Fish shoal dynamics in north-eastern North Island
POP2019-02

Objective 2: Using distributions of inshore pelagic schooling finfish species from the aerial sightings database *aer_sight*

Paul Taylor
Statfishtics

Objective

- Analyse fish shoal data from purse seine fishery database (*aer_sight*) and develop a model of East Northland, Hauraki Gulf and Bay of Plenty oceanography to explore relationships between fish shoal abundance and the physical and biological aspects of the marine environment to better understand fisheries pressures on seabird population trends.

Overview

- This component of POP2019-02 will explore changes in aggregations of pelagic schooling finfish species over time by investigating relationships between environmental features and distributions of sightings of these species from the *aer_sight* database, along with other characteristics (e.g., school size, school number) of the sightings recorded in the database, and determine whether these changes reflect known changes in seabird abundance.
- Success with this work could provide the first step to later investigating whether a relationship can be demonstrated between fishing activity using commercial and recreational catch data, and trends in seabird population size.

Scope of the work

1. **Finalise methodology including the hypotheses to be tested within the limitations of the data available from *aer_sight* data.**
2. Determine relevant *aer_sight* data; request extract from MPI; organise into appropriate data structure.
3. Characterise changes in schooling aggregations over time i.e., size of schools, tonnage of sightings, number of schools.
4. Identify relevant bathymetric, oceanographic and environmental factors; gain access to the data and organise into appropriate data structure. This is the first step towards developing a model of the East Northland, Bay of Plenty and outer Hauraki Gulf bathymetry (reefs, channels, shelf edges), topographical features (islands, island groups and headlands), temporal changes (annual, seasonal), and SST and Temperature Fronts...
5. **... plus a series of predictor variables developed elsewhere (Stephenson et al, 2018).**

Predictor variables

- The following 18 high-resolution gridded environmental variables, described by Stephenson et al (2018) may be available for use as predictors.

TABLE 1 Environmental variables used as predictors in Gradient Forest analyses

Abbreviation	Full name	Description	Original resolution	Units	Source
Bathy	Bathymetry	Depth at the seafloor was interpolated from contours generated from various sources, including multibeam and single-beam echo sounders, satellite gravimetric inversion and others (Mitchell et al., 2012)	250 m	m	CANZ (2008)
Beddist	Benthic sediment disturbance	Combination of seabed orbital velocities (estimates the average mixing at the seafloor as a consequence of orbital wave action, calculated from a wave climatology-derived hindcast (1979–1998) of swell wave conditions in the New Zealand (NZ) region; Gorman, Bryan, & Laing, 2003) and friction velocity for seabed types (based on grain size). Benthic sediment disturbance from wave action was assumed to be zero where depth ≥200 m	1 km	Unitless	NIWA, unpublished
BotTemp	Temperature at depth	Annual average water temperature at the seafloor (using NZ bathymetry layer) based on methods from Ridgway, Dunn, and Wilkin (2002). The oceanographic data used to generate these climatological maps were computed by objective analysis of all scientifically quality-controlled historical data from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Atlas of Regional Seas database (CARS2009)	250 m	°C/km	NIWA, unpublished
BotNi	Bottom nitrate	Annual average water nitrate concentration at the seafloor (using NZ bathymetry layer) based on methods from Ridgway et al. (2002). Oceanographic data from CARS2009 (2009)	250 m	µmol/L	NIWA, unpublished
BotOxy	Dissolved oxygen at depth	Annual average water dissolved oxygen concentration at the seafloor (using NZ bathymetry layer) based on methods from Ridgway et al. (2002). Oceanographic data from CARS2009 (2009)	250 m	ml/L	NIWA, unpublished
BotSal	Salinity at depth	Annual average water salinity concentration at the seafloor (using NZ bathymetry layer) based on methods from Ridgway et al. (2002). Oceanographic data from CARS2009 (2009)	250 m	Psu	NIWA, unpublished
BotSil	Bottom silicate	Annual average water silicate concentration at the seafloor (using NZ bathymetry layer) based on methods from Ridgway et al. (2002). Oceanographic data from CARS2009 (2009)	250 m	µmol/L	NIWA, unpublished
Disorgm	Coloured dissolved organic matter (CDOM)	Indicative of coloured dissolved organic matter (CDOM) absorption at 440 nm. Based on SeaWiFS ocean colour remote sensing data; modified Case 2 atmospheric correction; modified Case 2 inherent optical property algorithm (Pinkerton et al., 2005)	4 km	Indicative of CDOM absorption at 440 nm $a_g(440)$ (m^{-1})	Pinkerton (2016)
Roughness	Roughness	Roughness of the seafloor calculated as the standard deviation of depths in a surrounding 3 × 3 km neighbourhood (Leathwick et al., 2012)	250 m	Unitless	Leathwick et al. (2012)NIWA, unpublished data
SeasTDiff	Annual amplitude of seafloor temperature	Smoothed difference in seafloor temperature between the three warmest and coldest months. Providing a measure of temperature amplitude through the year	250 m	°C/km	NIWA, unpublished data

(Continues)

TABLE 1 (Continued)

