Ministry for Primary Industries
Manatū Ahu Matua

Developing indices of relative abundance from observational aerial sightings of inshore pelagic finfish; Part 2, expanding the dataset and producing annual indices for KAH 1 and TRE 1

New Zealand Fisheries Assessment Report 2014/35

P.R. Taylor
I.J. Doonan

ISSN 1179-5352 (online)

ISBN 978-0-478-43715-7 (online)

August 2014

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

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Taylor, P.R.; Doonan, I. (2014). Developing indices of relative abundance from observational aerial sightings of inshore pelagic finfish; Part 2, expanding the dataset and producing annual indices for KAH 1 and TRE 1.

***New Zealand Fisheries Assessment Report 2014/35*. 45 p.**

The work documented here was funded under MFish Project SEA2010-17. It was based on preliminary analyses of the aerial sightings data (MFish database *aer_sight*) carried out under MFish Project SAP2006-10 when a standardisation method was developed using only recent data under close scrutiny of the Northern Inshore Working Group (NINSWG). The WG guided development and accepted the method with the recommendation that it be used to produce annual indices of relative abundance for trevally and kahawai in the Bay of Plenty (BoP) using all data collected and also in east Northland if sufficient data were available.

The *aer_sight* database contains information on schools of inshore pelagic schooling finfish species collected by spotter pilots working in the domestic purse-seine fishery since 1976. The data used in the analysis include date, pilot, information on flight time, the flightpath followed during a day's flying, information on the sightings recorded on individual flights, the species composition and size of the schools making up each sighting, and data on each fishing operation.

The work described here aimed to analyse existing aerial sightings data to the end of the 2010–11 fishing year (i.e., 30 September 2011) and to carry out separate sightings per unit effort (SPUE) standardisations for stocks of kahawai and trevally in the Bay of Plenty (BoP) and east Northland, these two areas defining KAH 1 and TRE 1. The overall aim was to produce SPUE-based indices of abundance for KAH 1 and TRE 1. However, exploratory analysis of the volume and distribution of flights in east Northland indicated that the volume of data was too low to produce reliable estimates of relative abundance for this area.

For the BoP analyses, data were restricted to flights exclusive to this area, to fishing years from 1986–87 to 2010–11, to a single pilot who had collected most of the data in the area throughout that period, and to the first spotting flight of the day to avoid double counting. The surrogate for target species in the model was the species with highest representation in the purse-seine catch from the BoP on a given day — catch data were selected from a dataset of combined catches from the MPI catch-effort (*warehou*) and historic purse-seine (*fsu_new*) databases.

The approach adopted here was based on the two-component, binomial-lognormal approach often used for catch-per-unit-effort standardisations. The SPUE analysis is a catch per unit effort analogue for which effort is an important component. Search effort can potentially be derived from one of two sources within the aerial sightings data. Flightpath data is perhaps the more attractive because it is adjusted for non-search or operational effort, but it is not available at the level of the individual flight. To avoid possible confounding from double counting, data selection was restricted to the first flight of the day which precluded use of the flightpath data. Therefore flying effort was available only from the alternative source, flight length, or, to use a definition that eliminates the ambiguity of flight distance, flight time. To account for non-search time, flying time was adjusted by the number of operations and the total number of sightings, using a linear model fit outside of the main modelling method.

The main analysis was performed using the generalised additive model (GAM) to standardise observed tonnages of each of the two inshore schooling pelagic species, trevally and kahawai. Modelling adopted a two-component approach for each species: a binomial GAM to model the presence-absence of sightings of the species of interest on each flight, and a lognormal GAM to

standardise observed tonnages. Predictors included the adjusted effort variable along with fishing year, month, time of day, southern oscillation index, sea surface temperature, target species, and moonphase.

Results of the standardisations showed reasonable fits with no clear violations of model assumptions and the NINSWG accepted the estimated indices of relative abundance. Effort was not accepted by any of the four models; target was accepted for both the lognormal and binomial kahawai models but for neither of the trevally fits. Levels of variability explained for the selected models were 25.6% and 11.9% for the trevally lognormal and binomial fits respectively, and 32.3% and 17.7% for the kahawai lognormal and binomial fits respectively.

The final trevally indices showed an overall decline in the BoP since 1986. This is consistent with anecdotal reports from fishers and spotter pilots, and a contracting age distribution in catches made by the bottom trawl fishery. For kahawai, binomial models of the annual proportion of flights with zero sightings, and models of the tonnage sighted, both suggest that abundance of this species, in terms of both number and size of schools, has increased following the reduction of commercial catch in the early 2000s. This observation is also consistent with anecdotal reports from commercial and recreational fishers, and spotter pilots.

The NINSWG concluded that models of SPUE for kahawai and trevally probably do reflect, to some degree, the abundance of these two species in the Bay of Plenty. They recommended that SPUE indices should be used for stock assessment, with stock assessment model diagnostics employed to gauge the quality (and appropriate weight) of the abundance indices.

Summary for web
Annual indices of relative abundance for trevally and kahawai in the Bay of Plenty were estimated using aerial sightings data to the end of 2010–11. There were insufficient data collected in east Northland so indices were not developed there.

1. OVERVIEW

Stock assessments of inshore schooling pelagic species have been hampered by our inability to produce a measure of annual relative abundance. The main target fishery method for these species in QMA 1 is purse-seine, and, for reasons discussed below, using catch per unit effort from purse-seine fisheries as a stock index is unlikely to be reliable. Aerial sightings data offer a source of information that has the potential to provide annual relative abundance indices, and which is cost-effective to sustain.

The aerial sightings database (*aer_sight*) has been maintained by agencies of the Minister for Fisheries since 1976. It contains data on schooling pelagic species recorded by pilots assisting in the purse-seine fishing operation and dates almost to the beginning of this fishery in 1974. The database is in electronic format and has, until September 2011, been administered by NIWA for the Ministry of Fisheries using the relational database environment, EMPRESS.

The *aer_sight* database contains the longest available time series of information for the six main inshore schooling pelagic species taken by purse-seine: trevally (*Pseudocaranx dentex*), blue mackerel (*Scomber australasicus*), jack mackerel (*Trachurus declivis*, *T. murphyi*, and *T. novaezelandiae*), and kahawai (*Arripis trutta*), and for the highly migratory species skipjack tuna (*Katsuwonas pelamis*), on which the domestic purse-seine industry was founded. Flying effort has been quite consistent although some variation is evident particularly since 2004 (Taylor, unpublished results). By contrast, purse-seine catch and effort data have been collected only since 1982, and are unreliable during the period of transition (1988–89) from the Fisheries Statistics Unit (FSU) to the present Quota Management System (QMS). Therefore, the aerial sightings data are the longest and most consistent time series of information for some species of schooling pelagic species in New Zealand waters.

1.1 Aim of the study and scope of the report

The work documented here was funded under MFish Project SEA2010-17. The aim of the study was to analyse existing aerial sightings data to the end of the 2010–11 fishing year (i.e., 30 September 2011) and to carry out sightings per unit effort (SPUE) standardisations for stocks of kahawai and trevally in the Bay of Plenty (BoP) and separately for stocks of kahawai and trevally outside the BoP. The overall aim was to produce SPUE-based indices of abundance for KAH 1 and TRE 1 through investigation of the aerial sightings data.

At a meeting on 31 August 2009, the Northern Inshore Working Group had made the following recommendations regarding the investigation and development of relative abundance indices for small pelagic fishes using the aerial sightings data.

- The development of an aerial sightings index for the northern purse-seine fishery should be undertaken in three progressive stages.

- <u>Stage 1</u>
  The first stage to be based on the following data set:
    a. data collected on the new form (i.e., since April 1998);
    b. data collected by Pilot #2 only (who collected most of the data);
    c. data collected during the first flight of each day and from flights exclusive to the Bay of Plenty (BOP).

- <u>Stage 2</u>
  If the first stage appears to have been successful (based on diagnostic tools) the analysis will be expanded to include all years for which data exist for the BOP.

- <u>Stage 3</u>
  In stage three the analysis is to be expanded to include other areas and data collected by all pilots.

Stage 1 was completed under SAP2006-10 and reported by Taylor (2014) as "Part 1" to reference its place in the reporting process. The work reported here is for Stages 2 and 3, and also includes the production of preliminary annual stock indices for KAH 1 and TRE 1, all of which is referred to in the title of this report as "Part 2". Consideration of the patchy nature of data available from *aer_sight* for most areas other than the BoP resulted in the NINSWG concluding that Stage 3 should only include sightings data from east Northland. Extensive detail regarding the aerial sightings dataset, its collection and salient features, and a number of important investigations as part of the exploratory data analysis carried out under SAP2006-10 were recorded by Taylor (2014) and are not repeated here.

1.2 Review of preliminary study to standardise aerial sightings data

The most important commercial species in the domestic purse-seine fishery has been skipjack tuna, which has a roughly summer–autumn presence in New Zealand waters and is fished accordingly. Kahawai was the second most important commercial target, being fished mainly in the winter–spring when skipjack is unavailable, but, since catch limits were set for this species in 1990–91, more attention has been given to jack mackerel and blue mackerel, with the latter targeted preferentially. Blue mackerel have been more valuable as a commercial species than jack mackerel, although jack mackerel have been important as a high volume, low value fishery. More recently the market price of jack mackerel has increased and stabilised, resulting in closer parity in the preference for mackerel species. Trevally was fished consistently through early years of the fishery, but catches declined rapidly, so that total TACs are now very low (3932 t total for all fish stocks).

The work documented here is a continuation of that carried out under SAP2006-10, which investigated the efficacy of producing relative abundance indices for the inshore schooling species listed above (except skipjack tuna) from the aerial sightings data (Taylor, 2014). Jack mackerel are managed as separate species but, because they are not separated by species in the data they were removed from the list by the Northern Inshore Working Group (NINSWG) during discussions about SAP2006-10. Blue mackerel were also removed when preliminary analyses indicated high interannual variation in relative abundance indices, suggesting that aerial sightings were only indexing the abundance of that part of a larger stock that was present on the fishing grounds.

For the SAP2006-10 analyses, data were restricted to flights exclusive to the Bay of Plenty, where the greatest density of data was centred, to fishing years between 1998–99 and 2007–08 for which information on the number of fishing operations was readily available from *aer_sight*, to a single pilot who had collected most of the data in the area since 1976 (pilot #2), and to the first real working (i.e., fish-spotting) flight of the day to avoid problems of double counting. "Double counting" refers to repeated sightings of the same fish and would be more correctly named "multiple counting".

Effort was considered an important factor in the SAP2006-10 analyses because the approach adopted there was based on the two-component, binomial-lognormal approach often used for catch-per-unit-effort standardisations. Search effort can potentially be derived from one of two sources within the aerial sightings data. Flightpath data is perhaps the more attractive because, in the case of Pilot #2, it is adjusted for non-search or operational effort, but it is not available at the level of the individual flight. However, the restriction to data from the first flight of the day mentioned in the previous paragraph precluded use of the flightpath data. Flying effort was therefore available only from the alternative source, flight length, or, to use a definition that eliminates the ambiguity of flight distance, flight time.
It was known from the pilots that not all flying time was search time, so flight length was not an accurate representation of search effort. To account for non-search time, flying time was adjusted by the number of fishing operations and the total number of sightings observed using a linear model fit outside of the main modelling method. Values of adjusted effort were estimated for each flight and its performance was compared with the unadjusted effort by forcing each into separate model fits. These two effort regimes were compared with a third in which the adjusted effort was offered for selection to a third model fit, rather than being forced.

The main analysis was performed using the generalised additive model (GAM) to standardise observed tonnages of each of the two inshore schooling pelagic species, trevally and kahawai. Modelling adopted a two-component approach for each species: a binomial GAM to model the presence-absence of sightings of the species of interest on each flight, and a lognormal GAM to standardise observed tonnages. Predictors included each of the three effort variables along with fishing year, month, time of day, southern oscillation index, sea surface temperature, target species, and moonphase.

Results of the standardisations showed reasonable fits with no clear violations of model assumptions. Levels of variability explained for the selected models ranged from 20.4 to 23.8% for the trevally binomial, 19.4 to 26.4% for the kahawai binomial, 50.2 to 51.6% for the trevally lognormal, and 39.5 to 44.6% for the kahawai lognormal. Comparisons of results showed no clear advantage of using one measure of effort over the other. An unusual outcome within the trevally binomial fit occurred with pilchard as the target species. This was shown to result from the fact that no trevally was sighted when pilchard was the target species.

A second run of the models was performed using purse-seine catch for the period 1998–99 to 2007–08 as a surrogate for target species. The aim here was to inform discussion on whether the research should be extended back further in time. This was necessary because target species from the purse-seine fleet (which was used in the SAP2006-10 standardisation) is not available from catch data earlier than 1998. The results of this analysis showed little difference from those produced using the modal target data (for the same period) and it was concluded by the NINSWG that catch does provide an acceptable surrogate for target.

The results from SAP2006-10 (Taylor 2014) showed that reasonable indices can be expected for trevally and kahawai. The method was accepted by the Northern Inshore Working Group and recommendations were made to extend the work with the aim of producing annual indices of relative abundance for trevally and kahawai within QMA 1 over the longest possible time frame. The ultimate aim was to use these series as annual relative stock indices of abundance in stock assessment models for these two species in QMA 1.

2. METHODS

All analyses and most data manipulation/grooming were performed in the *R* statistical modelling environment (R Core Team 2012). Some manipulation was performed at the time of data extraction from the aerial sightings database (*aer_sight*) using the EMPRESS Standard Query Language (SQL).

2.1 Data extracts and processing

Data for the analysis included extracts from three databases: *aer_sight*, the MPI catch-effort database (*warehou*), and the Fisheries Statistics Unit database (*fsu_new*) (see Figure A1).

Central to the data extract was *R* code originally developed by Middleton et al. (2010) for analyses providing information contributing to NINSWG discussions relevant to SAP2006-10. This code was designed to investigate the number of records available from *aer_sight* that had associated target-species data from the purse-seine fleet available. The list of daily flying identifiers (*flt_grp*) so produced were used as a basis for data extracts from *aer_sight* for the SAP2006-10 analysis. This *R* code is referred to hereafter as the targetInv code (for "target investigation"). The targetInv code was modified for the SEA2010-17 analysis as described below.
Broadly speaking, the modifications expanded the time frame and provided a tool for producing a list of *flt_grp* identifiers for the BoP and east Northland. An underlying requirement of the work undertaken here was to determine whether there were sufficient data for an analysis. Work under

SAP2006-10 had shown that there were enough for both the trevally and kahawai analyses in the BoP for the period 1998 to 2008, and it was expected that this would remain the case for the additional years included here. However, the volume of available data for these species in east Northland was unknown. Consequently, summaries of data were tabulated to ensure that sufficient were available in both areas on an annual (fishing year) basis. These summaries are presented and discussed in the appropriate sections below.

2.1.1 Preliminary step — testing the modified code

Following its modification, the targetInv code was tested using an extract from *aer_sight* for the SAP2006-10 timeframe. Middleton et al. (2010) had extracted data directly from a database using the *R* database access package RODBC, but that method was not used here. Instead, data were extracted from *aer_sight* as ascii files that were then read into *R*, before being fed into the modified targetInv code to investigate its correct operation and that the extract from *aer_sight* was correct. The resulting output was compared with that of Middleton et al. (2010).

2.1.2 Subsequent steps — extracting and processing the data

Data extraction and processing was a complex procedure comprising a number of steps. Because of their complexity, these steps are not described in the body of this document, but, to ensure that a record is available, they are presented in Appendix A.

2.2 Further data considerations

2.2.1 Pilot and time frame

To avoid difficulties arising from the effects of multiple pilots (e.g., different methods of recording effort, see *Features of the Data* in Taylor (2014)) and to capture particular features largely exclusive to his data collection method (e.g., individual tonnages of the component species in mixed schools), data for pilot #2 only were extracted from the aerial sightings database. For aerial sightings data collected before the latest revised form was introduced (June 1998), operational data were available from the private collection of Pilot #2 and extended the available operational data back to 4 June 1983. They were entered into the *aer_sight* database under project DAT2008-01A and used in the present analysis. So, to utilise operational data for adjusting the flight effort, the extract was restricted to data collected after 3 June 1983, and this was further adjusted to post December 1985 when the revised data collection form with improved flightpath data was introduced (see Taylor 2014). Therefore, the dataset covered the period 1 January 1986 to 30 September 2011 with additional restrictions and omissions as described below.

