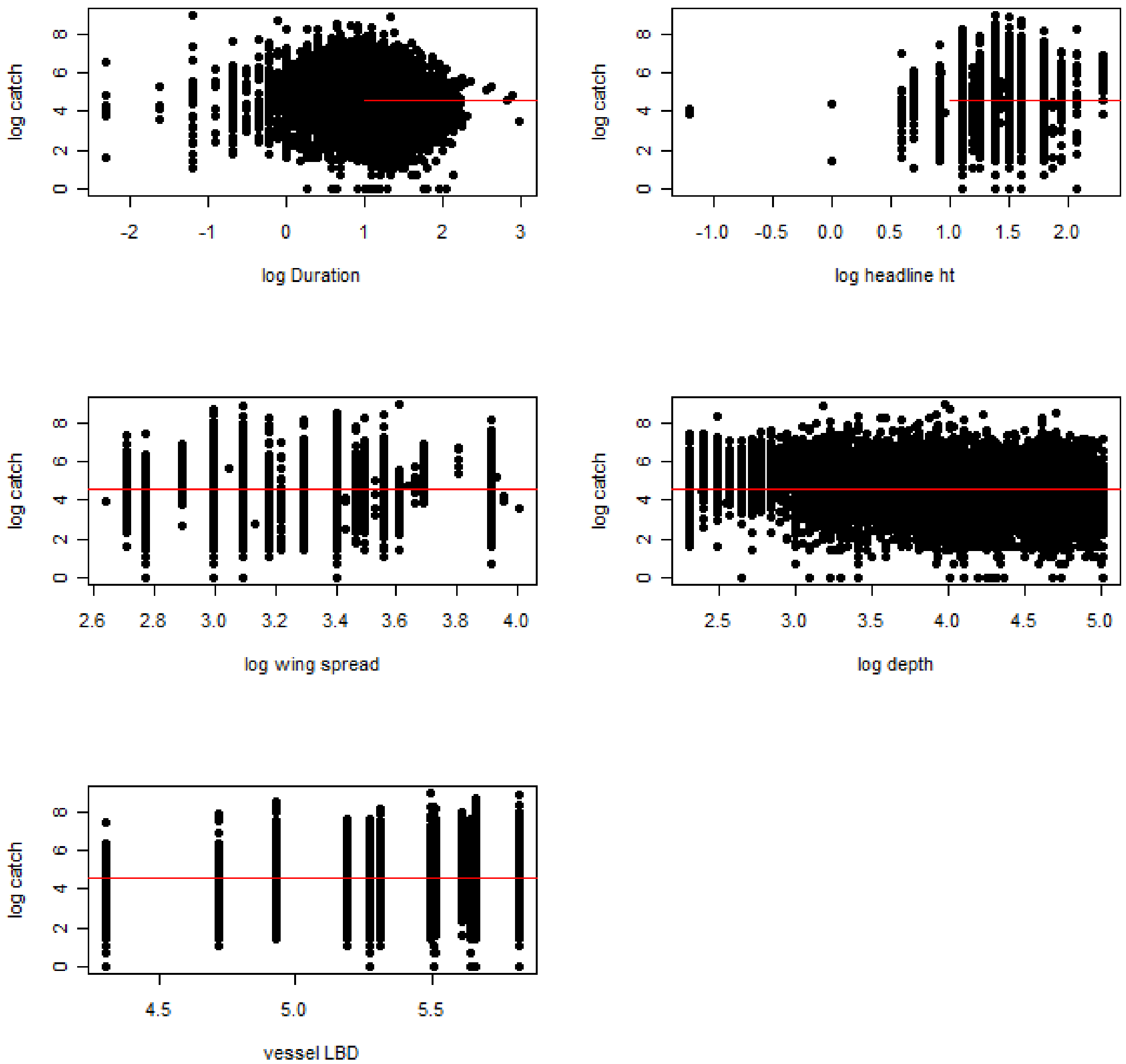
The first series of four plots for each area are all relevant to the assumption of normality. Three of these plots illustrate the distribution of the catch data, which should be similar to that of a normal distribution. The plot in the bottom right hand corner is a Q-Q plot, which plots the quantiles of the catch data against those of a theoretical normal distribution (if the catch data were exactly normal the data series would lie along the y = x line indicated on the graph). The second series of plots for each area are all relevant to the assumption of homogeneous variances. In these plots catch is illustrated across the different levels of categorical explanatory variables. A similar spread of catch across all levels would indicate homogeneous variances. The third series of plots for each area illustrate the presence of any correlations between explanatory variables offered to the model. The red trend lines and the size of the correlation coefficients opposite the relevant graphics indicate the strength of any correlations that may be present.