Abbreviation	Full name	Description	Original resolution	Units	Source
Sed	Sediment type	Seabed sediment and rock data which were obtained from research surveys around the NZ region (≈6,000 data points) were combined with sediment type data from a global online database (Jenkins, 2010), providing a total of about 30,000 data points for the study area. Because of the uneven distribution of these data, interpolation was required to provide a value for each cell. Interpolation was carried out using a kriging process in ArcGIS 10.3.1 (ESRI, 2015) providing coarse categorical classification of sediment type around NZ	polygon data—compiled from maps at various scales	(1) calc-gravel, (2) calc-mud, (3) calc-sand, (4) Clay, (5) deep ocean clays, (6) gravel, (7) mud, (8) sand, (9) siliceous ooze and (10) volcanic	Anderson et al., 2016; NIWA, unpublished data
SstGrad	Sea surface temperature gradient	Smoothed magnitude of the spatial gradient of annual mean SST. This indicates locations in which frontal mixing of different water bodies is occurring (Leathwick et al., 2006). Derived from Sea-Viewing Wide Field-of-view Sensor (SeaWiFS) satellite imagery (Pinkerton et al., 2005)	1 km	°C/km	
SuspPM	Suspended particulate matter	Indicative of total suspended particulate matter concentration. Based on SeaWiFS ocean colour remote sensing data (Pinkerton & Richardson 2005); modified Case 2 atmospheric correction; modified Case 2 inherent optical property algorithm (Pinkerton et al. 2006)	4 km	Indicative of total suspended particulate matter concentration (g/m^3)	Pinkerton (2016)
TidalCurr	Tidal current speed	Maximum depth-averaged (NZ bathymetry) flows from tidal currents calculated from a tidal model for New Zealand waters (Walters, Goring, & Bell, 2001)	250 m	m/s	NIWA, unpublished data
VGPM	Productivity Model	Provides estimates of surface water primary productivity based on the vertically generalized productivity model of Behrenfeld and Falkowski (1997). Net primary productivity by phytoplankton (mean daily rate of water column carbon fixation) is estimated as a function of remotely sensed chlorophyll concentration, irradiance and photosynthetic efficiency estimated from remotely sensed Sea-Viewing Wide Field-of-view Sensor (SeaWiFS) satellite imagery (M. Pinkerton, NIWA, pers. Comm.)	9 km	$mgC\ m^{-2}\ day^{-1}$	NIWA, unpublished
*DynOc**	Dynamic oceanography	Mean of the 1993–1999 period sea surface above geoid	250 m	m	NIWA, unpublished
*BotOxySat**	Oxygen saturation at depth	Annual average oxygen saturation at the depths	250 m	μmol/L	NIWA, unpublished
*OxyUt**	Apparent oxygen utilization	The difference between the measured dissolved oxygen concentration and its equilibrium saturation concentration in water with the same physical and chemical properties	250 m	μmol/L	NIWA, unpublished

Note. Variables not used in the final analyses because of their high correlation with other variables are identified with an asterisk.

Function of the *aer_sight* dataset

Provides 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)*
- 3 species of jack mackerel *(Trachurus declivis, T. murphyi, and T. novaezelandiae)*
- kahawai (*Arripis trutta)*

and for the oceanic migratory species skipjack tuna *(Katsuwonas pelamis)* on which the domestic purse-seine fishery was based.

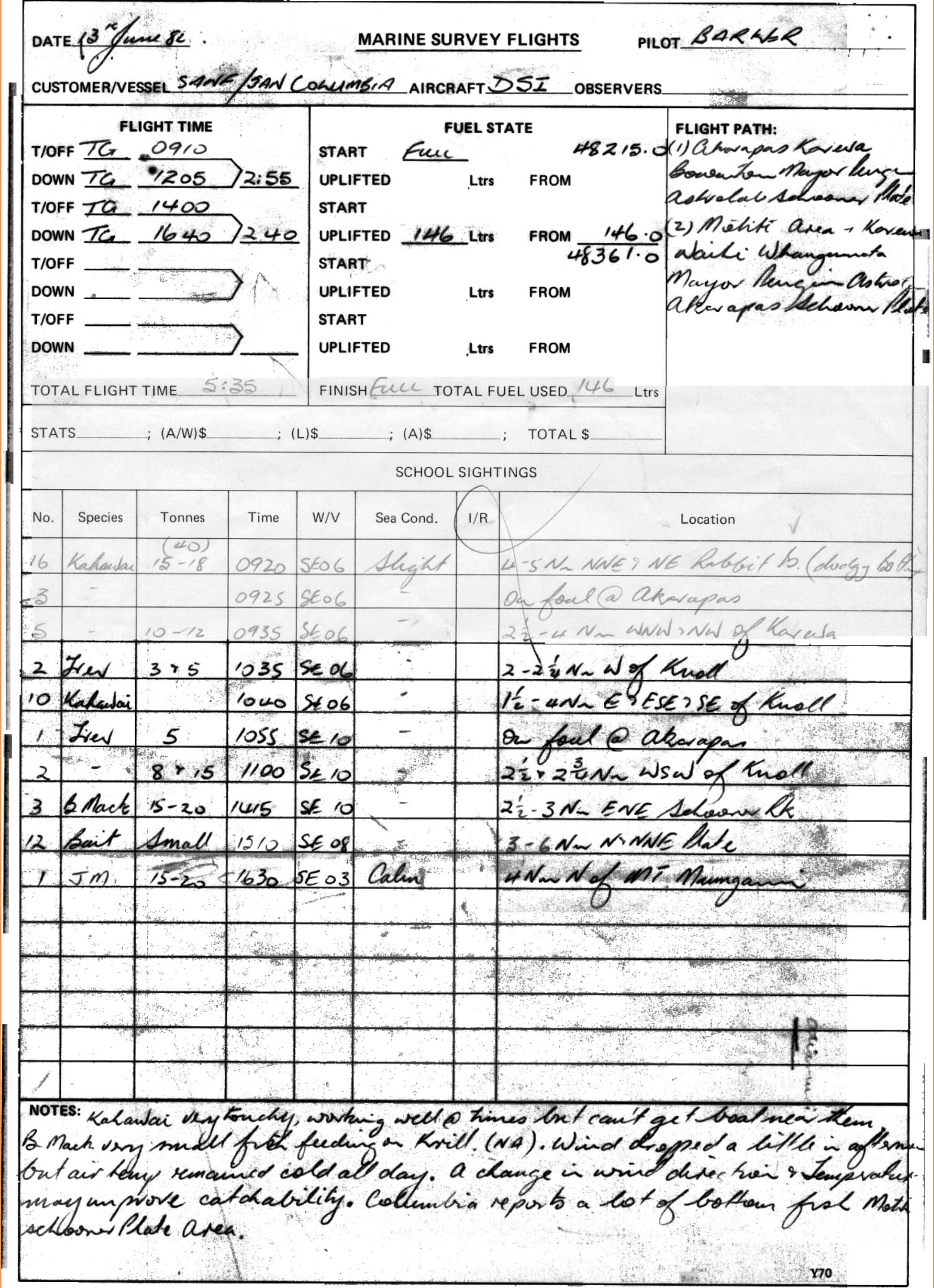
Collection of the data

- Pilots in fixed wing aircraft are an integral part of the purse-seine fishing operation.
- Pilots identify schools of particular size and species composition and assist boats to capture them.
- In addition, pilots opportunistically record sightings of (all?) schools they see.
- The data are stored in a relational database administered by MPI.

History of the data collection

- Data collection since June 1976.
- Two revisions to the data collection form: 1986 and 1998.
- In 1986 a map with grid squares was added for recording flightpath and flying time within the squares.
- In 1998 GPS lat and long could be recorded for each sighting; also, the addition of operational data allowed estimates of pilot error.