2.2.2 Area and flight number

Data selection for the analysis was limited to the first flight of the day and days when all flights were exclusive to the Bay of Plenty (BOP). This was achieved by limiting selection to those days on which flightpath data from Panel 5a (see Figure 1) indicated that flying was exclusive to the BOP. Selected fields included date, fishing year (fsyr), month, flight index (a unique flight identifier), species code, number of sightings of the species of interest, tonnage (pilot's tonnage estimate for about 97% of records, and an estimate using school number, minimum and maximum tonnage estimates for the remainder), time of takeoff, flying effort (in decimal hours), half degree grid square code, number of fishing operations (nops), sea surface temperature (sst), southern oscillation index (soi), and moonphase (moon) (proportion of disc illuminated). The total number of sightings of all species (*totsit*, see Section 2.3) was calculated and added at Step 5 (Figure A1, see Appendix A) along with target species.

2.2.3 Target species

Target species, in the form of catch data, were available from *Warehou* and from *fsu_new* (Fisher & Sanders 2011). The catch data were processed to provide target species as the species with the highest estimated catch on any day. Where more than one species met this criterion, the code "MIX" was assigned.

2.2.4 Environmental data

Several environmental covariates were included in the SAP2006-10 analysis, including sea surface temperature (sst), southern oscillation index (soi), and moonphase (proportion of the disc illuminated). For the present work, soi was updated from the National Center for Atmospheric Research's (NCAR) web pages on Climate and Global Dynamics[1] and moonphase was updated from the Astronomical Applications Department of the US Naval Observatory website[2]; sst data were updated from the Leigh Laboratory dataset, which was available until May 2011. Missing data for the final months (June – September, 2011) were imputed using the following standard offset method:

a. Calculate the long-term mean for all months.
b. Using months in the year (2011) for which data are available, subtract the respective long-term mean from each month and then use the mean of those differences as an offset to calibrate the missing monthly values from the respective long-term mean values.

2.2.5 Vessel

The *fsu_new* database contained records for 31 vessels (Table 1). To ensure that target was determined only from vessels that Pilot #2 had assisted (and had therefore been included in the aerial sightings data in some way), records from vessels that were not part of the domestic fleet were omitted from the dataset. So, of the 31 vessels for which data were available, 9 were accepted as meeting the criterion.

2.2.6 Mixed schools

The bulk of the data in *aer_sight* comprise sightings of single-species schools. Not all pilots record mixed schools, but data collected by Pilot #2 also include sightings of mixed schools and tonnage for each species comprising the school. Mixed schools are sightings of the species of interest (e.g., kahawai) mixed with other species (e.g., blue mackerel) (see Taylor 2014). Mixed schools of the species of interest were included in both of the model components used here (log normal and binomial; see below). For the binomial model, mixed schools were simply included as a sighting of the species of interest, but the requirement of tonnages for the lognormal model meant an extra link to the newly created database table containing names and tonnages of component species in schools of a mixed sighting.

[1] cgd.ucar.edu/cas/catalog/climind/soi.html

[2] http://aa.usno.navy.mil/data/docs/MoonFraction.php

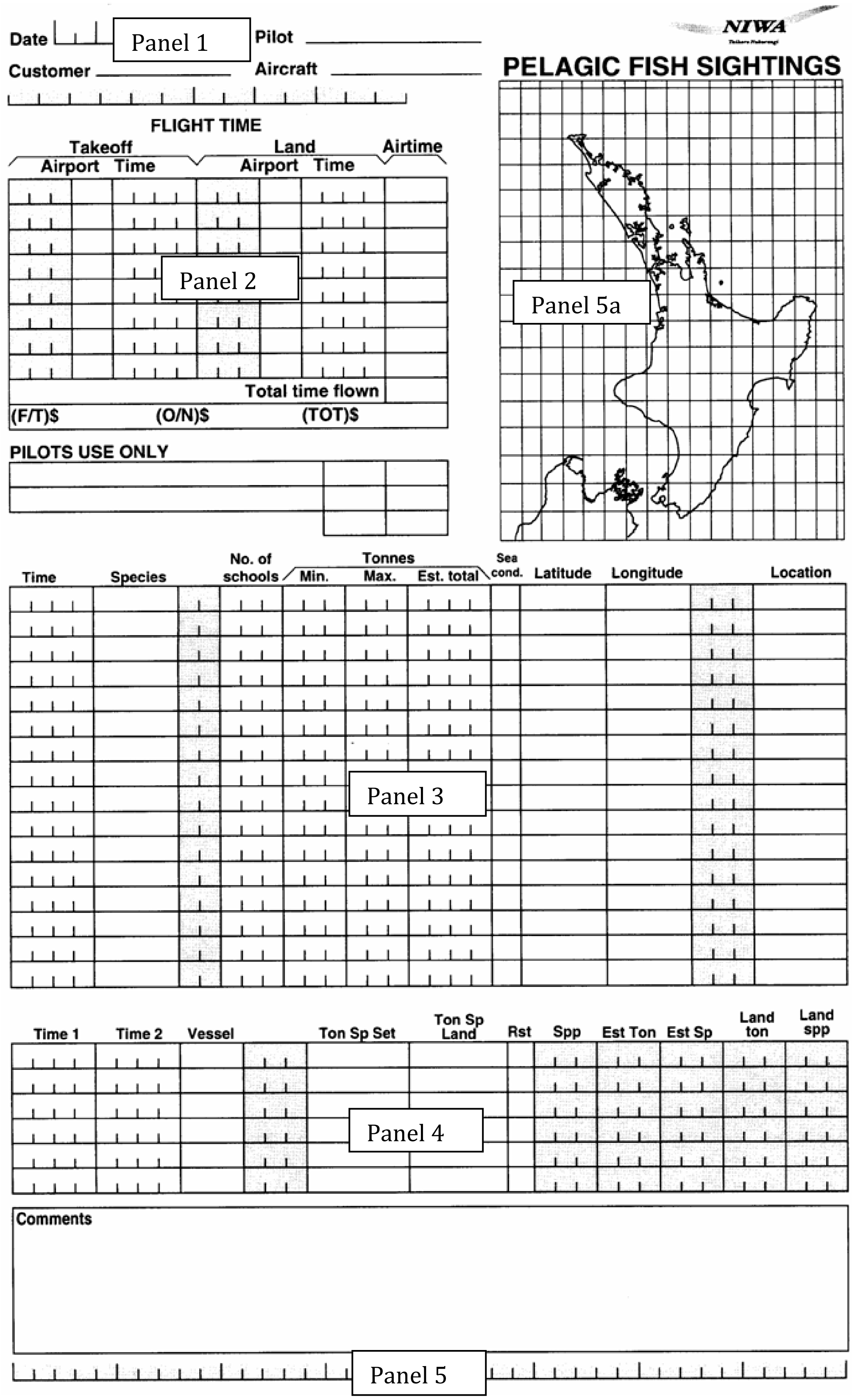
Date
Panel 1
Pilot
Customer
Aircraft
NIWA
PELAGIC FISH SIGHTINGS
FLIGHT TIME
Takeoff
Airport Time
Land
Airport Time
Airtime
Panel 2
Panel 5a
Total time flown
(F/T)$ (O/N)$ (TOT)$
PILOTS USE ONLY
Time Species No. of schools Tonnes Min. Max. Est. total Sea cond. Latitude Longitude Location
Panel 3
Time 1 Time 2 Vessel Ton Sp Set Ton Sp Land Rst Spp Est Ton Est Sp Land ton Land spp
Panel 4
Comments
Panel 5

Figure 1: The aerial sightings data-collection form — an explanation of the panels is included in Appendix B.

Table 1: Vessels appearing in the FSU-new database, the *aer_sight* database, their ownership, and an indication of Pilot #2 spotting for them.

FSU dataset	*Aer_sight*	Owner	Pilot #2 spotted
Lindberg	Lindberg	Sanford	✓
San Columbia	Columbia	Sanford	✓
Western Ranger	Western Ranger	Vela/Watties	✓
Waihola	Waihola	Sanford	✓
Michelangelo	NA	USA	x
Kerri M	NA	USA	x
South Pacific	NA	USA	x
Finisterre	NA	USA	x
Adriatic Sea	NA	USA	x
Zapata Discoverer	NA	USA	x
Voyager	NA	USA	x
San Benito	Benito	Sanford	✓
Marine Countess	NA	Vela	x
Janet D	NA	Vela	x
Shemara	Shemara	Sealords	✓
Island Princess	NA	USA	x
Lone Wolf	NA	USA	x
Western Pacific	Western Pacific	USA/Vela	✓
Tortugas	San Tortugas	Sanford	✓
Western Pacific (NZ)	Western Pacific	Vela	✓
Tifaimona	NA	Vela	x
Frontier	NA	USA	x
Jeanette C	NA	USA	x
Royal Pacific	NA	USA	x
Capt M J Souza	Souza	USA	x
Pacific Princess	Pacific Princess	USA	x
Capt Frank Medina	NA	USA	x
White Star	NA	USA	x
Cindy Ann	NA	USA	x
Montana	NA	USA	x
Western Pacific (US)	NA	USA	x

2.2.7 Years with missing data (1989; 1995–1997)

Because the Fisheries Statistics Unit (FSU) was disbanded in 1989 and a hiatus occurred before the new group managing fisheries data was fully operational, most purse-seine catch-effort data (along with most other fisheries data) are unavailable for 1989. Consequently, there were no surrogate target data available for 1989 and that year was dropped from the dataset.

Between 1995 and 1997 (inclusive), flightpath data were unavailable for Pilot #2, so flights exclusive to the BoP could not be identified for this period. Consequently, data from *aer_sight* for these years could not be included in the analysis.

2.2.8 Omitted years (1985–86; 2006–07)

During discussion at the 29 March 2012 and 16 April 2012 meetings, the NINSWG identified three indices that appeared to be outliers, probably as a result of the low numbers of flights that were the basis for these estimates. Consequently, the NINSWG requested that the first year in the series for both kahawai and trevally be dropped from the analysis as well as the 2006–07 year in the kahawai series.

2.3 Adjusting effort and calculating the time of day variable

Fundamental to the analysis was inclusion of flying effort. Work under SAP2006-10 included the adjustment of flying effort to allow for process time, which is represented in the data by the number of fishing operations and the total number of sightings of all species (Taylor 2014). In other words, flying time was adjusted for portions of flights that were not search time. To accommodate the process time idea, flight time (*feff*) was regressed against both the number of operations (*nops*) and the total sightings (*totsit*),

$$\text{feff} = b * \text{nops} + c * \text{totsit}.$$

The estimated slopes from this regression were used to adjust flight time into search time (*efft*) for the lognormal and the binomial regressions,

$$\text{efft} = \text{feff} - \text{nops} * b - \text{totsit} * c.$$

Time-of-day in decimal hours (called *dchr* in the analysis) was calculated as the time at the mid-point of the flight using the takeoff time plus the flying effort (i.e., flight time) divided by 2.

2.4 Density of flights by target species

Flight density plots were created for each target species using the flight path information to determine the extent to which the BOP was covered by flights, and whether the area flown varied with target species. This was achieved by extracting effort as the number of 10–15 min periods by grid square for each day that target data were available, and plotting, for each target species, the proportion of total effort spent searching for that species by grid square as expanding circle plots on a background of grid squares.

2.5 Producing the standardised indices — the analysis

The analysis was carried out using the generalised additive model (GAM) (Hastie & Tibshirani 1990) within the *R* package *mgcv* (Wood 2006) following a two-component approach based on the SAP2006-10 analysis (Taylor 2014). For the first component, a binomial fit was used to standardise the presence-absence of schools of the species of interest (trevally or kahawai) on the flight; for the second a lognormal fit was used to standardise observed tonnages of each species.

Because the aim was to produce annual indices of relative abundance, fishing year (*fsyr*) (categorical) was forced into all model runs at the start. A forward stepwise approach was used to include other explanatory variables, and models were constrained to include explanatory variables accounting for at least 3% of the variability (i.e., those increasing the R^2 by no less than 3%). In addition to effort (continuous), six explanatory variables were offered to each of the model runs (Table 2).

Table 2: Explanatory variables used in the regressions.

Explanatory variable	Abbreviated name	Description	Variable type
Fishing year	fsyr	Fishing year	categorical
Adjusted effort	efft	Flying effort adjusted for operations and sightings	continuous
Unadjusted effort	feff	Flying effort	continuous
Month	cmth	Calendar month	†continuous
Time of day	dchr	Decimal hour of the day	continuous
Southern oscillation index	soi	Troup's index — monthly values	continuous
Sea surface temperature	sst	Daily temperature collected at Leigh	continuous
Target species	targt	Species with maximum catch from the purse-seine fleet	categorical
Moonphase	moon	Proportion of the disc illuminated	continuous

†Calendar month must be included as continuous data for the cyclic smoother to function.

Terms were added to the model as follows:

y ~ fsyr + s(cmth,bs="cc") + s(efft) + s(sst) + s(soi) + targt + s(dchr) + s(moon)

where "s()" is a smoother and "cc" specifies a cyclic smoother. For the binomial model, *y* was tonnage>0 (i.e., 1 or 0), and for the lognormal model, *y* was log(tonnage). A separate analysis was done for each species (trevally and kahawai).
The final step in the analysis was to combine the year effects from the lognormal and binomial fits to produce a set of annual relative abundance indices for each of the two species.

2.6 Further analyses

At a meeting of the NINSWG on 29 March 2012, following presentation of the results of the original analysis, the working group identified several points of interest where additional plots would be useful in clarifying and resolving uncertainties arising from the original outputs. We report on the follow items.

2.6.1 Investigating apparent inconsistencies in the binomial indices

The WG identified that the binomial index for kahawai appeared to be inconsistent with the data presented on the annual numbers of flights that recorded kahawai and no kahawai and requested plots of the raw annual proportions of flights with zero kahawai sightings to resolve this apparent inconsistency.

2.6.2 Evaluating the SPUE indices — Kahawai

The WG determined that plots showing the annual proportions of flights with zero kahawai sightings, as well as the catch history and total allowable commercial catch (TACC) changes for kahawai, would be useful for evaluating the SPUE index as an index of abundance. The WG also determined that a comparison of the kahawai SPUE with recreational catch per unit effort (CPUE) from boat ramp surveys would be useful in this regard and requested that these be presented.

2.6.3 Evaluating the SPUE indices — Trevally

The WG determined that presentation of the catch history for TRE 1, along with the standardized SPUE indices, would assist in evaluating whether the indices reflect relative abundance.

2.6.4 Comparing trends of the trevally lognormal and binomial indices

The WG determined that models of positive (lognormal) and non-zero (binomial) sightings of trevally appeared to have similar trends and requested that the binomial and lognormal curves be plotted on the same set of axes for comparative purposes.

2.6.5 Investigating the low explanatory power of effort in the models

Effort seemed to have little explanatory power in the models. This may have been because most flights were of similar duration. The WG requested a figure showing the distribution of flight duration (adjusted to remove time associated with assisting fishing operations) to examine this suggestion.

3. RESULTS

3.1 Data extracts and processing

This section provides results for data extracts and processing as is appropriate. Outputs from the initial steps are useful and are presented. Processing at subsequent steps either add nothing in terms of data summaries (e.g., Step 3A), or are similar to the final datasets (see Sections 3.1.5 and 3.1.6) and are not included.

3.1.1 Testing the modified code

The tests of the targetInv code showed two variations (Table 3) in the output of *flt_grp*s on days when modal target data were available compared with the records extracted by Middleton et al (2010). The first occurred as a result of a modification to the code to avoid including four erroneous records: one each for target species octopus (OCT), garfish (GAR), paddle crab (PAD) and one day of flying that was on the west coast North Island and not in the BoP (i.e., included in mixed category in Table 3). The second variation occurred as a result of new data that had been added for the period June 1998 to July 2009, since the SAP2006-10 analysis had been completed. These data (i.e., "Additional data" in Table 3) included 2, 42, and 17 records for blue mackerel, jack mackerel, and skipjack tuna respectively, and were from data collection forms that had been submitted by Pilot #2 after the SAP2006-10 analysis was completed.