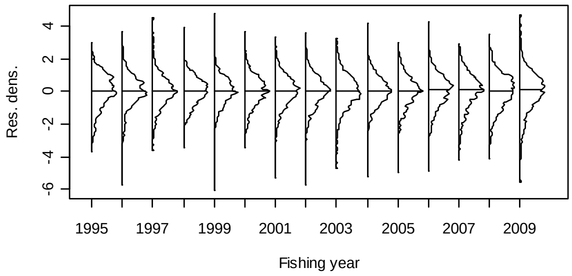


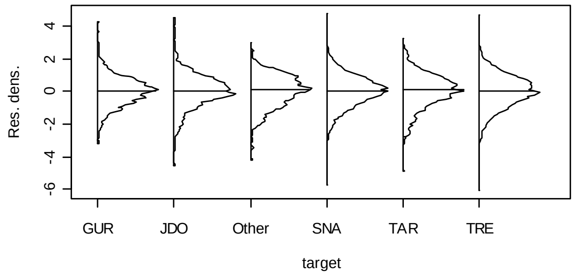
lcpue
effort_height
effort_width
effort_depth
SOI
MI
SST
0.3
-0.3

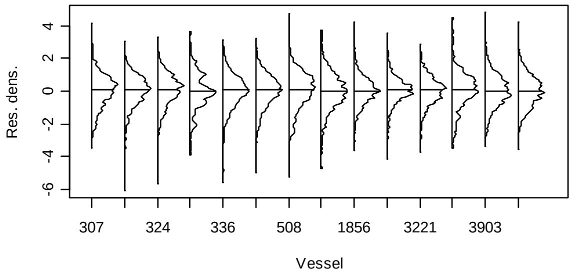

log catch by factor month
1 2 3 4 5 6 7 8 9 10 12
log catch by factor vessel
307 324 336 508 1856 3221 3903
log catch by factor fish_year
1995 1998 2001 2004 2007 2010
log catch by factor stat
8 9 10 107
log catch by factor target
GUR JDO Other SNA TAR TRE
log catch by factor season
AUTUMN SPRING SUMMER WINTER
0 2 4 6 8

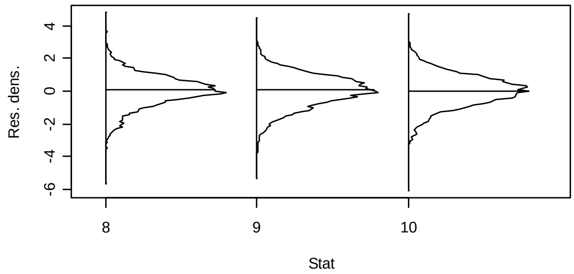

log catch
log Duration
log headline ht
log wing spread
log depth
vessel LBD

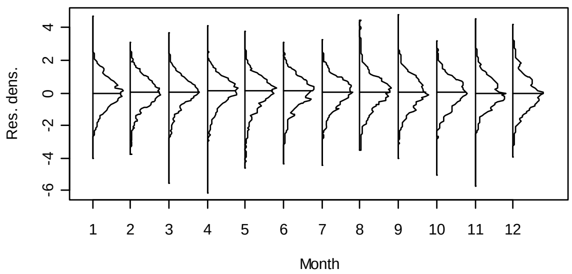
Appendix 25: Bay of Plenty Trawl Residual plots from the final regression model again testing for violations in the homogeneity of variance assumption from the fitted model.