The original data collection form – pre 1986



DATE 13th June 86 MARINE SURVEY FLIGHTS PILOT BARKER

CUSTOMER/VESSEL SANF/SAN COLUMBIA AIRCRAFT DSI OBSERVERS

FLIGHT TIME				FUEL STATE				FLIGHT PATH:
T/OFF	TG	0910		START	Full		48215.0	(1) Akarapas Karewa Bowentown Mayor Penguin Astrolabe Schooner Plate
DOWN	TG	1205	2:55	UPLIFTED	Ltrs	FROM		(2) Motiti area – Karewa Waihi Whangamata Mayor Penguin Astrolabe Akarapas Schooner Plate
T/OFF	TG	1400		START				
DOWN	TG	1640	2 40	UPLIFTED	146 Ltrs	FROM	146.0 / 48361.0	
T/OFF				START				
DOWN				UPLIFTED	Ltrs	FROM		
T/OFF				START				
DOWN				UPLIFTED	Ltrs	FROM		

TOTAL FLIGHT TIME 5:35 FINISH Full TOTAL FUEL USED 146 Ltrs

STATS ; (A/W)$; (L)$; (A)$; TOTAL $

SCHOOL SIGHTINGS

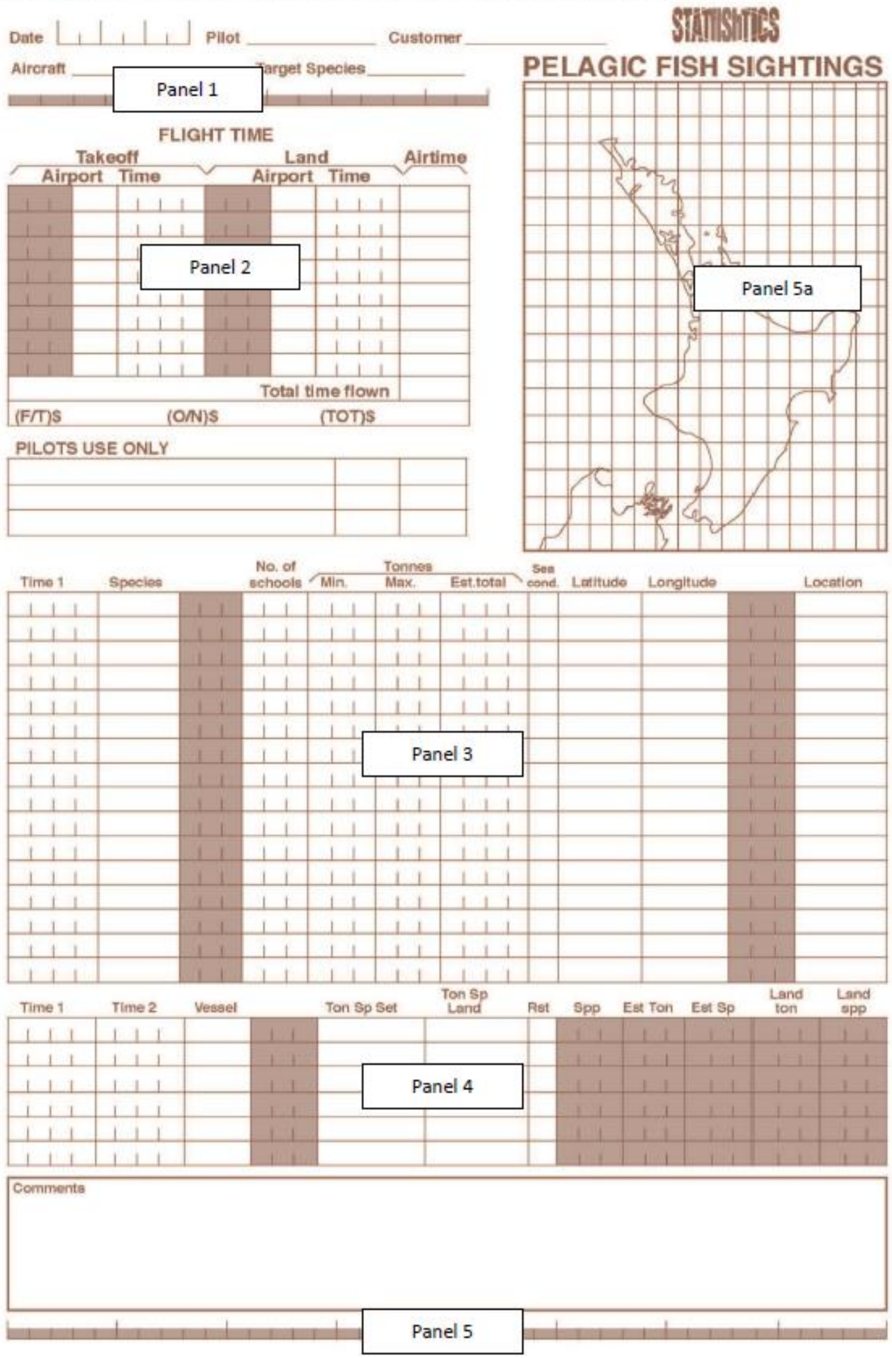
No.	Species	Tonnes	Time	W/V	Sea Cond.	I/R	Location
16	Kahawai	15-18 (40)	0920	SE06	Slight		4-5 Nm NNE & NE Rabbit Is. (dodgy bottom)
3	–		0925	SE06	–		On foul @ Akarapas
5	–	10-12	0935	SE06	–		2½-4 Nm WNW & NW of Karewa
2	Trev	3 & 5	1035	SE06	–		2-2¼ Nm W of Knoll
10	Kahawai		1040	SE06	–		1½-4 Nm E & ESE & SE of Knoll
1	Trev	5	1055	SE10	–		On foul @ Akarapas
2	–	8 & 15	1100	SE10	–		2½ & 2¾ Nm WSW of Knoll
3	B Mack	15-20	1415	SE10	–		2½-3 Nm ENE Schooner Rk
12	Bait	Small	1510	SE08	–		3-6 Nm N & NNE Plate
1	JM	15-20	1630	SE03	Calm		4 Nm N of MT Maunganui

NOTES: Kahawai very touchy, working well @ times but can't get boat near them. B Mack very small fish feeding on Krill (NA). Wind dropped a little in afternoon but air temp remained cold all day. A change in wind direction & temperature may improve catchability. Columbia reports a lot of bottom fish Motiti Schooner Plate area.