3.1.2 Extract step 1A

The extracts for Step 1A (see Figure A1) included 14 700 records of flight data (flight group, flight number, flight index, takeoff time, takeoff airfield, landing time, landing airfield, length of flight in decimal hours, date), 76 158 records of flightpath data (flight group, grid square), 15 962 records of purse-seine setting/operation data (flight group, sighting time, set time, vessel code, result of set, species sighted, pilot's estimate of tonnage, pilot's estimate of species composition, landed tonnage, landed species composition), 75 324 records of school data (flight group, flight number, location/grid square), and ID lists for species and vessels. Flight group is a group of flights recorded on a single day, flight number is the integer code expressing chronological order of the flights in a flight group, and flight index is a code for the flight group and flight number of any flight.

Table 3: Number of flights by modal target species, extracted for SAP200610 and with revised code for SEA201017; additional data represents updates to *aer_sight* since the SAP200610 analysis was completed;

Target species	Species code	SAP200610	SEA101017
Octopus	OCT	1	0
Blue mackerel	EMA	58	58
Garfish	GAR	1	0
Jack mackerel	JMA	227	227
Kahawai	KAH	51	51
Mixed	MIX	47	46
Paddle crab	PAD	1	0
Pilchard	PIL	11	11
Skipjack tuna	SKJ	127	127
Trevally	TRE	15	15
	Totals	539	535
Additional data			
Blue mackerel	EMA		2
Jack mackerel	JMA		42
Skipjack tuna	SKJ		17
	Totals	0	61

3.1.3 Extract steps 1B and 1C

A total of 25 442 and 6518 catch records were available for the period 1975–76 to 2010–11 from the *warehou* and *fsu_new* databases respectively (i.e., a grand total of 31 960). Of these, 1804 records could not be linked to a date, so a total of 30 156 catch records with associated dates were combined to produce a dataset of catches for the domestic purse-seine fleet between 1975–76 and 2010–11 (Table 4). Between 1975–76 and 1982–83, data were sparse, with many months showing no recorded catches. A total of 27 344 records of effort data were available from the *warehou* and *fsu_new* databases for the domestic purse-seine fleet between 1975–76 and 2010–11 (Table 5); all had dates associated with them. As was noted for the catch data, the effort dataset was characterised by many missing months between 1975–76 and 1982–83. The hiatus in data collection that accompanied the change-over from FSU management of catch-effort data in 1988–89 (see Section 2.2.7) was characterised by the absence of records for most months of that year in both the catch dataset and the effort dataset.

Table 4: Number of available records of catch data with associated dates for the domestic purse-seine fleet between 1975–76 and 2010–11. Source: *warehou* and *fsu_new*.

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1975–76	0	0	0	0	8	1	0	0	0	0	0	0	9
1976–77	0	0	0	0	24	17	4	0	0	0	0	0	45
1977–78	0	0	0	13	18	6	0	0	0	0	0	0	37
1978–79	0	3	0	24	12	11	1	0	0	0	0	0	51
1979–80	0	1	2	18	20	22	0	6	5	0	0	0	74
1980–81	0	6	26	39	27	34	12	0	0	0	0	0	144
1981–82	0	0	0	13	27	72	32	0	0	0	0	0	144
1982–83	0	0	0	62	64	82	62	20	29	27	13	19	378
1983–84	36	60	36	69	84	73	66	42	28	42	36	28	600
1984–85	57	81	54	91	54	31	28	35	39	22	26	44	562
1985–86	70	94	43	99	80	86	75	55	29	11	13	29	684
1986–87	45	67	66	55	133	101	97	95	50	18	17	31	775
1987–88	67	117	100	119	101	100	75	146	148	87	50	88	1198
1988–89	13	0	0	0	0	0	0	0	0	0	0	0	13
1989–90	37	69	41	91	114	105	67	99	52	20	17	60	772
1990–91	94	92	60	90	106	76	61	116	38	40	34	59	866
1991–92	77	75	69	69	51	66	73	39	20	17	14	30	600
1992–93	55	111	76	83	104	100	57	102	78	89	21	47	923
1993–94	151	163	95	92	46	77	143	98	71	44	76	70	1126
1994–95	110	124	100	105	74	64	89	81	71	71	64	50	1003
1995–96	144	120	70	98	108	130	88	69	45	81	58	53	1064
1996–97	138	104	60	78	78	69	91	78	23	60	98	78	955
1997–98	88	119	55	131	101	82	50	55	27	20	47	121	896
1998–99	101	104	64	110	69	85	73	113	48	94	101	89	1051
1999–00	75	121	87	110	158	130	56	67	38	16	96	94	1048
2000–01	78	136	102	143	135	158	31	37	58	48	91	153	1170
2001–02	193	160	100	235	111	136	162	145	133	204	160	211	1950
2002–03	160	225	179	165	199	143	148	73	136	99	90	171	1788
2003–04	178	222	163	215	90	239	186	169	169	79	62	77	1849
2004–05	113	181	53	109	109	157	177	122	53	10	84	117	1285
2005–06	114	107	92	143	134	112	54	58	54	49	104	145	1166
2006–07	165	115	68	120	145	86	195	40	25	41	59	145	1204
2007–08	125	140	90	149	217	110	58	60	84	51	52	152	1288
2008–09	149	102	89	196	79	41	22	37	23	56	44	104	942
2009–10	118	127	83	139	162	86	70	35	14	77	57	88	1056
2010–11	171	162	136	164	254	132	124	50	49	42	59	97	1440
Totals	2922	3308	2259	3437	3296	3020	2527	2142	1637	1515	1643	2450	30156

Table 5: Number of available records of effort data for the domestic purse-seine fleet between 1975–76 and 2010–11. Source: *warehou* and *fsu_new*.

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1975–76	0	0	0	1	12	6	0	0	0	0	0	0	19
1976–77	0	0	0	6	46	36	6	0	0	0	0	0	94
1977–78	0	0	0	36	31	21	0	0	0	0	0	0	88
1978–79	0	6	0	63	55	31	1	0	0	0	0	0	156
1979–80	0	5	7	61	53	30	0	10	11	0	0	0	177
1980–81	0	15	38	61	43	77	32	3	0	0	0	0	269
1981–82	0	0	0	27	58	120	51	0	0	0	0	0	256
1982–83	0	0	0	93	102	96	71	28	31	48	51	29	549
1983–84	40	47	50	90	115	82	57	54	32	43	38	35	683
1984–85	76	108	53	107	67	21	28	24	29	34	51	67	665
1985–86	80	79	36	115	86	122	73	66	22	15	20	40	754
1986–87	62	57	59	88	148	96	98	111	53	36	51	44	903
1987–88	87	133	97	156	128	119	75	127	120	111	87	137	1377
1988–89	23	0	0	0	0	0	0	0	0	0	0	0	23
1989–90	21	35	33	96	99	87	37	57	51	14	19	58	607
1990–91	65	65	38	74	91	77	46	77	38	43	34	47	695
1991–92	73	72	43	52	35	55	61	29	24	21	18	31	514
1992–93	47	67	44	45	63	69	36	58	60	79	18	47	633
1993–94	106	134	58	99	72	81	111	74	65	48	87	68	1003
1994–95	130	113	72	69	63	61	59	56	58	65	70	41	857
1995–96	117	111	60	102	92	114	53	41	35	72	51	39	887
1996–97	101	88	50	96	96	69	69	86	19	52	77	71	874
1997–98	66	113	43	155	113	101	44	34	25	12	39	94	839
1998–99	85	99	71	129	86	94	56	88	41	71	73	53	946
1999–00	57	102	89	128	189	172	45	48	33	16	97	79	1055
2000–01	69	106	81	139	159	174	24	24	55	61	90	127	1109
2001–02	128	97	56	176	117	84	76	80	86	133	104	169	1306
2002–03	152	181	115	131	165	106	94	56	78	64	62	109	1313
2003–04	120	147	96	160	59	198	183	97	84	63	66	99	1372
2004–05	115	131	59	104	155	168	159	75	37	11	76	85	1175
2005–06	73	83	83	148	174	123	52	34	52	61	119	132	1134
2006–07	144	126	53	173	135	79	117	45	27	25	54	122	1100
2007–08	103	107	96	133	206	127	29	43	51	44	44	93	1076
2008–09	96	73	49	158	67	34	14	19	21	50	40	98	719
2009–10	118	112	68	110	137	76	69	34	12	62	48	79	925
2010–11	141	122	83	128	198	104	86	56	47	57	70	100	1192
Totals	2495	2734	1780	3509	3515	3110	2012	1634	1297	1411	1654	2193	27344

3.1.4 Extract step 2

A total of 2186 flights exclusive to the BoP on days with surrogate target species data were initially available throughout the study period (1985–86 to 2010–11) (Table 6). A total of 960 flights exclusive to east Northland on days with surrogate target species data were initially available throughout the study period (1985–86 to 2010–11) (Table 7). These totals were subsequently reduced as various omissions were necessary, particularly with the absence of flying effort between 1994–95 and 1996–97.

Table 6: Total number of flights in the Bay of Plenty on days with surrogate target species (catch) available.

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1985–86	17	5	2	14	14	16	10	14	8	15	10	7	132
1986–87	9	15	3	4	15	16	13	21	13	15	20	17	161
1987–88	14	21	4	10	2	12	12	17	18	24	21	21	176
1988–89	18	11	0	8	13	22	14	18	17	20	10	14	165
1989–90	16	20	12	8	0	16	11	21	19	0	19	13	155
1990–91	6	13	9	15	11	10	20	22	16	16	12	12	162
1991–92	17	20	12	21	18	16	21	17	17	15	12	18	204
1992–93	8	8	6	5	2	2	0	18	18	22	6	21	116
1993–94	1	20	5	0	8	13	11	8	4	2	3	6	81
1994–95	12	13	10	5	5	6	7	4	11	14	2	10	99
1995–96	5	7	1	2	4	9	6	5	11	8	8	2	68
1996–97	1	1	2	5	10	9	6	13	5	8	7	3	70
1997–98	5	7	3	6	16	12	9	1	11	3	9	9	91
1998–99	13	9	3	4	1	6	3	9	9	12	11	2	82
1999–00	4	4	11	15	14	18	8	6	7	6	10	9	112
2000–01	17	6	7	7	9	0	1	1	0	6	10	11	75
2001–02	10	5	14	8	2	0	2	2	5	10	10	7	75
2002–03	6	5	3	3	8	2	1	0	3	13	9	11	64
2003–04	6	4	11	5	5	2	0	8	8	6	4	0	59
2004–05	5	8	3	3	2	4	1	3	8	5	13	12	67
2005–06	11	6	1	4	1	3	7	7	10	10	10	3	73
2006–07	1	0	4	0	5	3	0	9	9	8	4	5	48
2007–08	5	14	5	10	16	3	2	7	13	6	4	11	96
2008–09	7	11	11	14	5	0	4	4	15	8	7	12	98
2009–10	9	3	6	7	7	5	0	9	5	15	13	9	88
2010–11	15	8	5	6	0	0	1	12	9	11	5	0	72
Totals	238	244	153	189	193	220	185	277	288	302	270	262	2689

Table 7: Total number of flights in east Northland on days with surrogate target species (catch) available.

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1985–86	23	16	18	3	1	0	0	0	0	0	0	3	64
1986–87	1	0	0	1	0	2	4	3	0	1	0	0	12
1987–88	0	3	0	0	10	1	2	3	1	4	2	8	34
1988–89	3	1	3	0	0	2	1	4	1	2	2	1	20
1989–90	0	5	1	0	4	7	10	6	3	0	4	1	41
1990–91	1	9	3	9	4	2	11	12	10	9	11	4	85
1991–92	11	5	0	6	6	10	13	12	9	1	4	2	79
1992–93	0	12	5	4	3	3	1	9	4	4	3	1	49
1993–94	3	4	0	6	3	1	8	6	6	5	12	5	59
1994–95	3	6	10	4	1	2	2	1	5	0	0	0	34
1995–96	0	0	0	0	1	0	1	0	7	0	0	0	9
1996–97	3	8	3	11	8	3	2	0	2	6	2	0	48
1997–98	3	3	5	1	0	0	0	0	0	0	0	0	12
1998–99	0	3	4	3	0	0	0	0	0	0	1	2	13
1999–00	3	4	3	0	2	2	5	0	0	0	0	0	19
2000–01	4	9	5	1	2	0	0	0	0	0	0	7	28
2001–02	12	7	4	1	0	0	0	0	0	0	0	3	27
2002–03	18	18	5	0	0	0	0	0	0	0	0	0	41
2003–04	10	18	7	0	0	1	4	1	0	0	2	11	54
2004–05	13	5	7	0	1	3	2	0	0	0	2	0	33
2005–06	4	5	6	0	1	3	0	2	0	0	0	18	39
2006–07	23	12	3	6	5	2	0	0	0	1	0	4	56
2007–08	11	2	4	0	2	2	0	0	0	0	2	6	29
2008–09	4	6	3	2	1	0	0	0	0	1	0	3	20
2009–10	4	18	0	2	1	1	0	1	0	0	1	1	29
2010–11	1	11	4	1	3	1	2	0	0	1	2	0	26
Totals	158	190	103	61	59	48	68	60	48	35	50	80	960

3.1.5 The east Northland dataset

Four data summaries to determine the volume of available data for east Northland were completed.

- The number of positive flights for kahawai (i.e., the number of flights for which surrogate target data were available and on which kahawai were sighted) recorded by Pilot #2 between 1986–87 and 2010–11 are shown in Table 8.
- The number of zero flights for kahawai (i.e., the number of flights for which surrogate target data were available and on which kahawai were not sighted) recorded by Pilot #2 between 1986–87 and 2010–11 are shown in Table 9.
- The number of positive flights for kahawai recorded by all pilots between 1986–87 and 2010–11 are shown in Table 10.
- The number of zero flights for kahawai recorded by all pilots between 1986–87 and 2010–11 are shown in Table 11.

Table 8: Kahawai – number of positive flights (Pilot #2 only).

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1986–87	1	0	0	0	0	0	0	0	0	0	0	0	1
1987–88	0	0	0	0	0	0	0	1	1	0	1	0	3
1989–90	0	0	0	0	0	0	0	0	0	0	0	0	0
1990–91	0	0	0	0	1	0	1	1	0	0	0	1	4
1991–92	0	0	0	0	0	0	0	3	0	0	0	0	3
1992–93	0	0	0	0	1	0	0	0	1	1	0	0	3
1993–94	0	0	0	0	0	0	0	0	0	0	0	1	1
1997–98	0	0	0	1	0	0	0	0	0	0	0	0	1
1998–99	0	0	0	0	0	0	0	0	0	0	0	1	1
1999–00	0	0	0	0	0	0	0	0	0	0	0	0	0
2000–01	1	1	1	0	0	0	0	0	0	0	0	1	4
2001–02	0	2	0	1	0	0	0	0	0	0	0	0	3
2002–03	1	0	0	0	0	0	0	0	0	0	0	0	1
2003–04	2	3	1	0	0	0	0	0	0	0	0	0	6
2004–05	2	1	2	0	0	1	0	0	0	0	0	0	6
2005–06	0	3	0	0	0	0	0	0	0	0	0	8	11
2006–07	3	1	0	1	0	0	0	0	0	0	0	0	5
2007–08	2	0	1	0	0	0	0	0	0	0	0	0	3
2008–09	2	2	1	0	0	0	0	0	0	0	0	0	5
2009–10	2	4	0	0	0	0	0	0	0	0	0	0	6
2010–11	0	1	1	0	0	0	0	0	0	0	1	0	3
Totals	16	18	7	3	2	1	1	5	2	1	2	12	70

Table 9: Kahawai – number of zero flights (Pilot #2 only).

