J Annala

New Zealand Fisheries Assessment Research Document 95/21

Catch per unit effort (CPUE) analysis of the Campbell Island Rise southern blue whiting (*Micromesistius australis*) trawl fishery from 1986 to 1994

J.K.V. Ingerson and S.M. Hanchet
NIWA
P O Box 14–901
Wellington

November 1995

Ministry of Fisheries, Wellington

Catch per unit effort (CPUE) analysis of the Campbell Island Rise southern blue whiting (*Micromesistius australis*) trawl fishery from 1986 to 1994

J.K.V. Ingerson and S.M. Hanchet
New Zealand Fisheries Assessment Research Document 95/21. 15 p.

1. EXECUTIVE SUMMARY

A standardised CPUE analysis of the Campbell Island Rise southern blue whiting trawl fishery from 1986 to 1994 is described. Two generalised linear models were used; a gamma log-link (GLL) model and a combination model.

Catch-effort data were extracted from the fisheries statistics database and included all tows on the Campbell Island Rise that targeted and/or caught southern blue whiting. The units of CPUE were kg per n. mile towed.

A total of 18 variables were introduced to both models, including 2 new categorical variables; time-depth and vessel class. The time-depth variable divides all tows into one of four categories, each being a combination of time of day (day or night) and distance of the net from the bottom. Vessel class divides tows into one of five vessel design categories.

Variables were selected to enter the models in a stepwise procedure until the improvement by adding a new variable was less than 1%. Seven variables were selected into the GLL model. The relative year effect indices showed a steady decline from 1.0 in 1986 to 0.22 in 1992, followed by a sharp increase to 0.57 in 1993 and a small decline to 0.52 in 1994.

The indices derived from the combination model were calculated by combining the indices from two separate analyses; a GLL model carried out on all successful (non-zero) tows, and a binomial model of the proportion of zero tows in each year. The indices resulting from the combination model were very similar to those of the GLL model.

2. INTRODUCTION

Chatterton (FARD in prep) calculated a standardised CPUE index for the period 1986 to 1993 based on aggregations of spawning southern blue whiting on the Campbell Island Rise. This index was obtained using multiple regression analysis of commercial catch and effort data. The models used were a lognormal linear (LNL) model, and two generalised linear models (GLMs), a gamma log-link (GLL) model and a delta lognormal (DLN) model. The two GLMs showed a better fit to the data and treated zero catches in a more plausible manner than the LNL model.

This paper updates the analysis with data from 1986 to 1994, using the two GLMs. The DLN model is more correctly named as the combination model. Further data checking and the inclusion of two new variables are also described.

3. THE DATA

Data for the 1994 fishing season were extracted and checked for errors as detailed in Chatterton (in prep.). All tows were examined for possible errors in recorded net depth. Many vessels, especially between 1986 and 1988, were incorrectly recording net depth on the forms. Instead of entering the depth of the groundrope as net depth, the vessels mistakenly entered depth of the headline. The error was noticed because these vessels never appeared to make any bottom tows (i.e. where net depth equalled bottom depth), whereas most other vessels in all years made many bottom tows. Incorrect tows thus identified were altered by adding the headline height to the recorded net depth. The final dataset contained details of 7318 tows (Table 1).

Table 1: Number of vessels involved in the fishery, the number of tows, and the number of zero catches in each year in the final dataset

Year	1986	1987	1988	1989	1990	1991	1992	1993	1994	Total
No. of vessels	15	17	15	23	33	28	33	20	14	111
No. of tows	893	637	843	1 008	994	1 057	1 091	411	384	7 318
Zero catches	41	35	60	47	78	39	204	44	26	574
(%)	4.6	5.5	7.1	4.7	7.8	3.7	18.7	10.7	6.8	7.8

4. THE VARIABLES

The variables used in the two models are the same as detailed by Chatterton (in prep.), except for two additional categorical variables and the exclusion of target species as a variable. Target species was excluded as there are only four tows in the dataset with a target species other than southern blue whiting, and these were in the first 3 years.