Y70

The current data collection form

Figure 12: Current aerial sightings data collection form; see explanation in Appendix A.


STATISHTICS

PELAGIC FISH SIGHTINGS

Date Pilot Customer

Aircraft Target Species

Panel 1

FLIGHT TIME

Takeoff: Airport, Time | Land: Airport, Time | Airtime

Panel 2

Total time flown

(F/T)S (O/N)S (TOT)S

PILOTS USE ONLY

Panel 5a

Time 1	Species		No. of schools	Tonnes Min.	Tonnes Max.	Tonnes Est.total	Sea cond.	Latitude	Longitude		Location

Panel 3

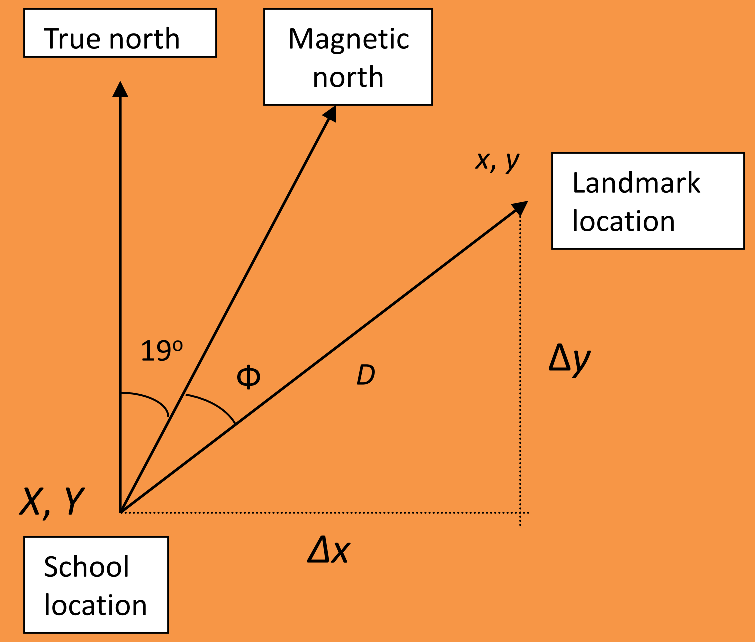
Time 1	Time 2	Vessel		Ton Sp Set	Ton Sp Land	Rst	Spp	Est Ton	Est Sp	Land ton	Land spp

Panel 4

Comments

Panel 5

Generating fine scale position data



Definitions

Information recorded by the spotter pilot.

- 1. *D* is the distance from the fish school to the known landmark.
- 2. Φ is the bearing from the fish school to the known landmark.

Calculations

1. Convert the bearing to be relative to true north (add local variation):

$$\theta = \Phi + 19$$

2. Convert *D* (nmi) to miles: $D = D_{orig} * 1852$

3. Lat/Long for landmarks (assume they are in WGS-84); convert to NZTM‡ (x, y form) in Microsoft Excel.

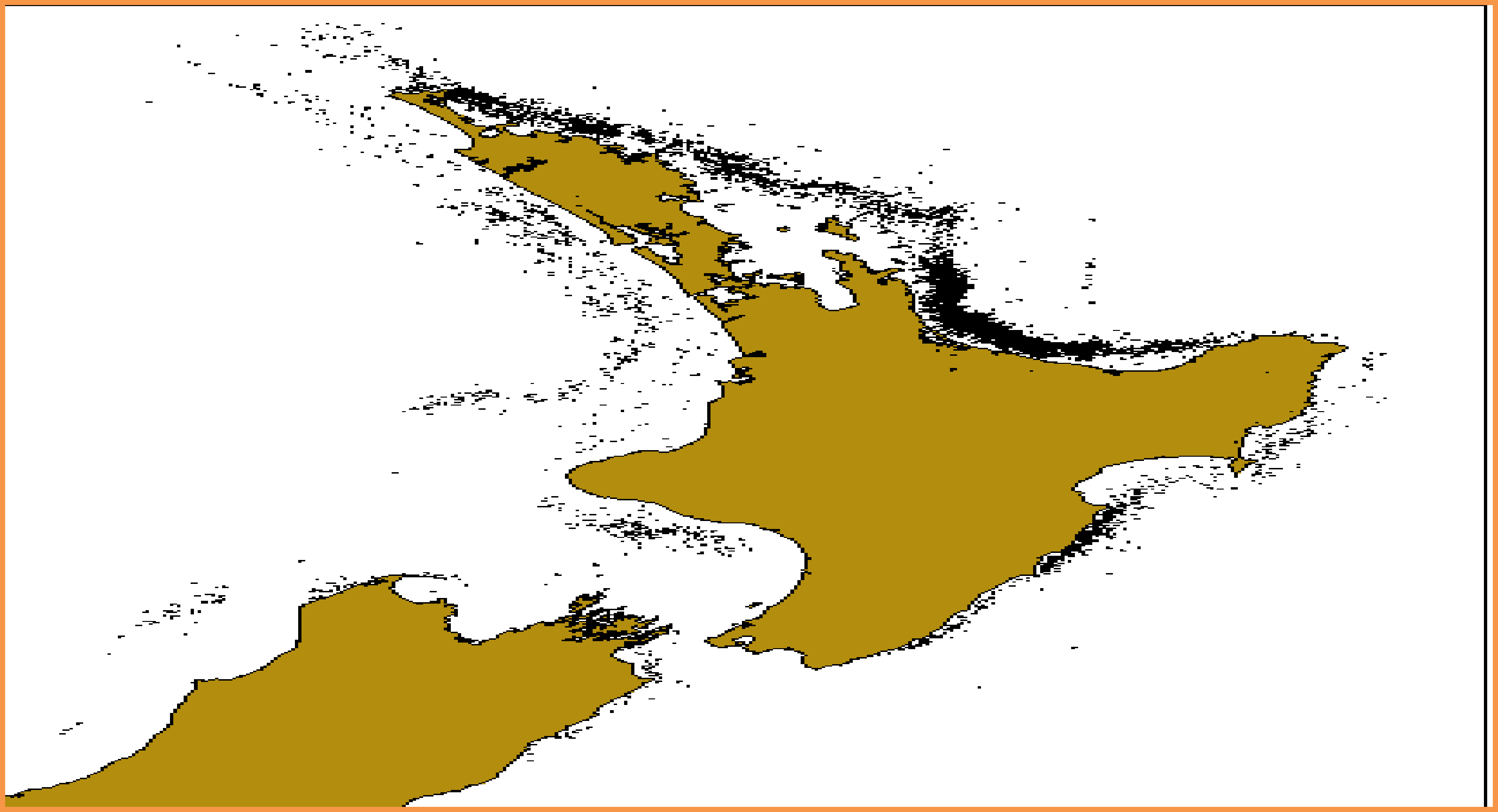
$$\Delta x = D * \text{Sin}[(\Phi + 19) * \text{Pi}()/180]$$
$$\Delta y = D * \text{Cos}[(\Phi + 19) * \text{Pi}()/180]$$

4. $X = x + \Delta x$; $Y = y + \Delta y$; i.e. bearing is from the fish school to the marker.