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1986–87	0	0	0	0	0	0	0	0	0	0	0	0	0
1987–88	0	0	0	0	1	0	0	0	0	0	0	0	1
1989–90	0	0	0	0	4	2	0	0	0	0	0	1	7
1990–91	1	2	0	1	1	1	1	0	0	0	0	1	8
1991–92	5	2	0	0	0	0	1	0	0	0	0	0	8
1992–93	0	0	0	0	0	0	0	0	0	0	0	0	0
1993–94	1	1	0	0	0	0	3	1	5	4	10	1	26
1997–98	1	2	1	0	0	0	0	0	0	0	0	0	4
1998–99	0	1	2	2	0	0	0	0	0	0	0	0	5
1999–00	0	1	0	0	0	0	0	0	0	0	0	0	1
2000–01	0	4	2	1	1	0	0	0	0	0	0	3	11
2001–02	5	2	0	0	0	0	0	0	0	0	0	2	9
2002–03	4	7	4	0	0	0	0	0	0	0	0	0	15
2003–04	0	2	0	0	0	0	2	0	0	0	0	5	9
2004–05	2	0	2	0	1	1	1	0	0	0	0	0	7
2005–06	0	0	0	0	0	1	0	0	0	0	0	1	2
2006–07	5	3	0	1	2	0	0	0	0	0	0	4	15
2007–08	3	0	3	0	0	0	0	0	0	0	2	2	10
2008–09	2	1	0	1	1	0	0	0	0	0	0	3	8
2009–10	1	4	0	0	0	0	0	0	0	0	0	0	5
2010–11	0	2	1	0	0	0	0	0	0	0	1	0	4
Totals	30	34	15	6	11	5	8	1	5	4	13	23	155

Table 10: Kahawai – number of positive flights (all pilots).

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1986–87	1	0	0	0	0	0	0	0	0	0	0	0	1
1987–88	0	0	0	0	1	0	0	1	1	0	1	2	6
1989–90	0	0	0	0	0	0	0	0	0	0	0	0	0
1990–91	0	0	1	1	2	0	1	1	0	0	0	1	7
1991–92	0	0	0	0	0	0	0	3	0	0	0	0	3
1992–93	0	1	1	0	1	0	0	0	1	1	0	0	5
1993–94	0	0	0	0	0	0	0	0	0	0	0	1	1
1994–95	0	2	3	0	0	0	0	0	0	0	0	0	5
1995–96	0	0	0	0	0	0	0	0	0	0	0	0	0
1996–97	1	0	0	0	0	0	0	0	0	0	0	0	1
1997–98	0	0	0	1	0	0	0	0	0	0	0	0	1
1998–99	0	2	0	0	0	0	0	0	0	0	0	1	3
1999–00	1	0	0	0	0	0	0	0	0	0	0	0	1
2000–01	2	2	1	0	0	0	0	0	0	0	0	1	6
2001–02	0	2	0	1	0	0	0	0	0	0	0	0	3
2002–03	1	0	0	0	0	0	0	0	0	0	0	0	1
2003–04	2	3	1	0	0	0	1	0	0	0	0	0	7
2004–05	3	3	2	0	0	1	0	0	0	0	0	0	9
2005–06	1	4	2	0	0	0	0	0	0	0	0	11	18
2006–07	3	1	0	1	0	0	0	0	0	0	0	0	5
2007–08	2	0	1	0	0	0	0	0	0	0	0	0	3
2008–09	2	2	1	0	0	0	0	0	0	0	0	0	5
2009–10	2	4	0	0	0	0	0	0	0	0	1	0	7
2010–11	0	1	2	0	0	0	0	0	0	0	1	0	4
Totals	21	27	15	4	4	1	2	5	2	1	3	17	102

Table 11: Kahawai – number of zero flights (all pilots).

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1986–87	0	0	0	0	0	2	1	0	0	0	0	0	3
1987–88	0	0	0	0	4	1	0	1	0	1	1	1	9
1989–90	0	0	0	0	4	4	0	0	0	0	0	1	9
1990–91	1	2	1	3	1	1	2	1	0	0	0	2	14
1991–92	9	2	0	1	0	1	1	0	0	0	0	0	14
1992–93	0	5	1	1	2	0	0	0	0	0	0	0	9
1993–94	3	2	0	4	0	1	5	1	5	4	11	4	40
1994–95	3	3	4	2	1	0	1	1	3	0	0	0	18
1995–96	0	0	0	0	1	0	0	0	0	0	0	0	1
1996–97	0	6	2	8	6	1	1	0	2	0	1	0	27
1997–98	1	3	2	0	0	0	0	0	0	0	0	0	6
1998–99	0	1	4	3	0	0	0	0	0	0	0	0	8
1999–00	1	4	2	0	2	0	3	0	0	0	0	0	12
2000–01	0	7	4	1	1	0	0	0	0	0	0	6	19
2001–02	9	5	2	0	0	0	0	0	0	0	0	2	18
2002–03	15	16	5	0	0	0	0	0	0	0	0	0	36
2003–04	2	11	5	0	0	0	3	0	0	0	2	7	30
2004–05	9	2	3	0	1	1	1	0	0	0	0	0	17
2005–06	0	1	4	0	1	2	0	0	0	0	0	5	13
2006–07	15	7	2	3	4	0	0	0	0	0	0	4	35
2007–08	8	2	3	0	1	0	0	0	0	0	2	6	22
2008–09	2	1	2	2	1	0	0	0	0	1	0	3	12
2009–10	2	14	0	2	0	0	0	0	0	0	0	0	18
2010–11	0	8	2	1	1	0	2	0	0	0	1	0	15
Totals	80	102	48	31	31	14	20	4	10	6	18	41	405

Aerial sightings data for east Northland were considered to be insufficient to proceed to a standardisation analysis. As a rule of thumb it was considered that, for the analysis to have a reliable basis there should be data available from an average of about 50 or more flights per year. Even when data from all pilots was considered, the east Northland dataset was still insufficient and so the analysis was abandoned.

3.1.6 The final BoP dataset

Data from 1289 of first flights of the day exclusive to the BOP on days that surrogate target data were available during the period of interest were extracted from *aer_sight*.

The total number of flights per year has varied between 21 in 2006–07 and 113 in 1991–92 (Table 12), although numbers were 35 or more in all years for which data were available except one, and 45 or more in 17 of the 22 years for which data were available. The mean of monthly totals over all years was 107 flights. The lowest monthly total over all years was 60 flights in April with 9 years contributing flights to the total, and only 3 years contributing flights since the break in the mid 1990s. The highest monthly mean of 6 flights occurred in July, September, and November, when the monthly totals (over the entire dataset) were 133, 135, and 125 flights respectively.

The number of sightings varies markedly between the two species of interest (Table 13) with a maximum grand total for kahawai of 2978 and 556 for trevally.

Table 12: Distribution of flights exclusive to the Bay of Plenty throughout the data period (1985–86 to 2010–11), by fishing year and month.

Fishing year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Totals
1985–86		no data		11	13	8	4	5	1	2	0	2	46
1986–87	2	1	1	3	14	3	0	6	4	3	2	8	47
1987–88	0	8	1	4	0	6	3	2	3	17	5	13	62
1988–89													no data
1989–90	6	14	6	3	0	13	7	14	17	0	5	11	96
1990–91	3	6	6	9	9	6	9	19	11	10	4	1	93
1991–92	4	11	7	13	11	11	18	5	9	11	6	7	113
1992–93	1	5	5	1	0	1	0	12	14	19	2	14	74
1993–94	0	18	3	0	1	5	9	4	2	0	0	1	43
1994–95													no data
1995–96													no data
1996–97													no data
1997–98	no data		3	3	15	11	5	0	4	2	3	6	52
1998–99	10	7	0	1	0	0	0	8	6	9	8	1	50
1999–00	2	3	7	12	13	18	2	2	4	0	7	8	78
2000–01	13	4	4	6	7	0	0	0	0	5	8	8	55
2001–02	5	2	6	7	1	0	0	1	2	4	5	4	37
2002–03	1	1	0	3	7	0	0	0	0	8	6	9	35
2003–04	6	2	7	5	2	1	0	6	3	2	1	0	35
2004–05	3	5	0	2	0	3	0	2	7	0	10	10	42
2005–06	7	3	0	4	0	1	3	5	6	6	9	2	46
2006–07	0	0	3	0	3	1	0	6	4	2	0	2	21
2007–08	3	12	2	10	15	1	0	4	11	5	3	9	75
2008–09	6	8	10	13	4	0	0	1	4	7	5	12	70
2009–10	6	0	6	6	4	2	0	5	4	14	8	8	63
2010–11	12	7	2	5	0	0	0	11	8	7	4	0	56
Totals	90	117	79	121	119	91	60	118	124	133	101	136	1289

An apparent discrepancy is evident for both kahawai and trevally between the number of sightings in the dataset used for the lognormal fit and the number in the dataset used for the binomial fit. These differences are highlighted in Table 13. In all cases the binomial dataset contains more sightings than the lognormal dataset. The differences are the result of missing mixed tonnage data on the original data collection forms. In the binomial dataset, it is the presence of a mixed school containing the species of interest that is required. These data are available from the *t_school_sight* table in the aerial sightings database. However, tonnages related to sightings of mixed schools are sometimes recorded incorrectly (e.g., two tonnages for a mixed school containing three species) or are missing altogether and, because a tonnage of the species of interest is required for each sighting in the lognormal component of the analysis, some mixed-school sightings (50 for kahawai; 20 for trevally) are missing from the lognormal dataset.

Table 13: Number of sightings of kahawai and trevally comprising the datasets used in the lognormal and binomial fits, by fishing year and month; highlighted cells indicate discrepancies in number of sightings between the two datasets for each species (see text).

Species	Dataset	Year	O	N	D	J	F	M	A	M	J	J	A	S
Kahawai	Lognormal	1985–86				9	3		1	11	3	5		7
		1986–87	11	3	7		3	1		19	9	9	3	26
		1987–88		33	5			4	3	3	8	71	19	31
		1989–90	23	25	4			8		42	61		6	27
		1990–91	7	11	9	1	3	5	10	72	20	19	6	
		1991–92	6	1	4	3	17	12	72	16	24	30	10	18
		1992–93		16	30	9				18	45	77	4	18
		1993–94		77	1			4	17	16	7			4
		1997–98			13	2	14	10	9		10	6	13	28
		1998–99	12	14						11	4	14	11	3
		1999–00	5	12	8	12	2	5	1	4	12		24	18
		2000–01	42	8	7	8	12					7	15	33
		2001–02	19	7	13	7						7	18	17
		2002–03	4	2		3	13					2	11	18
		2003–04	22	6	25	8	4			7	7	3		
		2004–05	12	18		4					16		38	27
		2005–06	49	12		4			11	14	23	9	20	7
		2006–07			7					24	7	4		15
		2007–08	14	42	12	13	20			15	38	24	11	35
		2008–09	28	39	29	4				1	7	17	12	46
		2009–10	16		34	10	4	3		17	13	21	19	45
		2010–11	54	39	5	9				20	14	15	8	
		Totals	324	365	213	106	95	52	124	310	328	340	248	423
		Grand total												2928

Table 13: continued

Species	Dataset	Year	O	N	D	J	F	M	A	M	J	J	A	S
	Binomial	1985–86				9	3		2	11	3	5		7
		1986–87	11	3	7		3	1		19	9	9	3	26
		1987–88		33	5			4	3	3	9	72	19	31
		1989–90	25	26	4			8		42	61		6	27
		1990–91	7	12	9	1	3	5	10	72	20	19	6	
		1991–92	6	1	7	5	18	13	72	16	24	30	10	18
		1992–93		16	31	9				18	45	77	4	18
		1993–94		76	1			4	21	17	7			4
		1997–98			13	2	14	10	9		10	6	13	28
		1998–99	12	14						11	4	14	11	3
		1999–00	5	12	8	12	2	5	1	4	12		24	18
		2000–01	42	10	7	8	12					7	15	33
		2001–02	20	7	13	7						7	18	17
		2002–03	4	2		3	14					2	11	18
		2003–04	22	6	26	8	4			7	7	3		
		2004–05	12	20		4					16		38	27
		2005–06	49	12		4			11	14	23	9	20	7
		2006–07			7					24	7	4		15
		2007–08	14	43	13	14	20			15	38	24	11	35
		2008–09	28	39	47	4				1	7	17	12	46
		2009–10	16		34	10	4	3		17	13	21	19	45
		2010–11	56	39	5	9				20	15	15	8	
		Totals	329	371	237	109	97	53	129	311	330	341	248	423
		Grand total												2978
Trevally	Lognormal	1985–86					2			1		2		
		1986–87		1	1			1		13	6	3		10
		1987–88		4				5			3	17	7	20
		1989–90		3	2			4	5	14	35		3	4
		1990–91		1		3				26	6	5		
		1991–92	2	1	3	2	3		26	4	5	2	1	
		1992–93			2					2	10	17	2	3
		1993–94						4	7	9	2			2
		1997–98			5		1		1		1			5
		1998–99	10	7						3	1	6	3	
		1999–00		3	2	7	5	2	1	1	3			3
		2000–01	12	1	4	3	2					1		5
		2001–02	14	1	3	3						1	1	1
		2002–03	2	4		2	5							
		2003–04	3		2	2	3			2	1			
		2004–05	4	5		2								2
		2005–06	2	2		2			3	1	3		2	
		2006–07			7									
		2007–08		5	1		1						1	
		2008–09	2		2								2	2
		2009–10			5	1	1	2					1	2
		2010–11	1	3	2								3	
		Totals	52	41	41	27	23	18	43	76	76	54	26	59
		Grand Total												536

Table 13: continued

Species	Dataset	Year	O	N	D	J	F	M	A	M	J	J	A	S
	Binomial	1985–86					2			1		2		
		1986–87		1	1			1		13	6	3		10
		1987–88		4				5			4	17	7	20
		1989–90		4	2			4	5	14	35		3	4
		1990–91		1		3				26	6	5		
		1991–92	2	1	3	3	3		26	4	5	2	1	
		1992–93			2					2	10	17	2	3
		1993–94						4	10	10	2			2
		1997–98			5		1		1		1			5
		1998–99	10	7						3	1	6	3	
		1999–00		3	2	7	5	2	1	1	3			3
		2000–01	12	2	4	3	2					1		5
		2001–02	15	1	3	3						1	1	1
		2002–03	2	4		2	5							
		2003–04	3		2	2	3			2	1			
		2004–05	4	7		2								2
		2005–06	2	2		2			3	1	3		2	
		2006–07			7									
		2007–08		6	1		1						1	
		2008-09	2		7								2	2
		2009–10			5	1	1	2				1	1	2
		2010–11	2	3	2						1		3	
		Totals	54	46	46	28	23	18	46	77	78	55	26	59
		Grand total												556

3.2 Density of flights in the BoP by target species

A particular spatial pattern dominated the flying effort distributions for most species (Figure 2) — the largest proportion of flying effort usually occurred in grid squares 147 and 164 with lesser amounts in grid squares 130 and 165, although that did not hold for kahawai which had greatest proportions in 147 and 165, and a little less in 164. The pattern for skipjack tuna also deviated from the general pattern. For all species there was flying in squares 112, 129, and 146, although in most cases the proportion of effort expended there was very low. However, it was clearly evident for pilchard in all three squares and for trevally in square 129.

Despite the "common" pattern, particularly for squares 147 and 165, there were subtle variations that characterised most species individually. For skipjack tuna the overall pattern contrasted markedly with that of the other species, with a greater proportion of coverage in the east resulting in a more dispersed coverage over a wider area.

3.3 Standardised indices for the BoP

3.3.1 Trevally — model fits (binomial and lognormal) and indices

The results of the trevally lognormal fits are shown in Table 14 and Figure 3. The final model is:

$$\log(tons) \sim fsyr + s(sst) + s(cmth, bs = \text{"cc"}).$$

Target species is not significant and the total amount of variability accounted for by the fit is relatively high at about 25.6%. The trend in the indices is variable with a marked drop in about 1997. There is no evidence of assumptions being violated in the diagnostic plots (Figure C1). Plots of partial effects are shown in Figure C2.