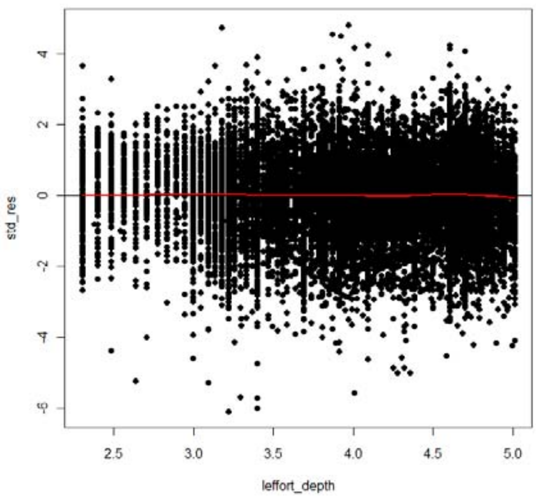




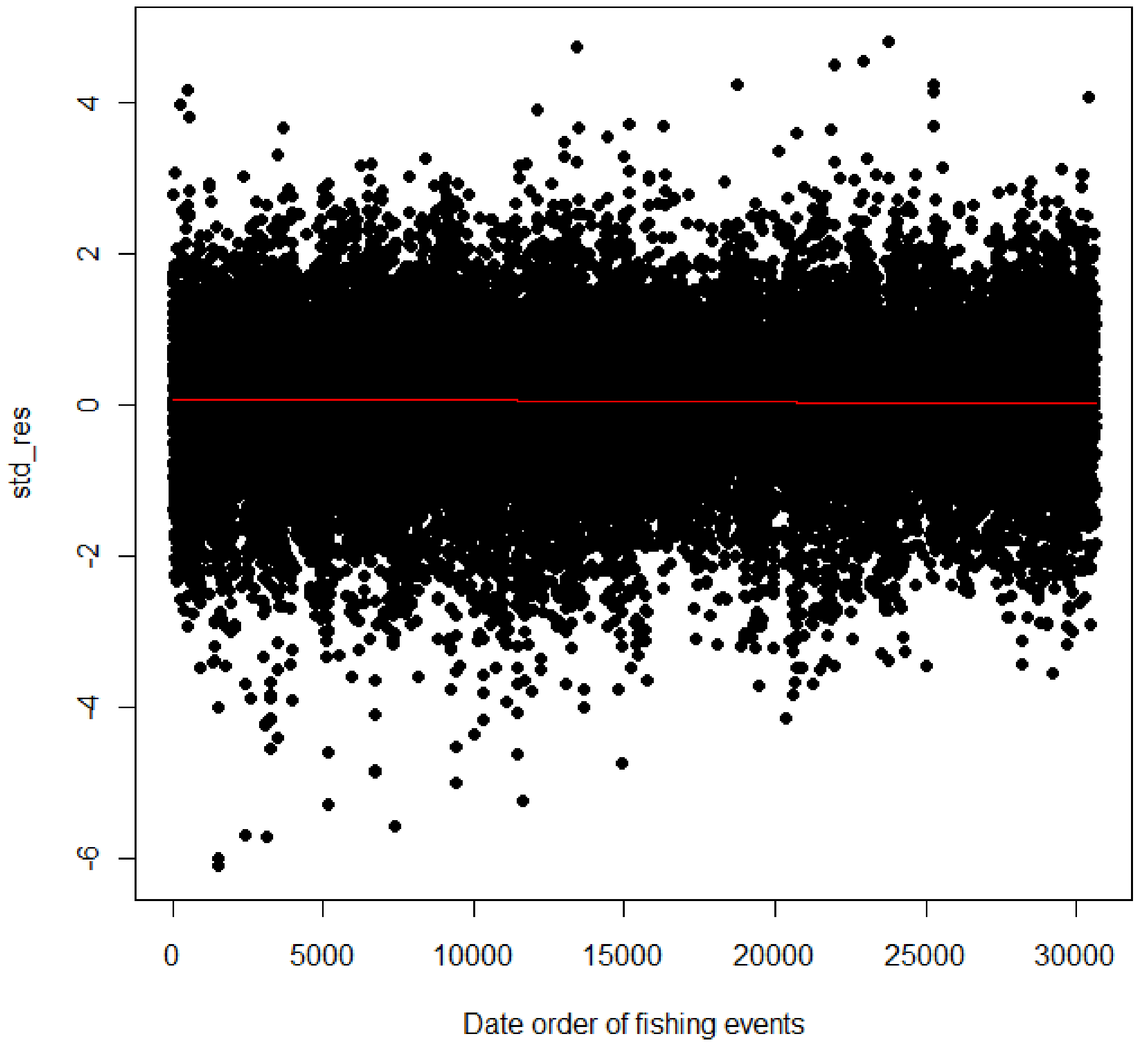




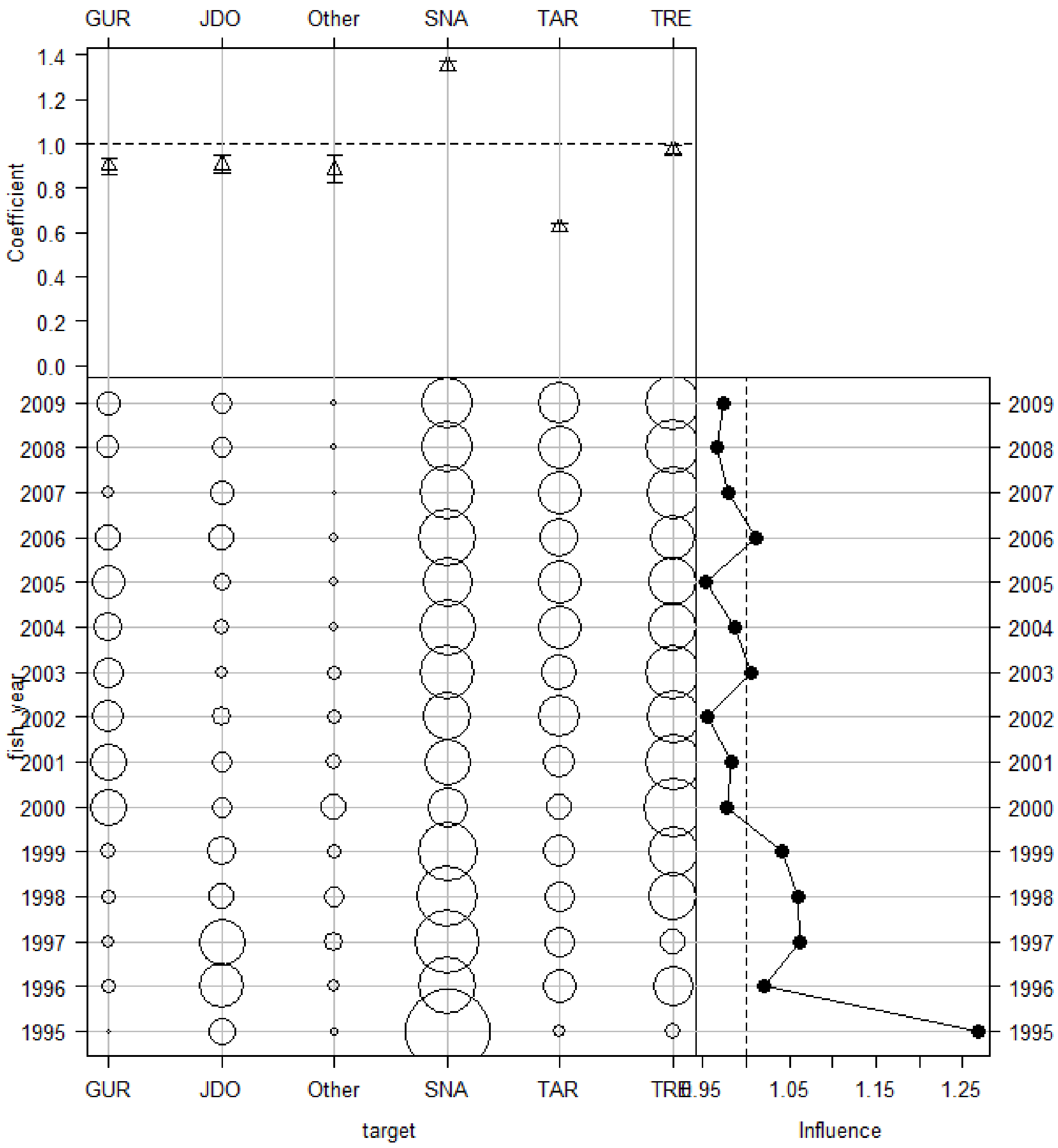