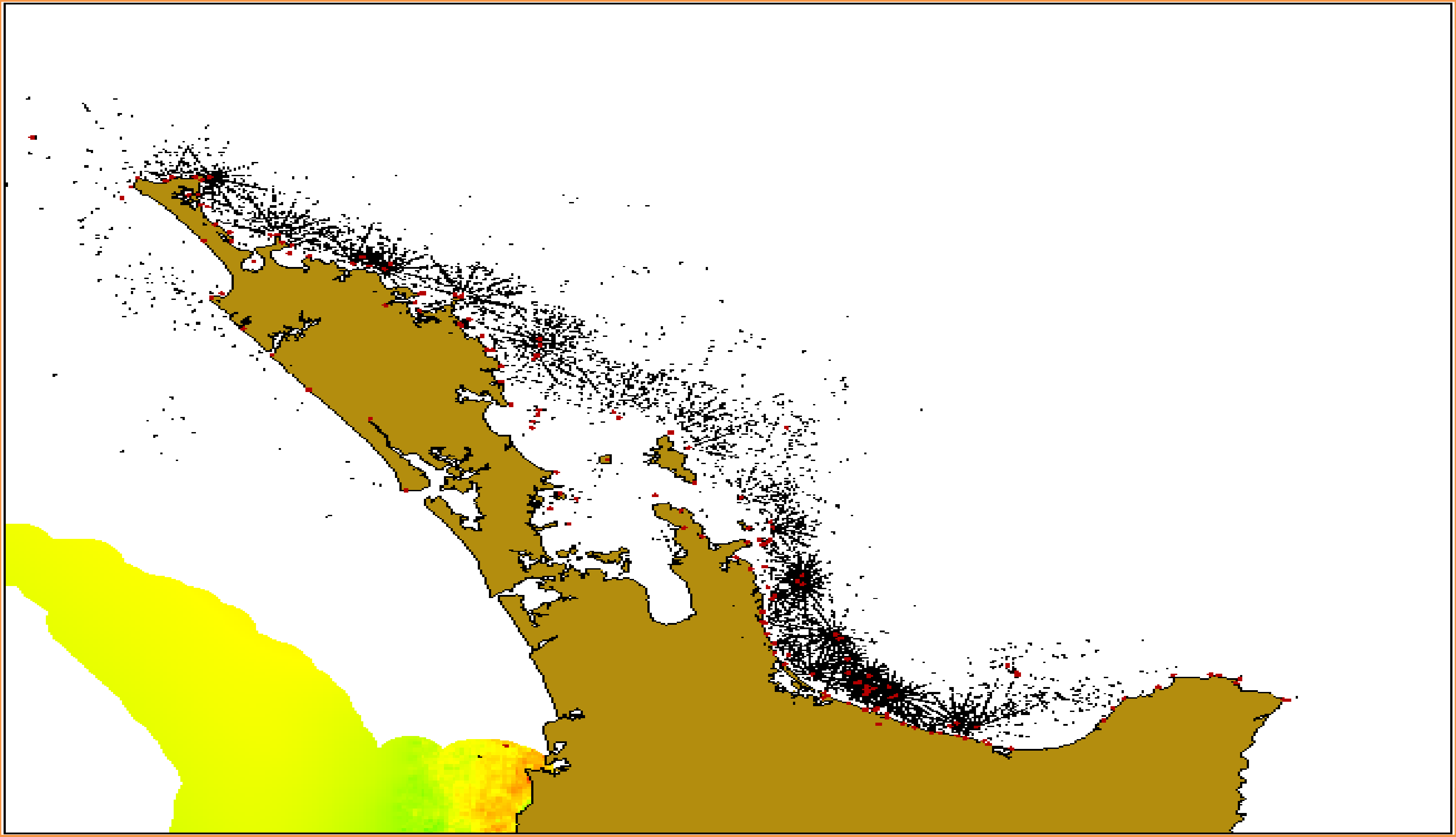
5. Convert *X*, *Y* back to latitude, longitude.

NZTM‡ is the replacement for NZMG, based on WGS-84 datum.