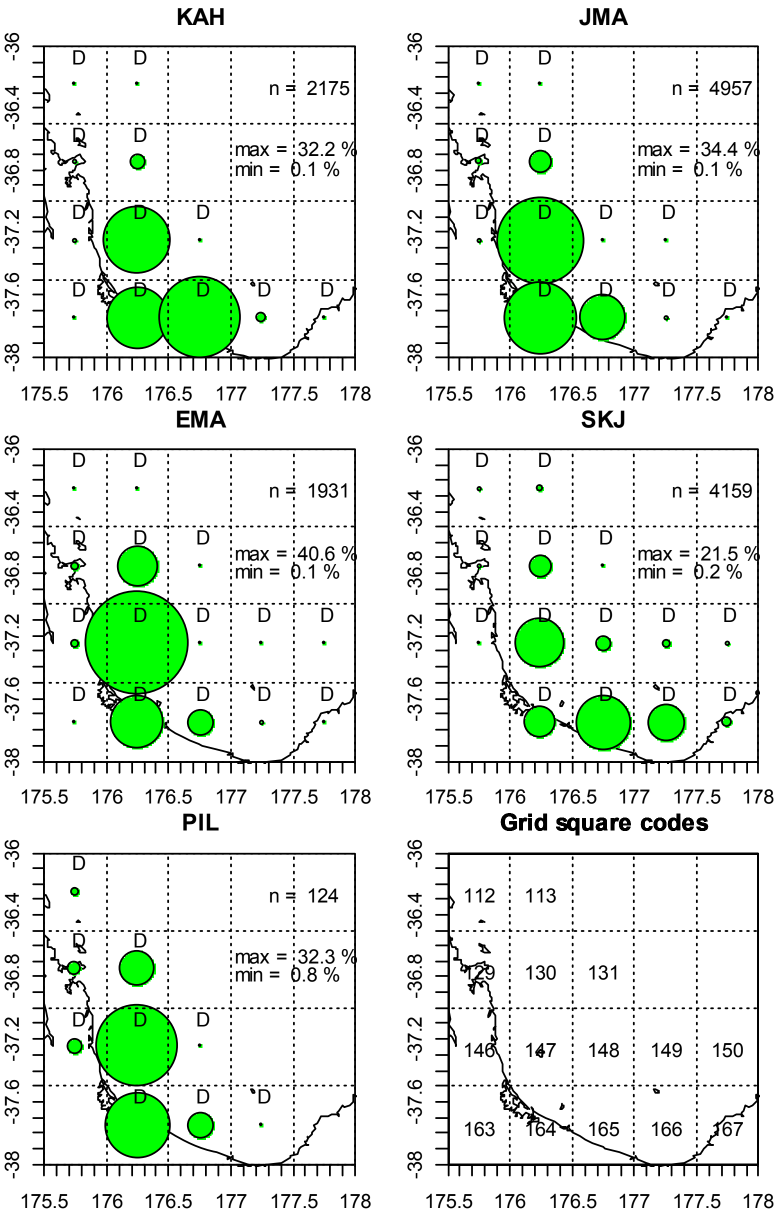

Figure 2: For each target species, flightpath density or the proportion of total flightpath ticks (10–15 min periods) recorded in each grid square visited during all flights throughout the period of interest (January 1986 to September 2011) in the Bay of Plenty; circles are centred on grid squares, their diameters are relative to proportions of ticks for that species, and the scale is constant for all plots; *n* is the total number of 10–15 min periods recorded during flights on days a particular target species was assigned, max is the largest proportion plotted for the relevant species, min is the smallest, D denotes squares where data were recorded for that species; EMA is blue mackerel (*Scomber australasicus*), JMA is jack mackerel (*Trachurus* species), KAH is kahawai (*Arripis trutta*), MIX refers to several minor target species, PIL is pilchard (*Sardinops neopilchardus*), SKJ is skipjack tuna (*Katsuwonas pelamis*), TRE is trevally (*Pseudocaranx dentax*); grid square codes are shown in the final plot for squares where data were recorded.

The results of the binomial model-fitting are shown in Table 14 and Figure 3. The final model is:

$$\text{tons}>0 \sim \text{fsyr} + s(\text{soi}) + s(\text{dchr}).$$

Table 14: Stepwise model fits (binomial and lognormal) for trevally; boldened rows indicate details of the final model in each case

	Predictor added	df	Deviance	AIC	R^2
Lognormal	fsyr	20.00	202	818	17.7
	s(sst)	25.74	192	812	22
	s(cmth, bs = "cc")	27.01	183	800	25.4
	targt	32.20	180	804	26.7
	s(efft)	33.88	178	804	27.4
	s(dchr)	36.09	176	805	28.3
	s(soi)	37.01	176	805	28.6
	s(moon)	37.99	175	807	28.6
Selected model	log(tons) ~ fsyr + s(sst) + s(cmth, bs = “cc”)				
Binomial	fsyr	20.00	1374	1414	4.8
	s(soi)	28.70	1323	1380	8.4
	s(dchr)	**37.34**	**1271**	**1345**	**12**
	targt	43.37	1236	1323	14.3
	+s(sst)	49.32	1211	1310	16.1
	+s(efft)	50.37	1207	1308	16.4
	+s(moon)	51.43	1204	1306	16.6
Selected model	tons>0 ~ fsyr + s(soi) + s(dchr)				

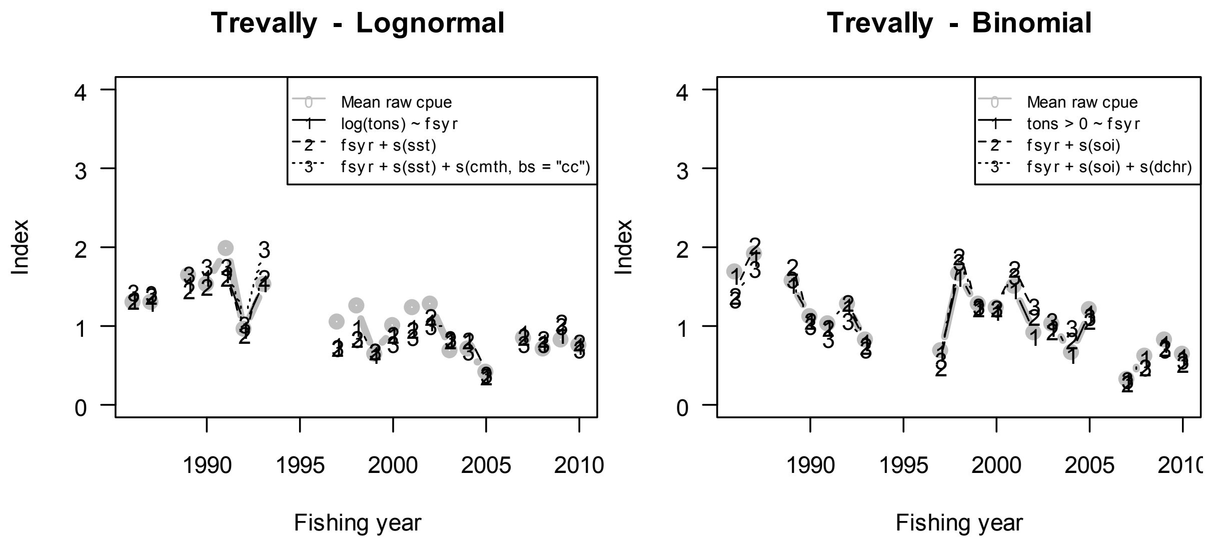

Figure 3: Stepwise standardised annual indices (SPUE) and mean raw (unstandardised) sightings from the trevally binomial and lognormal regressions; fishing year labels show first year of each couple e.g., 1998 is 1998–99.

A declining trend is evident in the time series. Target species is not significant in the model fit and the model accounts for less than 12% of variability. Diagnostic plots (observed proportion non-zero on expected proportion non-zero) (Figure C3) indicates no major deviations in the estimated values.

The combined trevally indices (Figure 4) show an overall decline throughout the period.

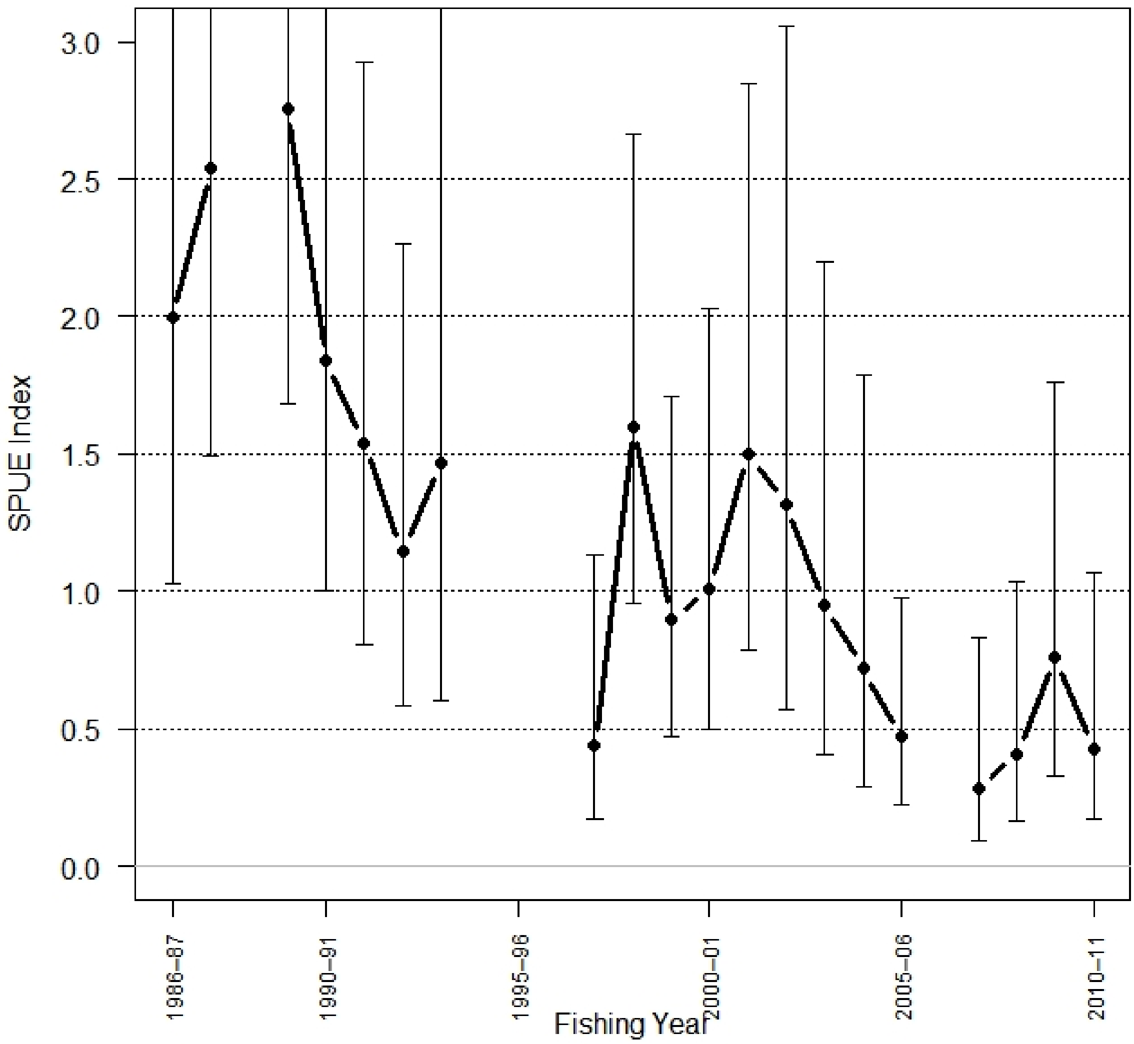

Figure 4: Final normalised combined indices of relative abundance (SPUE) for trevally generated as the combination of the binomial and lognormal regressions; vertical bars are the 95% confidence intervals.

3.3.2 Kahawai — model fits (binomial and lognormal) and indices

The results of the lognormal model-fitting are shown in Table 15 and Figure 5. The indices are bimodal with peaks at 1992–93 and 2007–08, although the second peak is considerably higher than the first. The final model is:

log(tons) ~ fsyr + s(cmth, bs = "cc") + s(targt)

Effort was not accepted into the model. Diagnostic plots (Figure C4) indicated no violation of model assumptions. Partial effects plots for the kahawai lognormal fit are shown in Figure C5.

The results of the binomial model-fitting are shown in Table 15 and Figure 5. The final model is:

log(tons) ~ fsyr + s(cmth, bs = “cc”) + s(targt).

The indices were flat throughout the time series. The model accounts for 17% of the variability. Effort is not accepted by the model. Diagnostic plots (observed proportion non-zero on expected proportion non-zero) (Figure C6) indicate no major deviations in estimated values. Plots of partial effects are included in this figure.

Table 15: Stepwise model fits (binomial and lognormal) for kahawai.

	Predictor added	Df	Deviance	AIC	R^2
Lognormal	fsyr	20.00	1186	2834	14.8
	s(cmth, bs = "cc")	24.96	1000	2691	28.1
	targt	**30.22**	**943**	**2649**	**32.3**
	s(sst)	35.76	912	2630	34.5
	s(dchr)	38.79	897	2621	35.6
	s(efft)	41.58	879	2609	36.9
	s(soi)	47.07	863	2603	38.0
	s(moon)	50.14	857	2603	38.4
Selected model	log(tons) ~ fsyr + s(cmth, bs = “cc”) + s(targt)				
Binomial	fsyr	20.00	1355	1395	4.6
	targt	**26.00**	**1168**	**1220**	**17.7**
	s(sst)	29.53	1127	1186	20.6
	s(efft)	34.94	1112	1182	21.6
	s(moon)	36.10	1108	1180	22.0
	s(dchr)	33.76	1107	1175	22.0
	s(soi)	39.63	1091	1170	23.2
Selected model	tons>0 ~ fsyr + s(targt)				

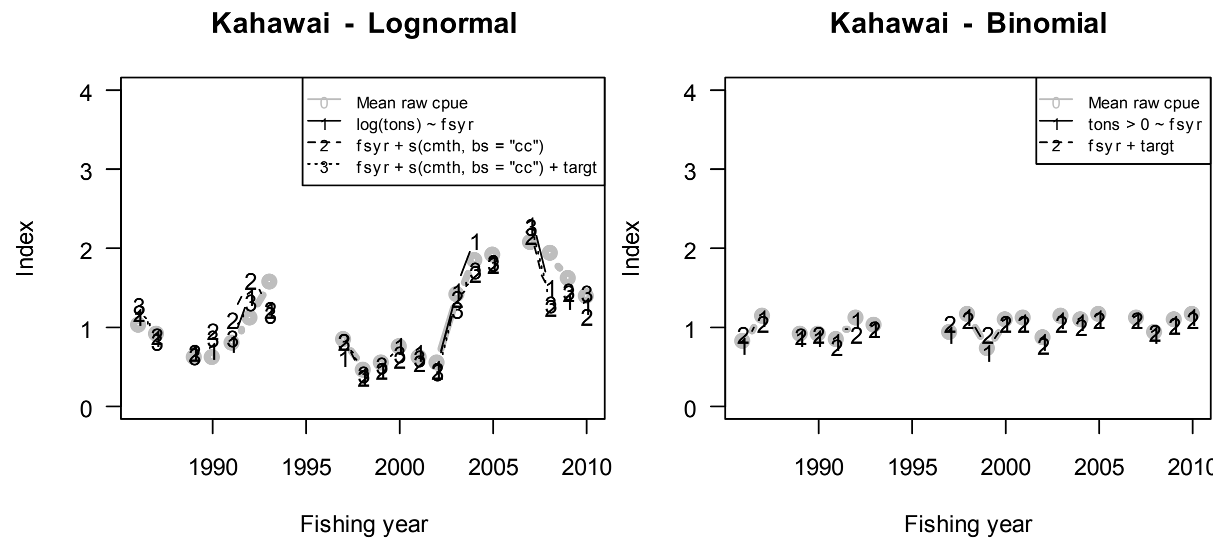

Figure 5: Stepwise standardised annual indices (SPUE) and mean raw (unstandardised) sightings from the kahawai binomial and lognormal regressions; fishing year labels show first year of each couple e.g., 1998 is 1998–99.

The combined indices are similar to those from the lognormal fits (Figure 6).