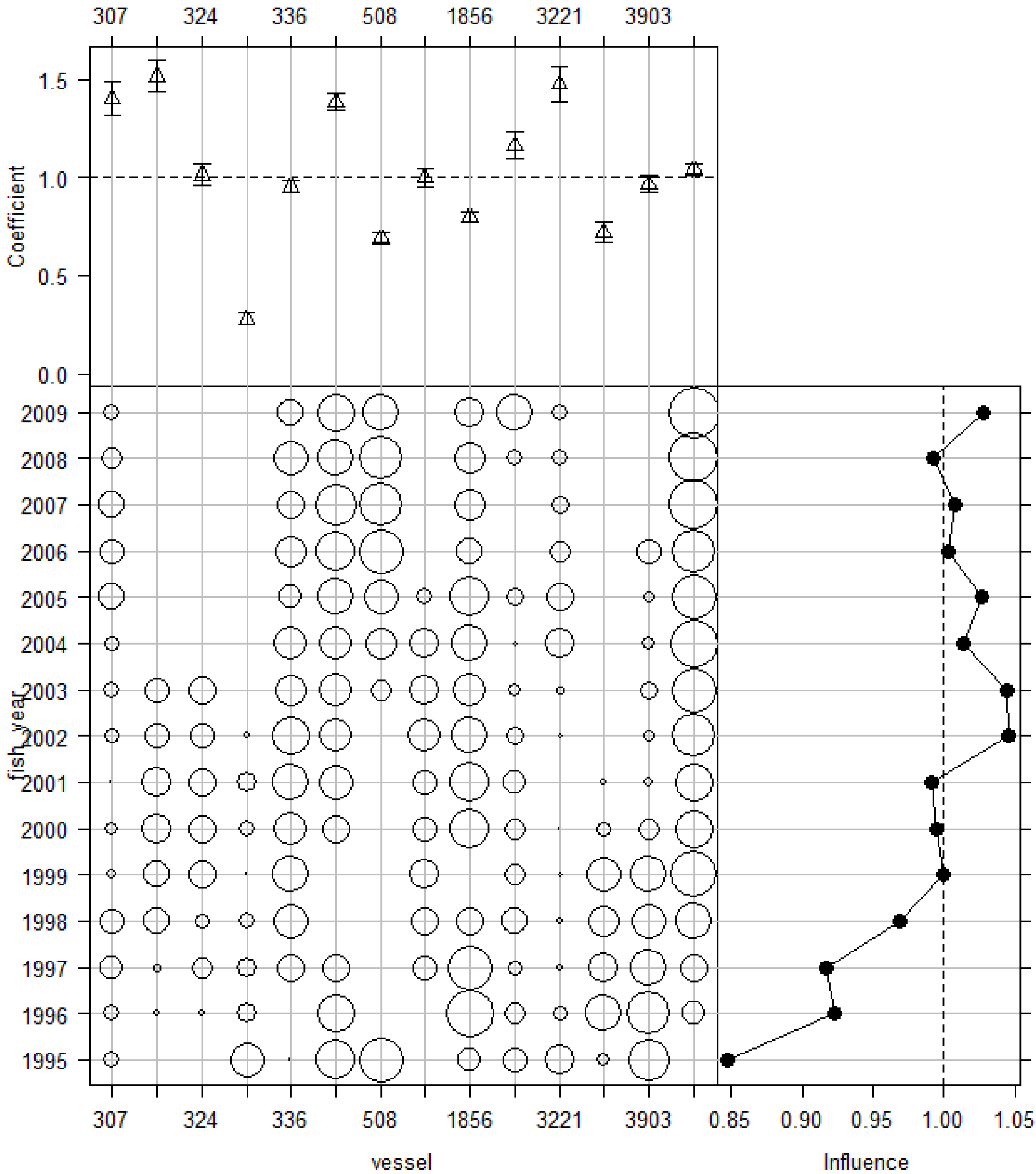


Appendix 26: Bay of Plenty single trawl residual plots from the final regression model testing for violations in the independence of the response data observations.