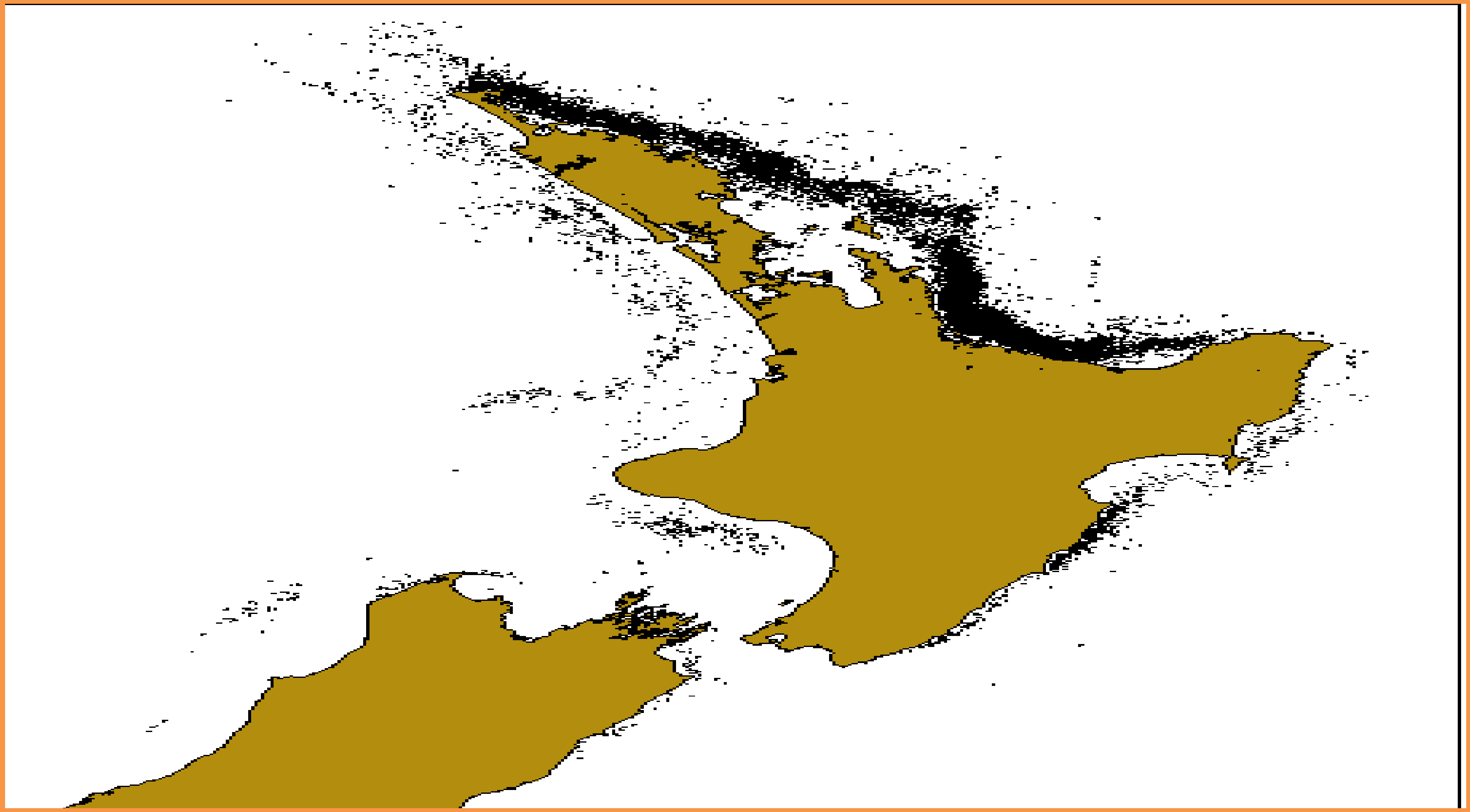
Sightings since 1997 (generated circa 2003)



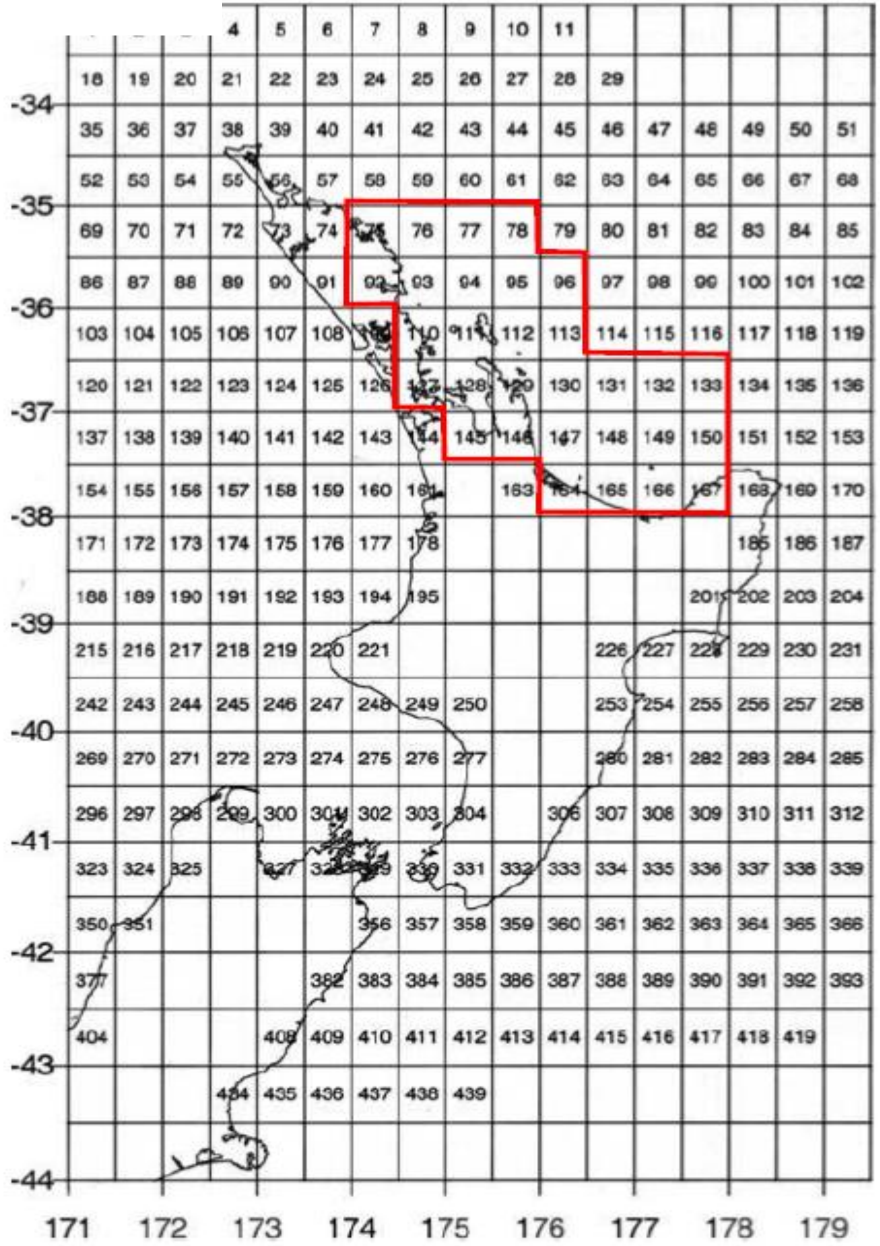
Sightings between 1986 and 1997 (16,571 sightings)



Fine scale data since 1986 (24,796 sightings)