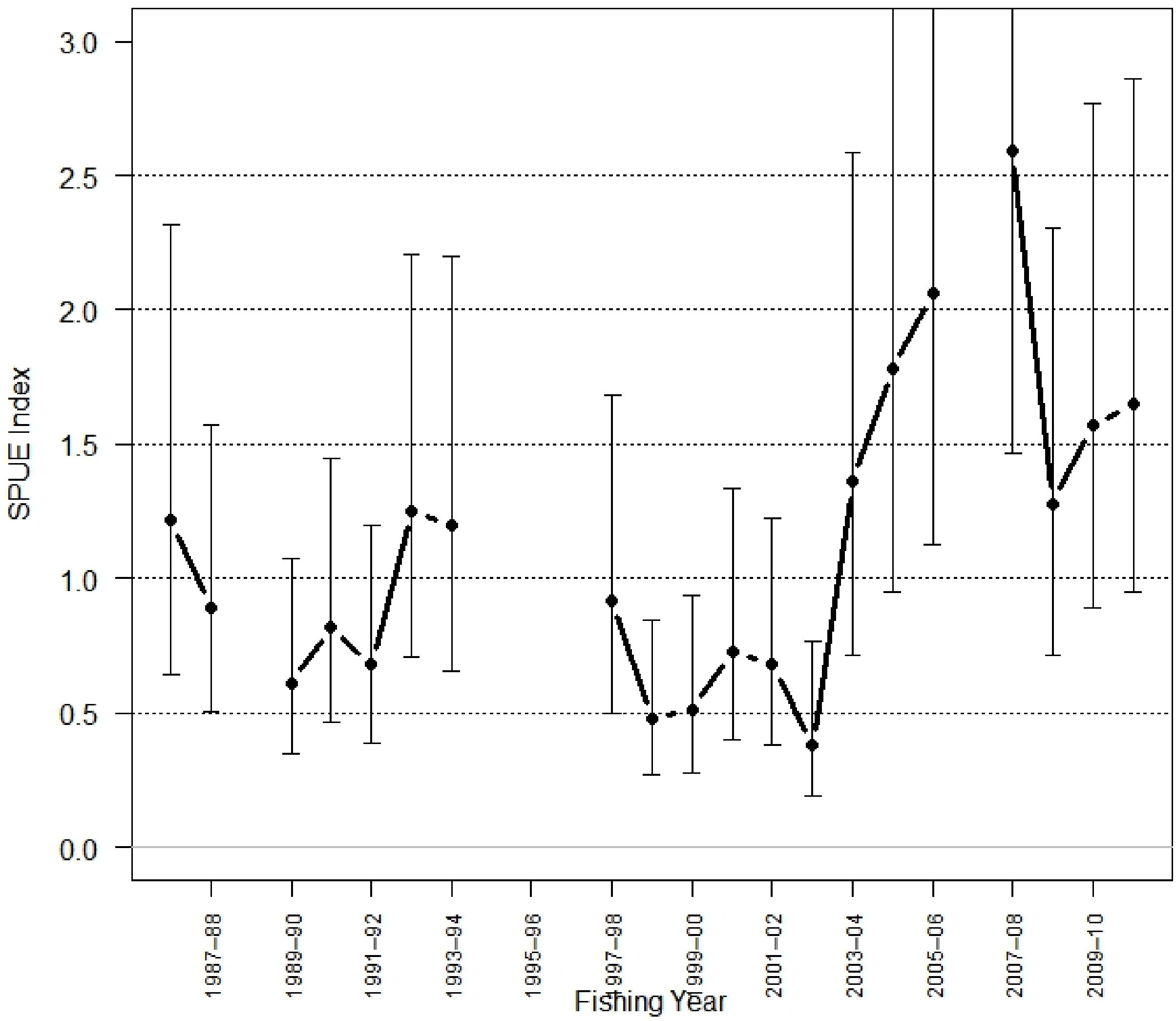

Figure 6: Final normalised combined indices of relative abundance (SPUE) for kahawai, generated as the combination of the binomial and lognormal regressions; vertical bars are the 95% confidence limits.

3.3.3 Final indices for kahawai and trevally in the BoP

The final annual relative abundance indices for kahawai and trevally in the BoP, along with their coefficients of variation (CVs) are summarised in Table 16.

Table 16: Final normalised combined abundance indices and CV for kahawai and trevally in the BoP, estimated from aerial sightings data.

Fishing year	Trevally Index	Trevally CV	Kahawai index	Kahawai CV
1986–87	2.00	0.34	1.22	0.33
1987–88	2.54	0.27	0.89	0.29
1988–89	No data			
1989–90	2.76	0.25	0.61	0.29
1990–91	1.84	0.31	0.82	0.29
1991–92	1.54	0.33	0.68	0.29
1992–93	1.15	0.35	1.25	0.29
1993–94	1.47	0.47	1.20	0.31
1994–95	No data			
1995–96	No data			
1996–97	No data			
1997–98	0.44	0.50	0.92	0.31
1998–99	1.60	0.26	0.48	0.29
1999–00	0.90	0.33	0.51	0.31
2000–01	1.01	0.36	0.73	0.31
2001–02	1.50	0.33	0.68	0.30
2002–03	1.32	0.44	0.38	0.36
2003–04	0.95	0.44	1.36	0.33
2004–05	0.72	0.48	1.78	0.32
2005–06	0.47	0.38	2.06	0.31
2006–07	No data			
2007–08	0.28	0.59	2.59	0.29
2008–09	0.41	0.49	1.28	0.30
2009–10	0.76	0.44	1.57	0.29
2010–11	0.43	0.48	1.65	0.28

3.4 Additional analyses

3.4.1. Investigating apparent inconsistencies in the kahawai binomial indices

Plots of the raw annual proportions of flights with non-zero kahawai sightings from the binomial model are shown in Figure 7. These plots suggest similar trends in the two curves.

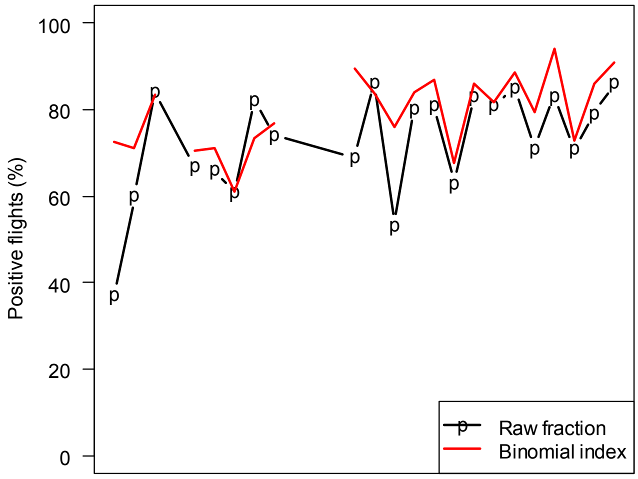

Figure 7: Annual proportion of non-zero flights, raw data (p) and that from the kahawai binomial model (re-normalised to the mean of the raw fractions over common years).

3.4.2. Evaluating the SPUE indices — Kahawai

The index of flights with zero flights plotted against the kahawai catch history and TACC is shown in Figure 8. The rationale here is that there should be fewer zero sightings because the catch is lower in later years.

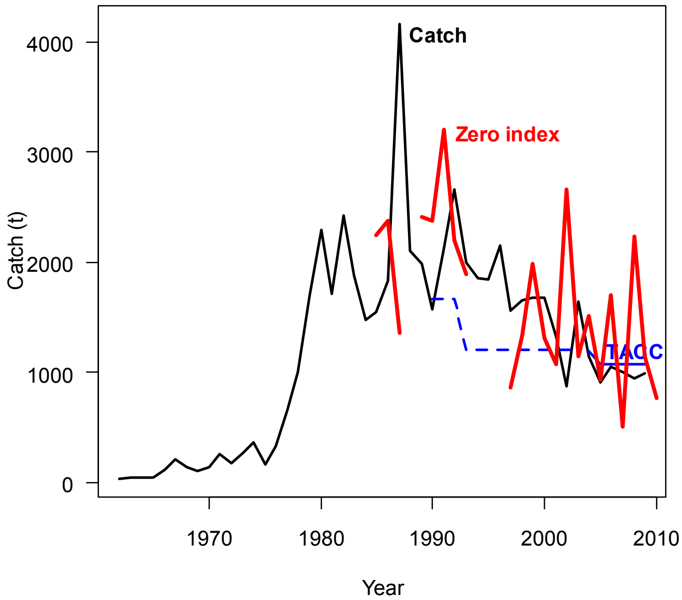

Figure 8: Kahawai catch history for fishing years 1985–86 to 2010–11, with TACC levels and index of flights with zero sightings of kahawai (re-normalised to the mean of the catch over common years).

The plot of kahawai SPUE against recreational CPUE from boat ramp surveys in the BoP is shown in Figure 9. Although not a perfect match, both indices show an increase in the later years.

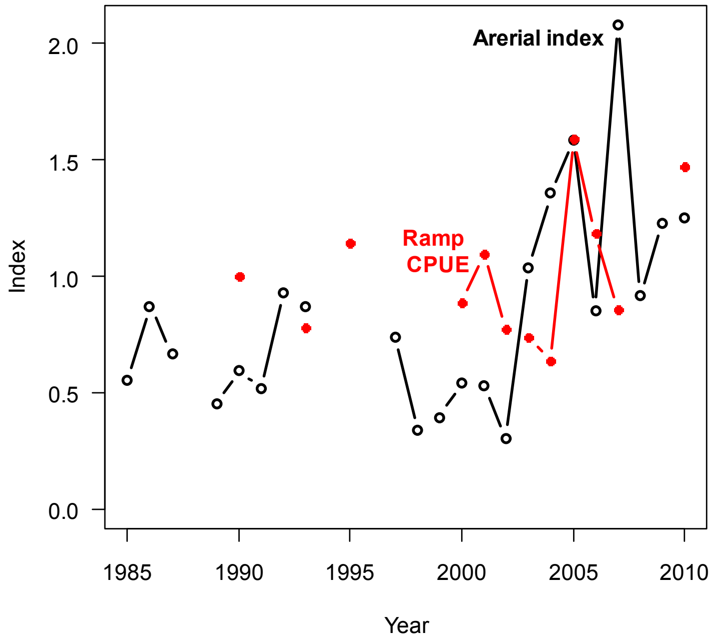

Figure 9: Kahawai SPUE (combined binomial-lognormal aerial index) and recreational CPUE from boat ramp surveys (Ramp CPUE) for the Bay of Plenty and fishing years 1985–86 to 2010–11. SPUE is re-normalised to the mean of the ramp CPUE over common years.

3.4.3. Evaluating the SPUE indices — Trevally

The trevally SPUE indices are shown in Figure 10 along with the trevally catch history constructed from table 1 in the trevally working group report (Ministry of Fisheries, 2011) and the trevally TACC. Here, the catch is not constrained by the TACC and it is decreasing with time. The SPUE also decreases showing that a lower population size could be the cause of the declining catches relative to the TACC.

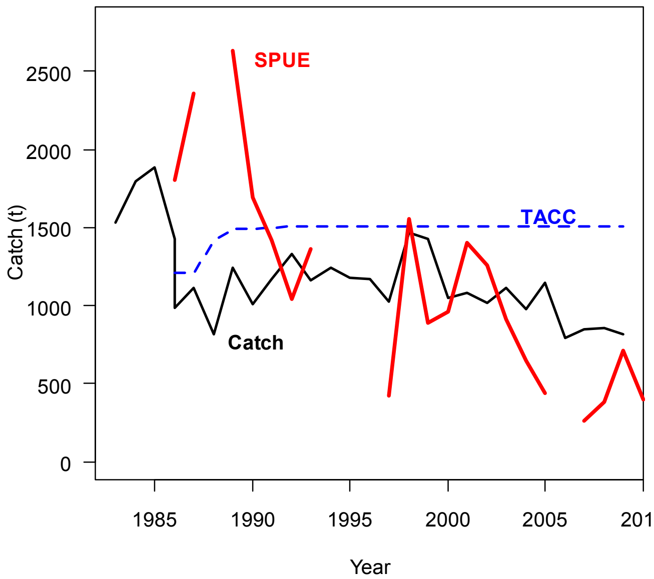

Figure 10: Trevally SPUE (aerial index) and recreational CPUE from boat ramp surveys (Ramp CPUE) for the fishing years 1985–86 to 2010–11. SPUE is re-normalised to the mean of the ramp CPUE over common years.

3.4.4 Comparing trends of the trevally lognormal and binomial indices

The plot of positive (lognormal) and non-zero (binomial) trends is shown in Figure 11. Here, both indices decline over time, which might be expected if the population was decreasing.

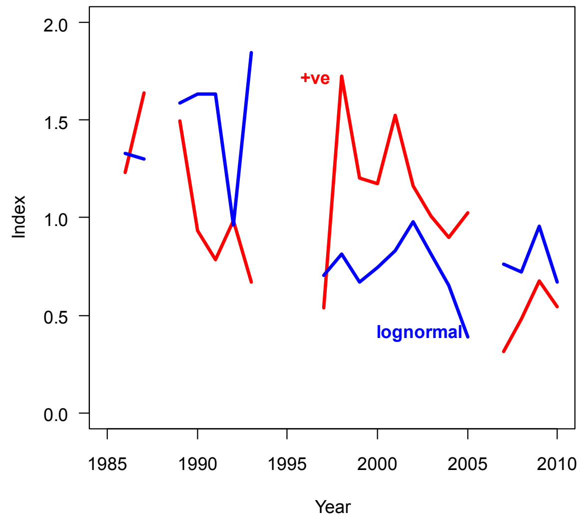

Figure 11: Standardised sightings for trevally for fishing years 1985–86 to 2010–11 from the binomial and lognormal fits. Binominal indexes non-zero sightings. Indices are re-normalised to the same mean over common years.

3.4.5 Investigating the low explanatory power of effort in the models

The adjusted effort data (flight duration for all flights in the kahawai analyses, adjusted to remove time associated with assisting fishing operations) is characterised by the following summary:

Minimum value	1st Quartile	Median	Mean	3rd Quartile	Maximum value
-1.4	0.70	1.50	1.7	2.50	7.3

The distribution of adjusted effort is shown in Figure 12. The majority of effort (87%) is represented by the four largest cells (from 0 to 4 hours).

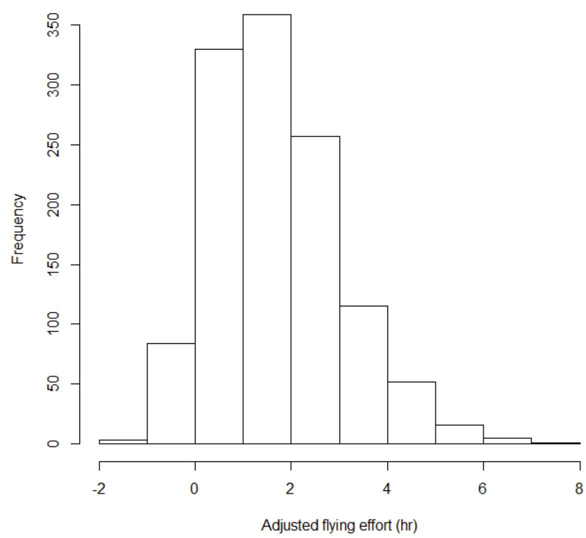

Figure 12: Distribution of flight duration for all flights in the kahawai analyses, adjusted to remove time associated with assisting fishing operations.

4. DISCUSSION

4.1 Factors limiting the time frame

The overall aim of the present study was to produce SPUE-based indices of abundance for KAH 1 and TRE 1 through investigation of the aerial sightings data. During discussions on preliminary work under SAP2006-10, the NINSWG had determined that development of the indices should proceed according to three stages. Under Stage 1, which investigated the utility of the aerial sightings data as a basis for producing relative abundance indices, a method was developed that was accepted by the NINSWG. The work documented here presents the results of Stages 2 and 3.

The aerial sightings dataset includes sightings of several species of inshore pelagic schooling finfish species from 1976 to September 2011. There have been two major revisions of the data collection form, the first in 1985 when there were improvements to the collection of flightpath data, and the second in 1998 when GPS positions were added as well as information on the fishing operation. These revisions and associated changes in the data have imposed a system of three periods on the dataset which resulted in certain limitations but also provided a means of breaking the data into workable pieces.

Thus, under Stage 1 development of an acceptable method was limited to data collected on the most recent form version. This portion of the dataset contained the greatest amount of information, with data on the fishing operation, and also the most accurate data on positions of the sightings. Although these positions were not used in the final method, their being available for the exploratory data analysis added to our understanding of the dataset as the method was developed.

Under Stages 2 and 3, the method and lessons determined during Stage 1 were applied to the earlier portions of the dataset. The method of adjusting flying effort for non-search time required operational data as well as estimates of the total number of sightings of the major species examined and recorded during each flight. Operational data were available on the data-collection forms only from 1989, but Pilot #2 had collected similar data since 1983 and made his annotated forms available. This allowed adjustment of flying effort to be extended back to 1983.