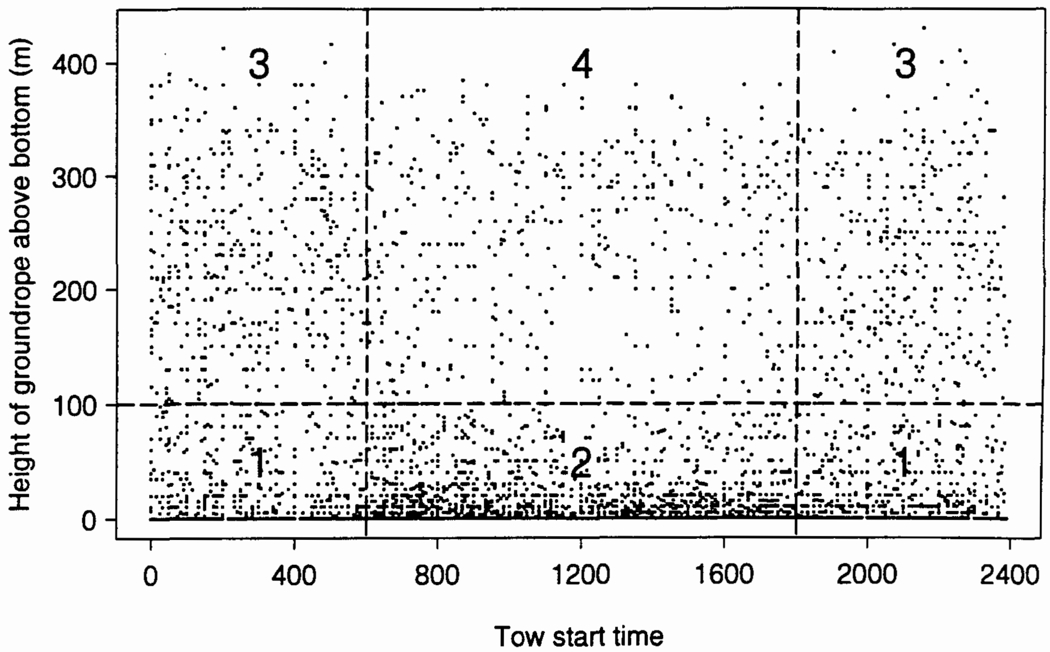

Figure 1: Distance of groundrope from the bottom by time of day for all tows. The dashed lines and numbers within the plot indicate the criteria used to generate the four categories of the time-depth variable.

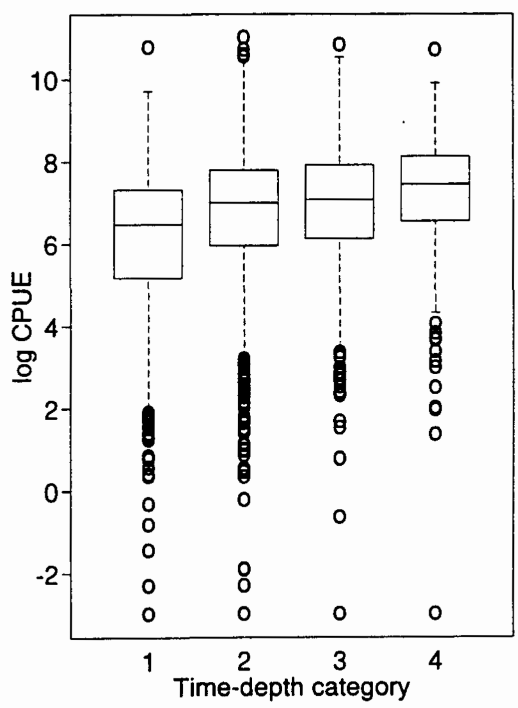

Figure 2: Boxplot of log CPUE for each time-depth category.

The categorical variable time-depth classified tows into four categories reflecting time of day and depth of the groundrope relative to the bottom (Figure 1). Tows were classified as day (tows started between 0600 and 1800 hours), or night (before 0600 or after 1800). These two categories were then split into two depth ranges; tows with the groundrope less than 100 m from the bottom, and tows with the groundrope more than 100 m from the bottom (Figure 1).

There is a gradual increase in CPUE from time-depth category 1 through to category 4. This suggests that CPUE is higher for tows in mid water (categories 3 and 4) than for those close to the bottom (Figure 2). Tows during the day (categories 2 and 4) also have higher CPUE than tows at night in the same depth (categories 1 and 3, respectively).

The depth of the net relative to the bottom changes throughout the season. Relatively more fishing is carried out in mid water around day zero, defined as the day at which 10% of the fish are running ripe, which is considered to be close to the peak of spawning (Chatterton in prep.). This pattern occurs every year (Figure 3), and particularly in 1989 when tows made about the time of peak spawning were almost exclusively in mid water. The curves fitted to the data show that fishing in all other years is still carried out on or near the bottom during this period.