Grid squares & codes; red boundary indicates study area



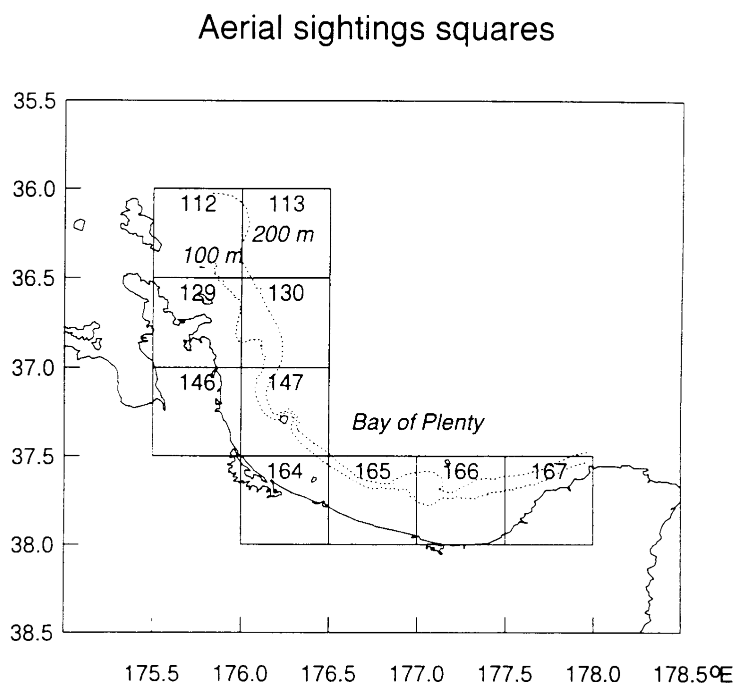


Figure 2: Map of the Bay of Plenty showing the aerial sightings half degree squares defining the Bay of Plenty region. The 100 m and 200 m depth contours are shown as dotted lines.

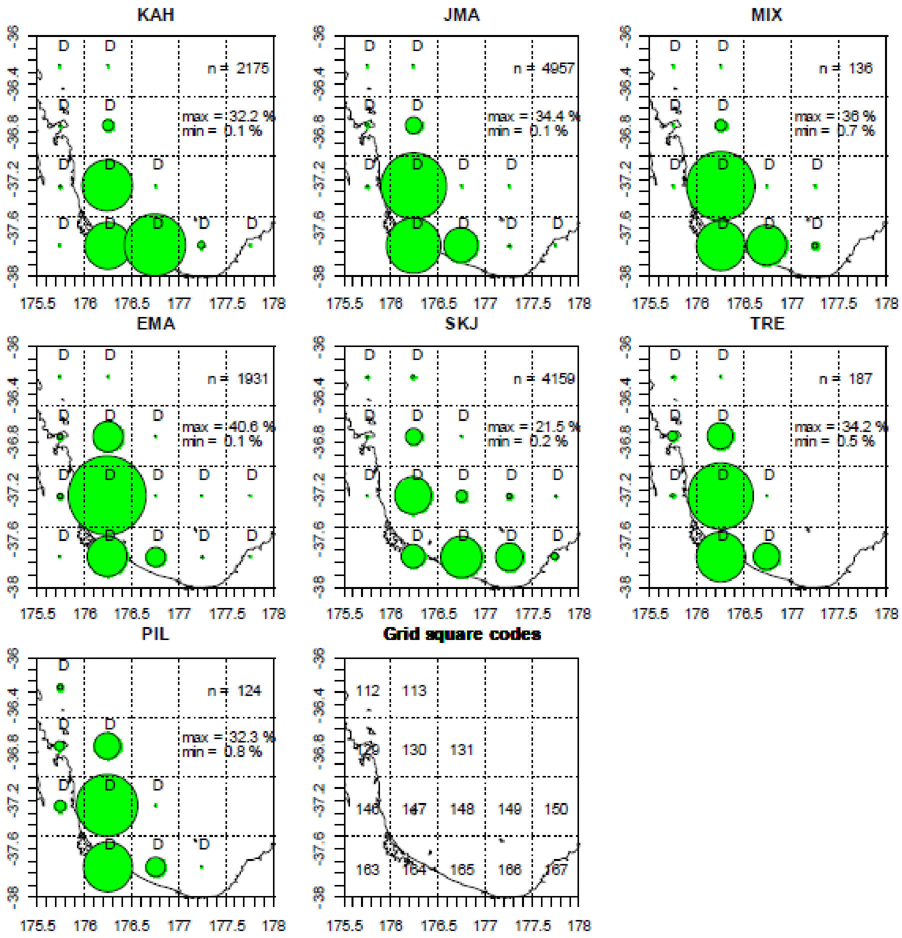


Figure 2: For each target species, flightpath density or the proportion of total flightpath ticks (10–15 min periods) recorded in each grid square visited during all flights throughout the period of interest (January 1986 to September 2011) in the Bay of Plenty; circles are centred on grid squares, their 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.

Previous work

- Until now the only real use of the data has been to produce indices of relative abundance for kahawai and trevally in the Bay of Plenty for use in stock assessment models for these two species.
- Jack mackerel were omitted because 3 species are managed as a single entity; also blue mackerel, when preliminary analyses indicated high interannual variation in relative abundance indices, suggesting that aerial sightings were indexing the abundance of only part of a larger stock present on the fishing grounds.
- For reasons referred to in the factors limiting use of these data, only sightings records for the first flight in the day and only for pilot #2 were used for this work.

Features of the dataset imposing limits to how the data can be used

1. Because they are aggregated over the entire day, flightpath data recorded since January 1986 can only be used with reference to the entire day - they cannot be used in the context of the individual flight. cannot be used as flying effort.
2. As is requested in the instructions for filling out the data collection forms, Pilot #2 “records one mark in the appropriate square on the map for each quarter hour (or part thereof) spent searching for fish. It is clear on forms from some other pilots that this is not done.
3. Changes related to the two revisions of the data-collection form have resulted in the data naturally falling into three sub-series.
4. There was a change in the way sightings were reported from about 1994.
5. Pilot #2 has attempted to avoid double counting on a daily basis by omitting any fish recorded during earlier flights.
6. Where a sighting is based on multiple schools, pilots always record the number of schools together with estimates for size of the largest and smallest. Some pilots (97% for Pilot #2) also include an estimate of the total, which is considered the “best estimate”.
7. Not all flying time is search time - pilots spend time identifying species composition of the schools comprising each sighting, determining the size (tonnage) of the schools, and assisting the vessel(s) to set on the chosen school. This component of non-search time is referred to here as *process time*.
8. Sightings can be divided into two categories based on whether they contain one species (referred to as single species, mono-specific, or pure schools) or more (mixed schools).
9. Target species was not recorded on the forms.