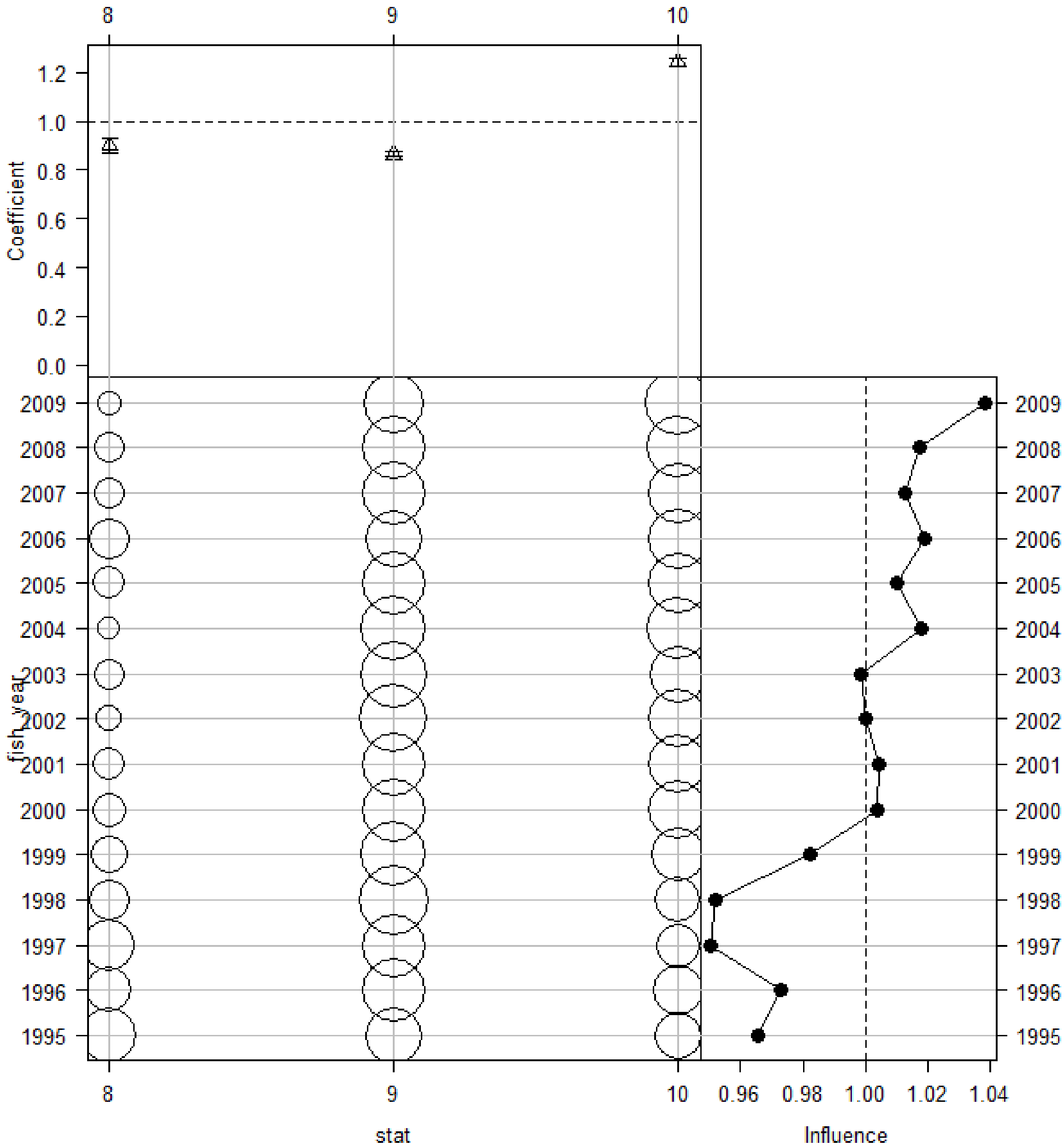


Appendix 27: Plots illustrating the influence of all significant variables in the Bay of Plenty single trawl log normal model (after Bentley et al. in press).