The revision to the data-collection form in 1985 had improved the flightpath data. Previously the flightpath had consisted of a list of landmarks visited or flown over during the day's flying. These included islands, headlands, shoals, and reefs. The improvements comprised adding the map of Panel 5a (see Figure 1) and recording the 10–15 min periods as strokes or ticks in the grid squares. This system allowed selection of flights exclusive to a particular area (e.g., the BoP) and this strategy became a key to the method developed under Stage 1. The disadvantage of this was that data collected before this revision was made could not be identified as exclusive to any particular area with any certainty. This major feature limited the analysis completed in the present study to the period beginning January 1986.

4.2 The estimated indices of relative abundance

Relative abundance indices have been produced here for kahawai and trevally in the BoP. In each case the series cover most years from the 1986–87 fishing year to 2010–11, with a gap for the three years 1994–95 to 1996–97. The kahawai series has an additional gap for 2006–07, when it was deemed by the NINSWG that the low number of records had produced an outlying estimate in the original series that should be removed. Similarly for both series (kahawai and trevally), the WG had requested removal of the first year in the original series because the estimate appeared to be erroneous as a result of low numbers of records.

The final time series of annual relative abundance indices are a combination of the indices from the binomial and lognormal fits. The same flights are used for both the kahawai and trevally analyses, but the ratio of positive to zero sightings is markedly different between the two species. The lognormal model is based purely on tonnage, while the binomial model is a two-score dataset — there is either a sighting or there is not — so all the data points contribute to the binomial model.

Diagnostic plots for the lognormal fits indicate that, generally the assumptions of the GAM fitting methodology are met. For the kahawai lognormal fit, the normal Q-Q plot is close to a straight line, suggesting that the distributional assumption is reasonable. The residual versus linear predictor plot indicates that the assumption of constant variance is not violated, and the histogram of residuals appears approximately consistent with normality, although it is skewed a little to the left. The response versus fitted value plot shows a reasonable degree of scatter and a positive linear relationship with little indication of the assumption of constant variance being violated.

For trevally there are fewer data points. There is a hint in the residuals versus linear predictor and response versus fitted values plots of variance increasing with fitted values, but this could be the result of a small dataset. The histogram of residuals appears normally distributed, particularly given the size of the dataset.

In most cases the plots of observed proportion non-zero versus expected proportion non-zero for the binomial fits for both trevally and kahawai cluster closely around the 1:1 line, indicating no major error in estimations from the fitting process.

Thus, results of the standardisations showed reasonable fits with no clear violations of model assumptions. Levels of variability explained for the selected models were 25.6% and 11.9% for the trevally lognormal and binomial fits respectively, and 32.3% and 17.7% for the kahawai lognormal and binomial fits respectively.

4.3 Conclusions of the NINSWG

Most of the conclusions documented here were reached by the NINSWG at the meeting on 16 April 2012, after the additional plots shown above in Section 3.4 had been presented. The conclusion regarding the east Northland dataset was made by the NINSWG at the meeting on 29 March 2012.

4.3.1. East Northland

- Given the low numbers of flights in East Northland, the WG did not recommend attempting to generate SPUE indices of abundance for this region.

4.3.2. Kahawai, BoP

- Binomial models of the annual proportion of flights with zero sightings, and models of the tonnage sighted, both suggest that kahawai abundance (in terms of both number and size of schools) had increased following the reduction of commercial catch in the early 2000s. This observation is consistent with anecdotal reports from commercial and recreational fishers, and spotter pilots.

- Some NINSWG members felt that it was unlikely that recent abundance should be greater than it was in the mid 1980s. The WG did, however, agree that this result was plausible if recruitment had been high in recent years and given that increased commercial catches in the period 1980–1985 may have reduced abundance prior to the index beginning in1986.

4.3.3 Trevally, BoP

- Trevally indices of abundance have declined since 1986, which is consistent with anecdotal reports from fishers and spotter pilots, as well as a contracting age distribution in the catches made by the bottom trawl fishery.

4.3.4. Overall conclusions

Models of SPUE for kahawai and trevally probably do reflect, to some degree, the abundance of these two species in the Bay of Plenty. The SPUE indices should be used for stock assessment, with stock assessment model diagnostics employed to gauge the quality (and appropriate weight) of the abundance indices.

4.4. Implications for stock assessments

The primary aim of this work is to produce annual indices of relative abundance for inshore schooling pelagic finfish. Because jack mackerel cannot be separated into their three component species, and because high levels of variation were evident in preliminary estimates for blue mackerel under SAP2006-10, these data cannot provide satisfactory indices for those species. The results from investigative work under SAP2006-10 (Taylor 2014) showed that reasonable indices could be expected for trevally and kahawai and, based on those results, annual indices of relative abundance have been produced in the present study for these two species in the BoP. The aim now is to use these series as stock indices in stock assessment models for trevally and kahawai in QMA 1. Initially, stock assessment diagnostics should be employed to determine the quality and appropriate weight of these indices for use in subsequent stock assessment modelling.

5. ACKNOWLEDGMENTS

- Many thanks to the fish-spotter pilots for continuing to supply data over the years, particularly Red Barker for his ongoing discussion on many issues related to the data collection and behaviour of the fish, and to John Reid and Brian Decke for other useful dialogue.
- This work was funded under MFish project SEA2010-17.
- We also thank Andy McKenzie and Marc Griffiths for independently reviewing this report, and Marianne Vignaux for comments on an earlier version of the manuscript.

6. REFERENCES

Fisher, D.O.; Sanders, B.M. (2011). Database documentation for the Ministry of Fisheries, Fisheries Statistics Unit database new_fsu. NIWA Fisheries Data Management Database Documentation Series.

Hastie, T.J.; Tibshirani, R.J. (1990). Generalized additive models. Chapman & Hall.

Middleton, D.A.J.; Starr, P.J.; Griffiths, M. (2010). Assigning target species to purse seine spotter flights. Report to the Northern Inshore Stock Assessment Working Group, NINSWG-2010-25.

Ministry of Fisheries (2011). Report for the Fisheries Assessment Plenary, May 2011: stock assessments and yield estimates. Ministry of Fisheries. 1178 p.

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Taylor, P.R. (2014). Developing indices of relative abundance from observational aerial sightings of inshore pelagic finfish; Part 1, exploring the data. *New Zealand Fisheries Assessment Report 2014/34*. 66 p.

Wood, S.N. (2006). Generalized additive models: An introduction with R. Chapman & Hall/CRC.

7. APPENDICES

Appendix A: An explanation of the steps required to extract and process the data used in the analysis

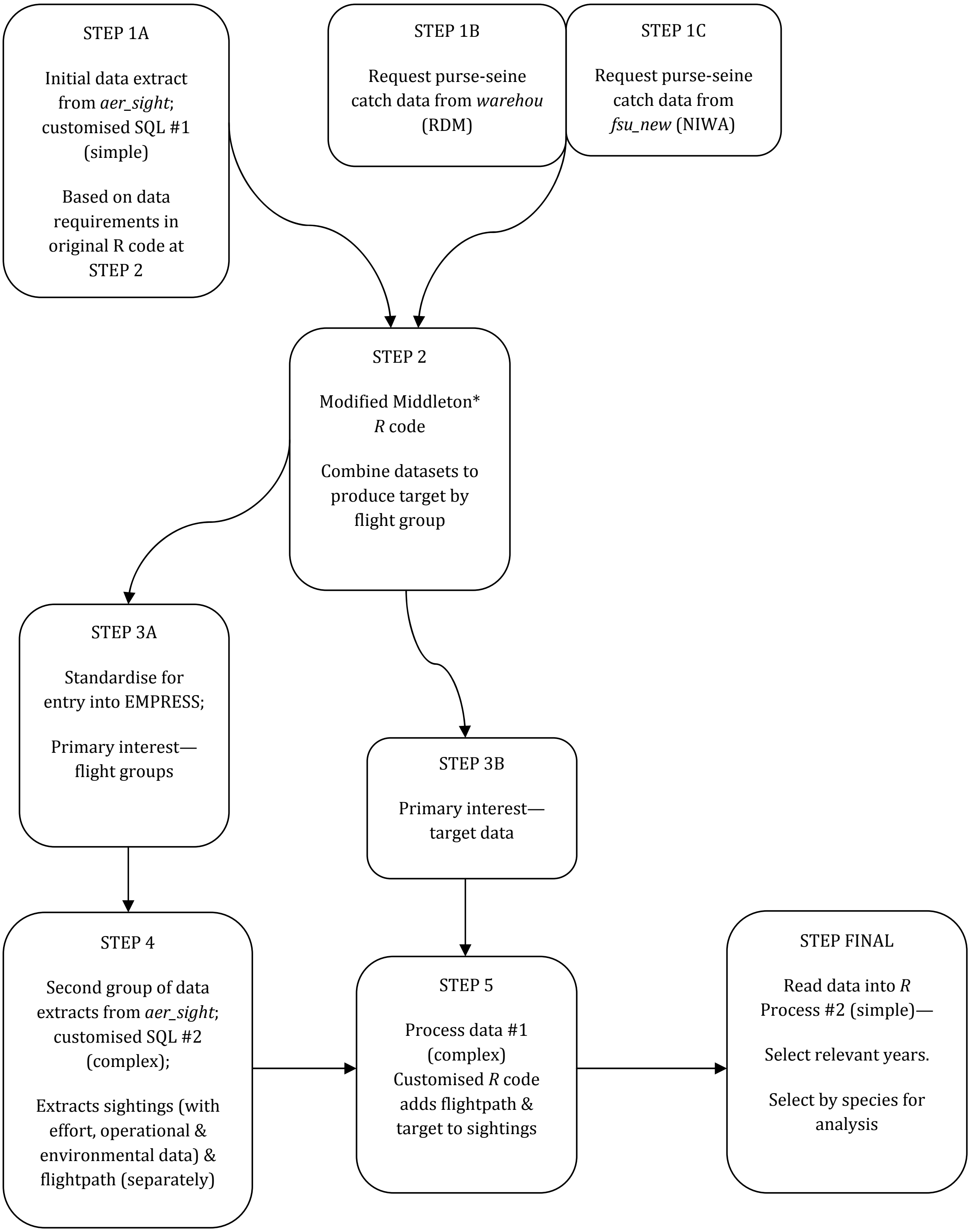

Figure A1: Steps in extracting and processing the data used in the analysis. See text for details. *Middleton et al. (2010).

A1. Extract steps 1 and 2

Once it was established that the SAP2006-10 data processing had been duplicated, the targetInv code was used to provide a list of *flt_grp* identifiers for those flights in the SEA2010-17 dataset on days when there were target data available. To achieve this, data for the SEA2010-17 analysis were first extracted from *aer_sight*, read into *R*, and fed into the targetInv code. Instead of covering the period June 1998 to July 2009 as in the SAP2006-10 extract, the SEA2010-17 dataset and modification to the *R* code expanded the period to cover from 1 January 1986 to the end of the 2010–11 fishing year (30 September 2011).

The second methodological change made at this point was inclusion of summarised catch data as a surrogate for target species instead of the modal target data that had been used in the SAP2006-10 analysis. The catch data were a combination of extracts requested of RDM (from *warehou*) and of the NIWA Fisheries Data Manager (from *fsu_new*).

The tasks described above are represented by Steps 1 and 2 in Figure A1: Step 1A is the initial data extract from *aer_sight*; Step 1B is represented by the data requested from RDM; and Step 1C is the data requested from the NIWA Data Manager. Data from Steps 1B and 1C were combined to produce a catch dataset for the entire period of interest and then fed into the targetInv code at Step 2, along with the aerial sightings data from Step 1A. The extract from *aer_sight* at Step 1A is referred to as a *simple* extract because it simply extracted required data from the *t*_flight, *t_flightpath*, and *t_set* tables and stored them in an ascii file which was read into *R*.

Two extracts, one each of catch and effort data from *Warehou* were requested from RDM (replog 8379) for Step 1B. The SQLs for these extracts were based on those defined by Middleton et al. (2010) (replog 7736) with modifications to extend the time frame as required for the current work. For Step 1C, extracts of purse-seine catch and effort data were requested from the NIWA Fisheries Data Manager based on the SQL used for Step 1B. An *R*-code script was written to combine the two catch and the two effort datasets into a single catch and a single effort dataset for the period and area of interest. The combined datasets were fed into the targetInv *R*-code at Step 2 along with the aerial sightings data extracted at Step 1A.

The aim at Step 2 was to produce a list of flight-group codes for which there were associated target species data. Consistent daily modal target species from the purse-seine fleet were not available before 1998, so daily purse-seine catch by species in the area of interest was being used here as surrogate for target for the entire period. Target species was selected as the species with highest catch, but if there was no clear “winner”, with more than one species contributing similar tonnages to the day’s total catch, then the category “MIX” was assigned for that day.

A2. Extract step 3A

A second data extract was taken from *aer_sight* and the standardising at Step 3A in Figure A1 was made in preparation for carrying it out. So, at Step 3A, the list of *flt_grp* identifiers produced by targetInv was standardised (i.e., EMPRESS delimiters were added) for entry into a temporary database table in *aer_sight*.

A3. Extract step 3B

Step 3B also used output from targetInv, but in this case the relevant target species for each day was listed along with the *flt_grp* identifiers. Hence the primary interest flag in Figure A1 of “target data” in comparison with the primary interest flag for Step 3A of “flight groups”.

A4. Extract step 4

A second group of extracts from *aer_sight* was carried out at Step 4. Generally, these extracts are referred to as *complex* (see Figure A1) because extensive manipulations were performed on the data as they were passed between several temporary files before finally having environmental variables added and the data extracted from *aer_sight* and stored in an ascii file. This group of extracts can be further described in terms of two broad categories: flightpath and sightings. For each area (BoP, east Northland), there was a single flightpath extract and a series of sightings extracts. Two sightings extracts were produced for each species (kahawai, trevally, jack mackerel, blue mackerel, skipjack tuna): one each for the lognormal and binomial models. Thus, for each area there were 11 separate extracts. In addition to providing sightings and associated covariate data for producing SPUE indices of the two central species, kahawai and trevally, the individual species extracts provided information on the total number of sightings which was used in the adjustment to effort (see Section 2.3).

A5. Extract step 5 and Step Final

Outputs from Steps 4 and 3A (Figure A1) were fed into customised *R* code that carried out extensive processing, adding flightpath and target species to the sightings data as well as calculating and adding the daily total number of sightings of all species. A final processing step followed where data for incomplete fishing years were removed from the dataset. This was included as a preliminary function call in the *R* script used for fitting the lognormal and binomial models. This completed preparation of the data for analysis.

Appendix B: An explanation of components of the aerial sightings data collection form

The aerial sightings data reside in an EMPRESS relational database that comprises five main relational tables and several ancillary tables. The latter contain environmental data, definitions for codes used in the main tables, and other information to facilitate grouping during data extracts (e.g., temporal periods — calendar year and month, fishing year and month). The main tables reflect the five main panels on the data-collection form (see Figure 1). The following is a brief description of the information recorded on each panel, including the database table in which each group of data are stored.

Panel 1
Description: meta-data for a group of flights.
Specific data: date, pilot, customer, aircraft call-sign.
Database table: *t_flight_group*.

Panel 2
Description: takeoff and landing data.
Specific data: takeoff airfield, takeoff time, landing airfield, landing time.
Database table: *t_flight*.

Panel 3
Description: various data on the sightings made while observing the group of schools comprising a sighting.
Specific data: time of the sighting (Time 1), species (or species mix) in schools comprising the sighting, number of schools in the sighting, the size of the smallest school in the sighting (ton_min), the size of the largest school in the sighting (ton_max), the pilots estimate of the total tonnage (Est. total), sea condition at the time the sighting was made, latitude and longitude (from GPS).
Database table: *t_school_sight*.

Panel 4
Description: operational data.
Specific data: original time of the sighting (Time 1; Note that this is the same time as in Panel 3 and allows position of the school and other information to be accessed), time that fishing on the school began (Time 2), the vessel name, the tonnage and species composition estimated by the pilot (Ton Sp Set), the tonnage and species composition determined by crew on the vessel after the school has been landed to the hold (Ton Sp Land), result of the fishing (Rst) — options are caught, saved, skunked, unknown, caught unknown amount (unavailable from the vessel), let go, burst net.
Database table: *t_set*.

Panel 5
Description: effort data — strokes recorded by pilots into the squares on panel 5a represent 10–15 min periods spent in particular grid squares, which are summed and recorded on panel 5 at the time of form processing.
Specific data: number of ticks (first two spaces), grid square code (spaces 3–5).
Database table: *t_flightpath*.