Previous work - strategy for data selection

- Flightpath used to select data according to days flying in a particular area e.g. BoP.
- Data limited to Pilot #2 only, because of the reliability and consistency of his data (e.g. flightpath records).
- Because of his strategy to avoid double counting, 1^{st} flight of the day only used, sometimes 2^{nd} (repositioning).
- East Northland analysis abandoned because minimum number of flights per year (50) criterion not met.
- Effort data (flight duration) adjusted with process data.
- Both mono-specific and mixed school data used.
- Target species generated from purse-seine catch data (Warehou – MPI- and the FSU database - Niwa).

Problematic issues - discussion

1. Low data coverage in east Northland.
2. Adjusting flying effort to search effort.
3. The need to include target species in the current work.

Numbers of flights with target species data available in the BoP 1985-86 to 2010-11

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

Numbers of flights with target species data available in east Northland 1985-86 to 2010-11

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

Adjusting flying effort

- Not all flying time is search time.
- Also process time: determining school size and composition, choosing appropriate school, assisting vessel to set on chosen school.
- Flight time (*feff*) regressed against both the number of fishing operations (*nops*) and the total sightings (*totsit*)

$$feff = b * nops + c * totsit.$$

- The estimated slopes were used to adjust flight time to search time *(efft).*

Previous work - Standardising for sightings per unit effort (SPUE)

- 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.
- 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.

Including purse-seine target species in the standardisation; 1. Kahawai

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)				

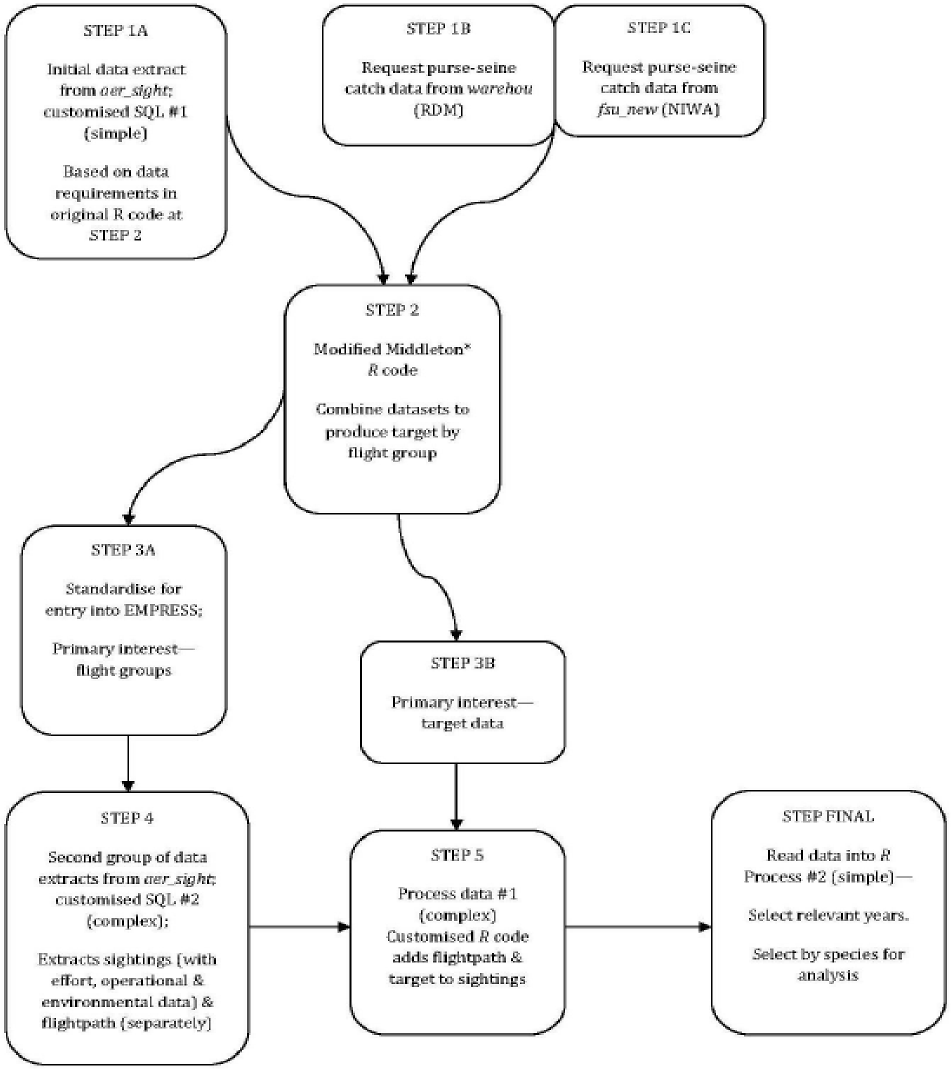
Including purse-seine target species in the standardisation; 2. trevally

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)				

Steps required for data extracts in previous work using *aer_sight* database.

Steps required for data extracts in previous studies using *aer_sight* database.



Current work

- Reasons for omissions of blue and jack mackerels probably irrelevant here.
- Restriction to Pilot #2 may be unnecessary here, but difficulty may arise if we wish to separate analyses into sub areas – problem related to flightpath recording by pilots other than Pilot #2.
- However, using boosted regression tree method may overcome too few data perception in east Northland.
- Relevant data are probably adjusted flight time (effort), sighting time, species, number of schools, total tonnage of sighting, latitude, longitude, as well as catch data from catch-effort database as surrogate for target.
- Need to standardise by flying effort and target species.

Overall strategy

- Investigate ways to increase the dataset size, particularly in east Northland: relaxing the need to separate by area would eliminate requirement for selection using flightpath thus allowing data from more pilots to be included.
- Alternatively, relax the rule of thumb applied in previous work (50 flights/yr minimum) and accept a higher level of uncertainty on the probability estimates.
- Determine the best avenue of extracting data – via RDM or directly from a copy of the database as with Niwa work which would allow the use of existing Empress SQL code; RDM extracts would require SQL code to be rewritten.

Acknowledgements

- John Leathwick.
- Fabrice Stephenson (Niwa):

Stephenson, F., Leathwick, J.R., Geange, S.W., Bulmer, R.H., Hewitt, J.E., Anderson, O.F., Rowden, A.A. and Lundquist, C.J., 2018. Using Gradient Forests to summarize patterns in species turnover across large spatial scales and inform conservation planning. *Diversity and Distributions*, *24*(11), pp.1641-1656.