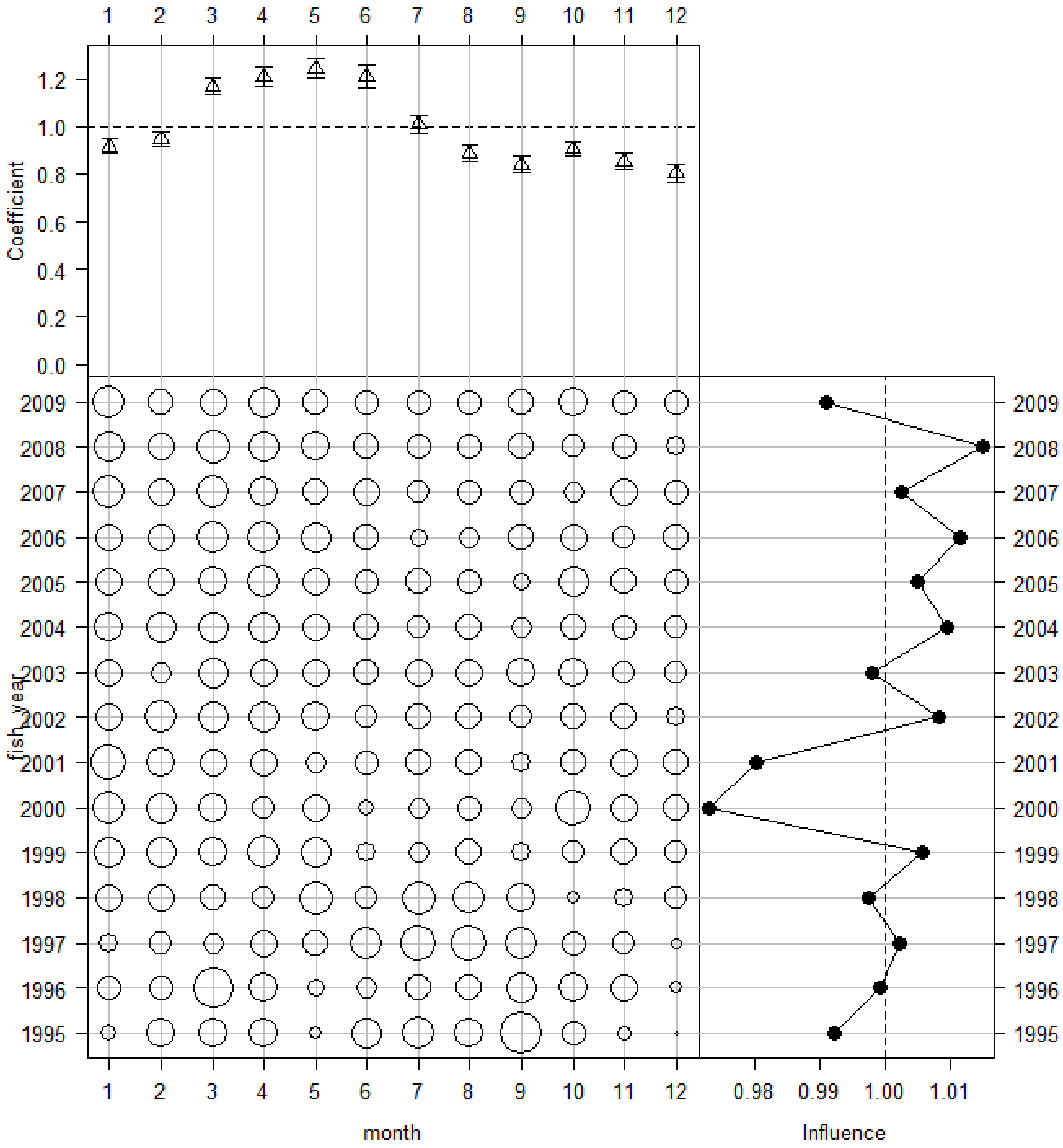


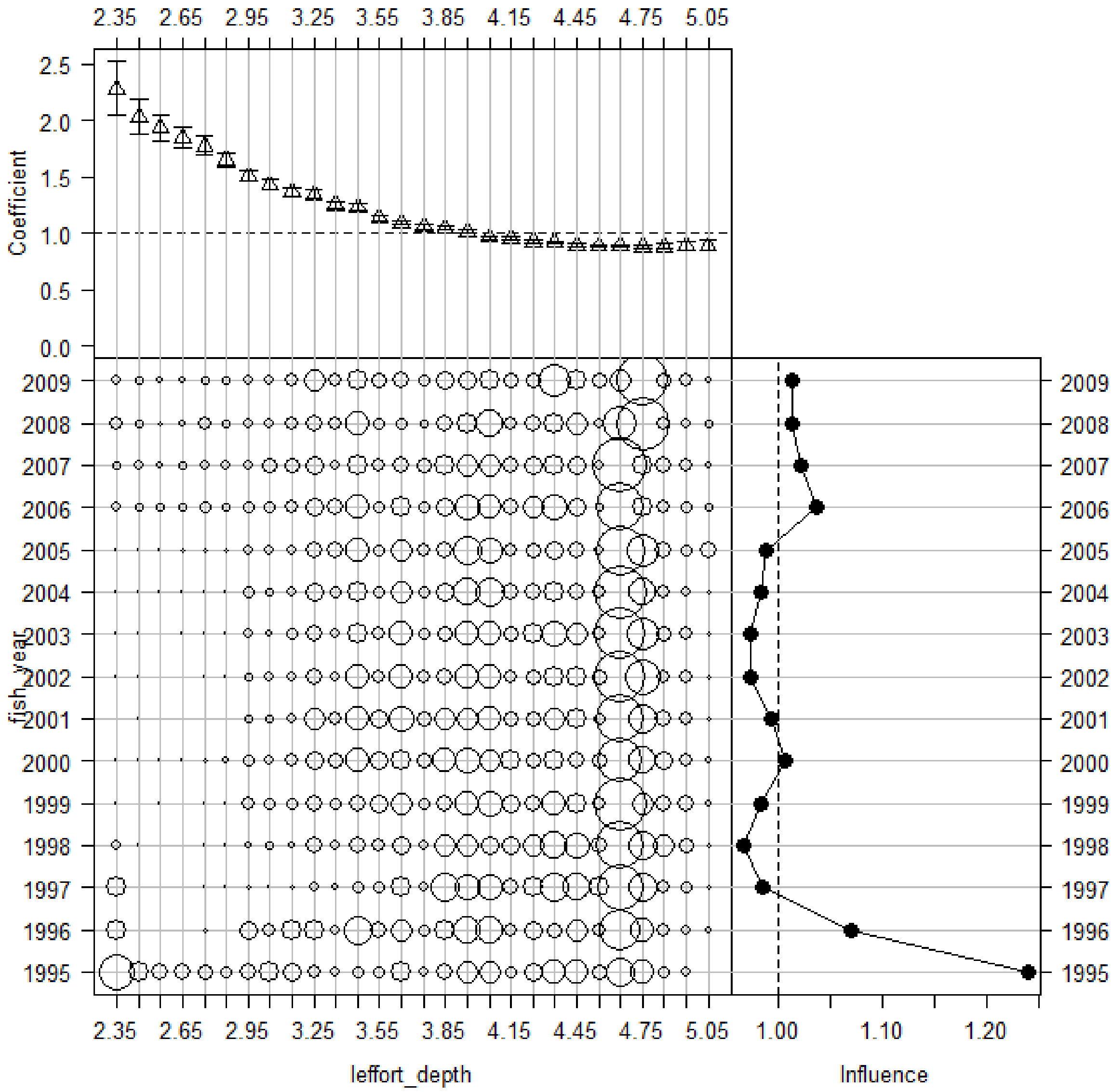


307
324
336
508
1856
3221
3903
Coefficient
1.5
1.0
0.5
0.0
fish_year
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
vessel
0.85
0.90
0.95
1.00
1.05
Influence



8
9
10
Coefficient
1.2
1.0
0.8
0.6
0.4
0.2
0.0
fish_year
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
0.96
0.98
1.00
1.02
1.04
stat
Influence


Coefficient
fish_year
month
Influence
1 2 3 4 5 6 7 8 9 10 11 12
1.2 1.0 0.8 0.6 0.4 0.2 0.0
2009 2008 2007 2006 2005 2004 2003 2002 2001 2000 1999 1998 1997 1996 1995
0.98 0.99 1.00 1.01



Coefficient
fish_year
leffort_depth
Influence

Appendix 28: Bay of Plenty single trawl CPUE indices (c.v.'s in brackets).

Fishing year	Log normal	Binomial	combined
1995–1996	0.909 (0.03)	0.256 (0.1)	0.554
1996–1997	1.006 (0.03)	0.327 (0.08)	0.694
1997–1998	0.902 (0.03)	0.44 (0.07)	0.709
1998–1999	0.949 (0.02)	0.794 (0.06)	0.908
1999–2000	0.948 (0.02)	1.379 (0.07)	1.025
2000–2001	0.942 (0.02)	0.798 (0.06)	0.902
2001–2002	1.003 (0.02)	1.073 (0.07)	1.032
2002–2003	1.091 (0.02)	1.283 (0.06)	1.165
2003–2004	0.948 (0.02)	1.26 (0.06)	1.008
2004–2005	1.015 (0.02)	1.747 (0.06)	1.14
2005–2006	1.057 (0.02)	1.653 (0.07)	1.177
2006–2007	1.084 (0.02)	1.674 (0.07)	1.21
2007–2008	1.241 (0.02)	2.235 (0.08)	1.439
2008–2009	1.018 (0.02)	1.54 (0.07)	1.122
2009–2010	0.9393 (0.02)	1.077 (0.06)	0.968