Appendix C: Diagnostic plots for the standardised indices, BoP

Trevally — lognormal fits

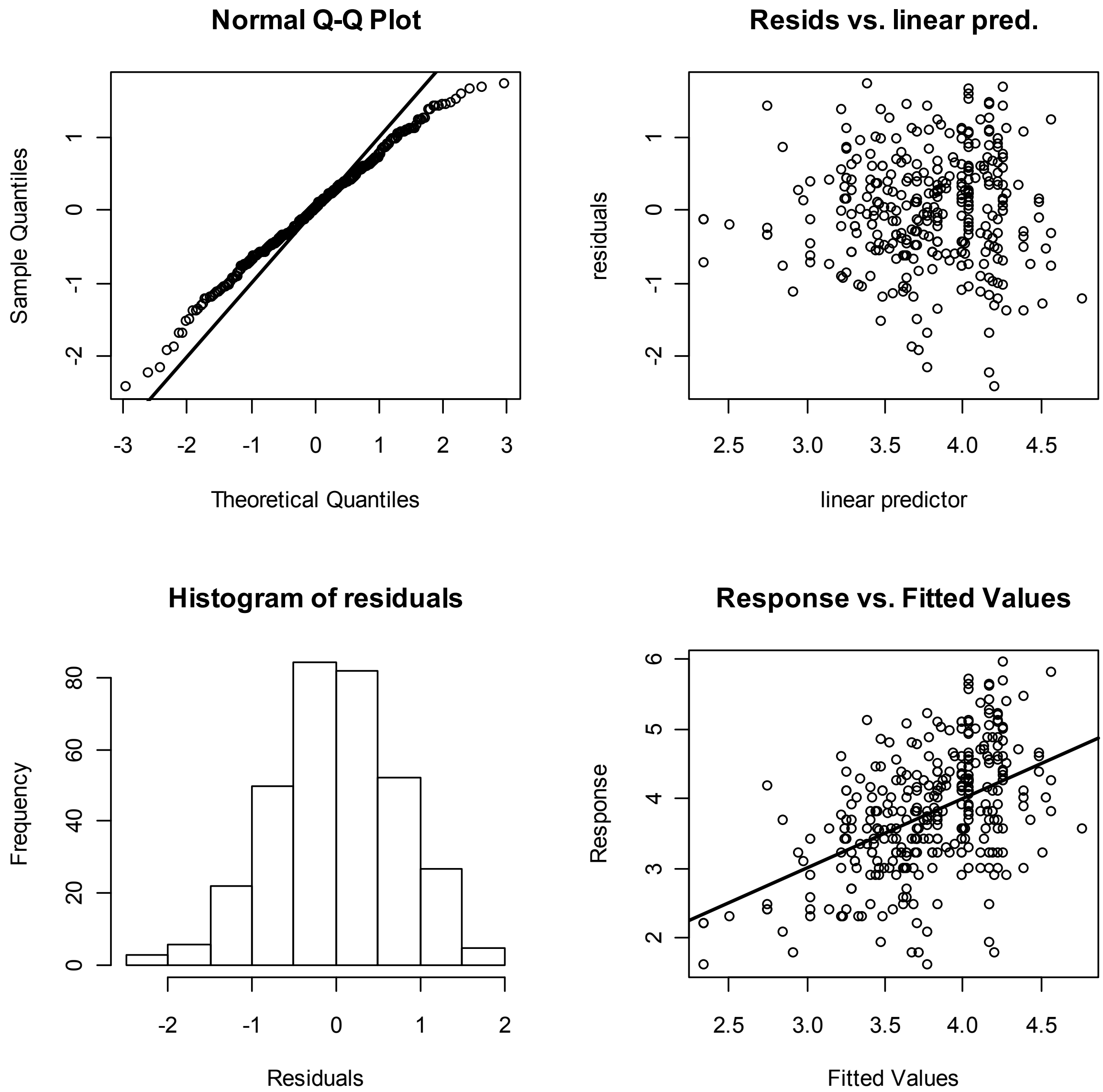

Figure C1: Diagnostic plots from the trevally lognormal fit.

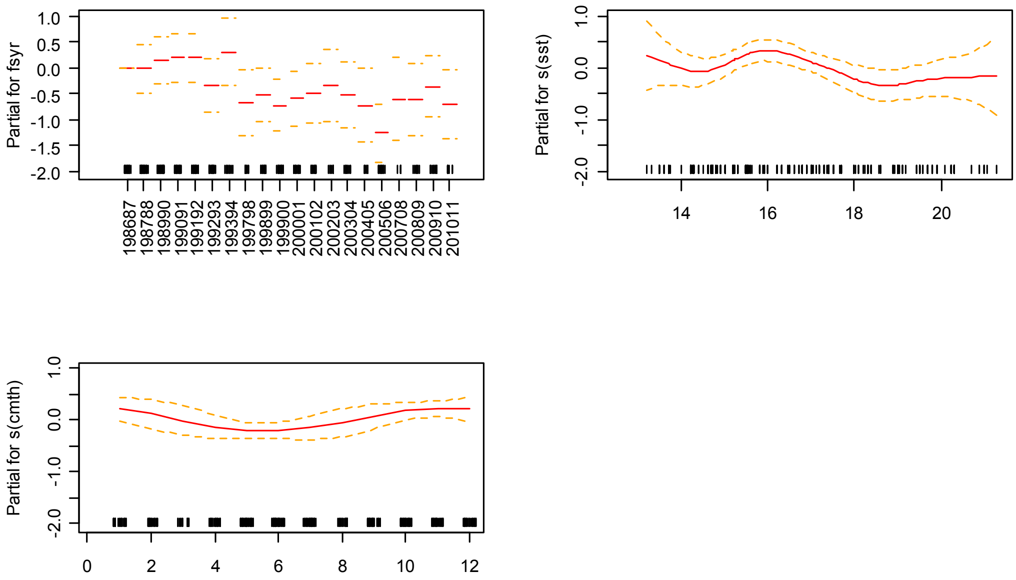

Figure C2: Partial effect plots for the trevally lognormal fit.

Trevally — binomial fits

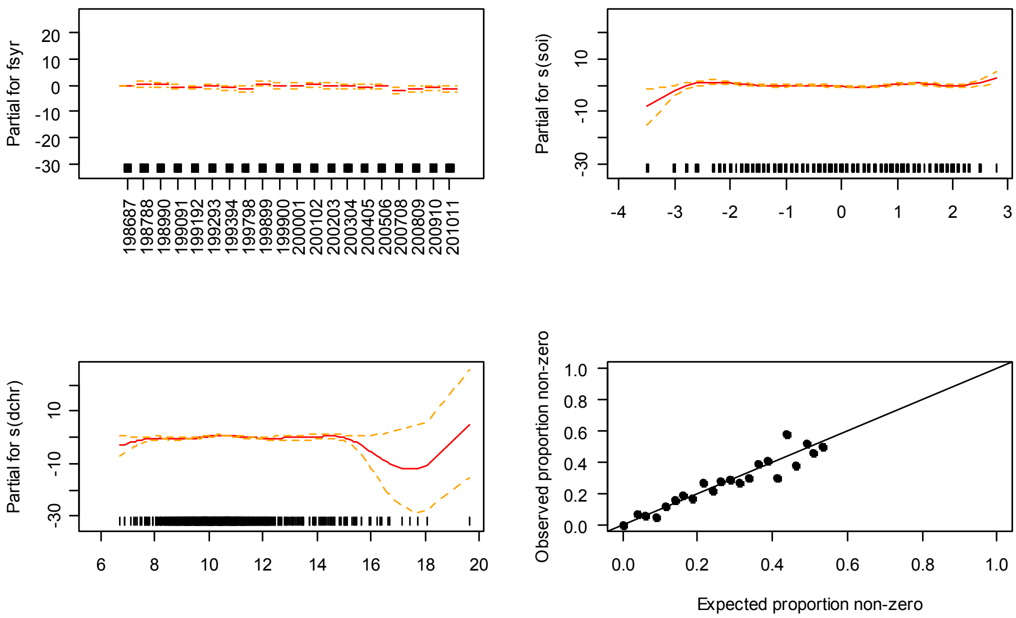

Figure C3: Partial effect plots and observed proportion non-zero on expected proportion non-zero, for the trevally binomial fit.

Kahawai — lognormal fits

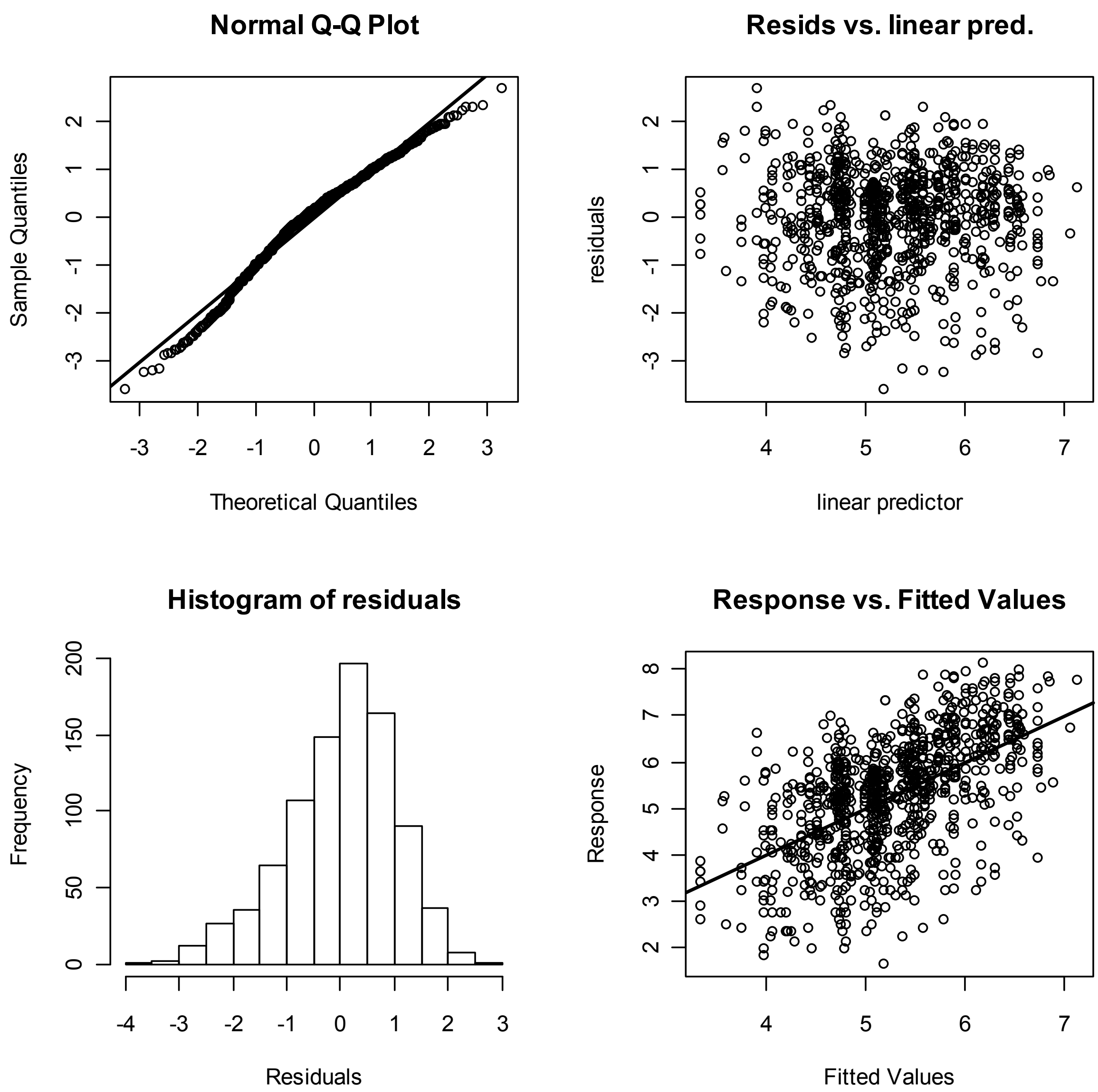

Figure C4: Diagnostic plots for the kahawai lognormal fit with unadjusted effort forced.

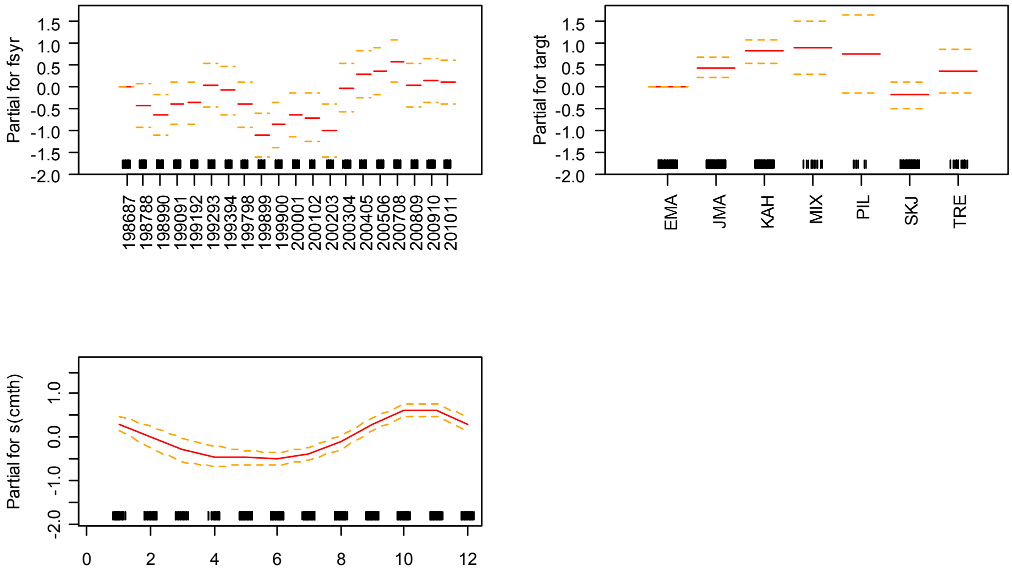

Figure C5: Partial effect plots for the kahawai lognormal fit with adjusted effort offered.

Kahawai — binomial fits

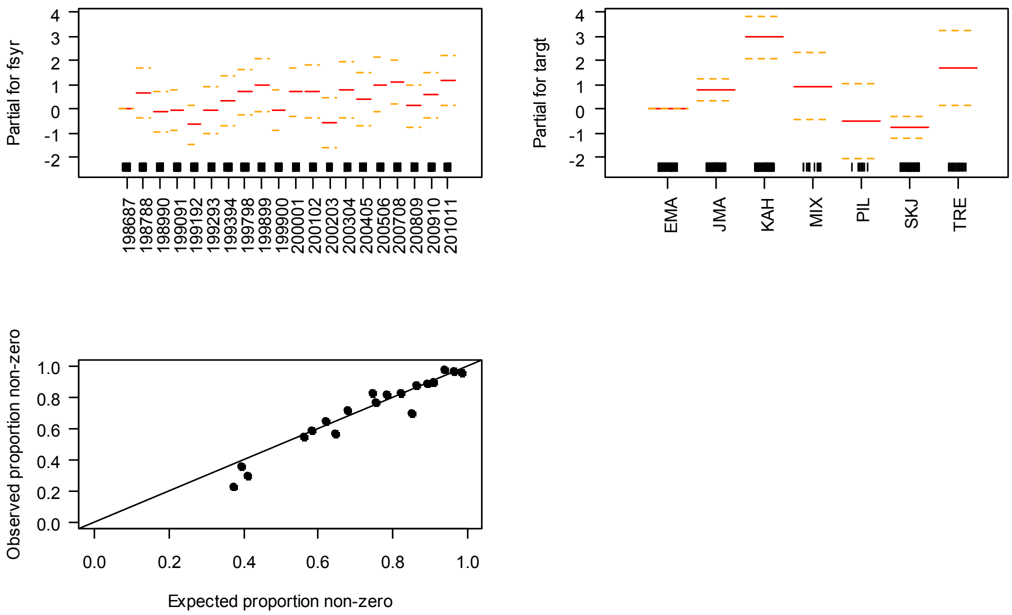

Figure C6: Partial effect plots and observed proportion non-zero on expected proportion non-zero, for the kahawai binomial fit with adjusted effort offered.