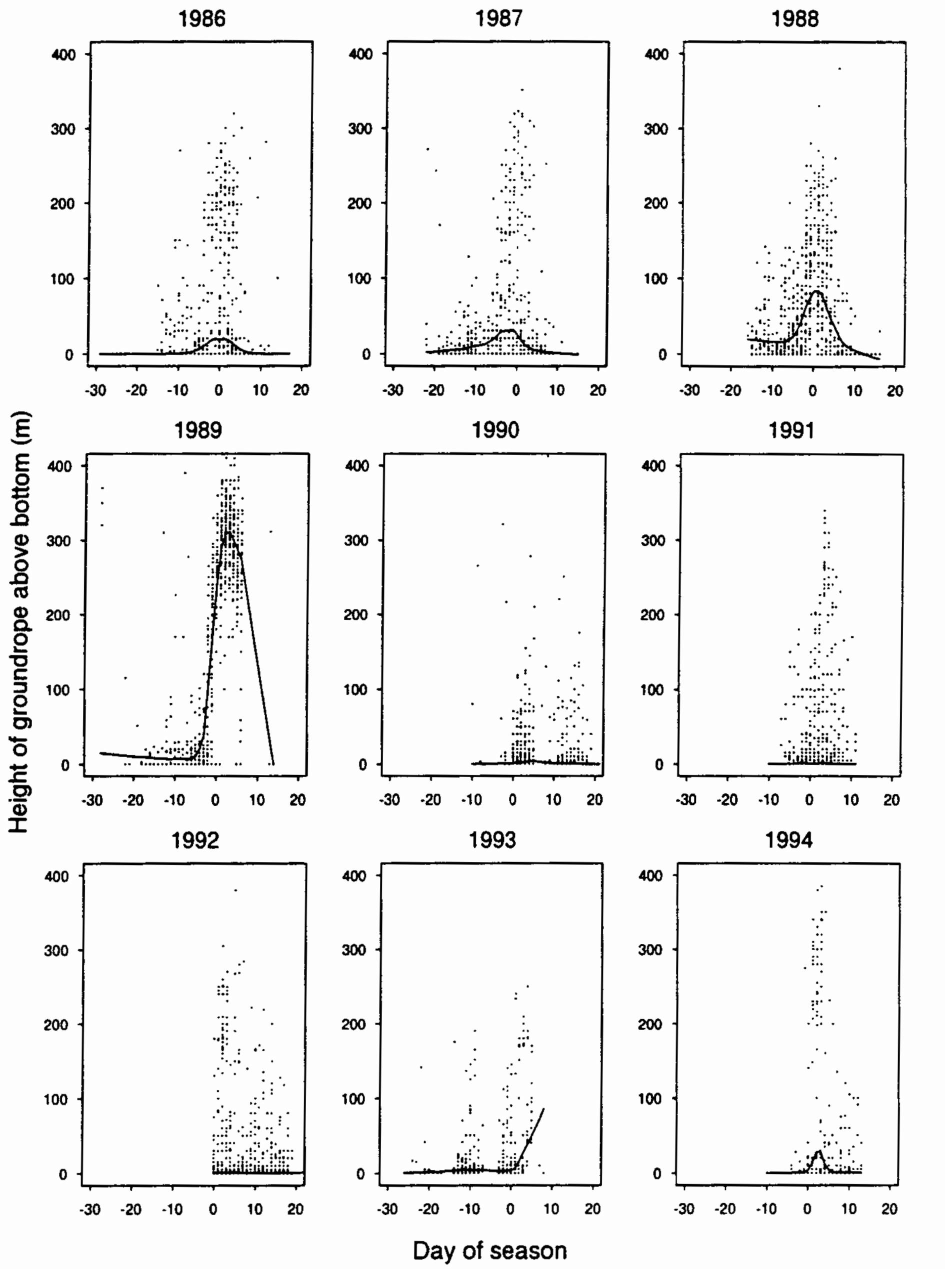

Figure 3: Distance of the net from the bottom (in metres) by day of season for each year, with lowess curves fitted. Day zero corresponds to peak spawning.

The pattern in change of CPUE throughout the season is similar for most years, peaking around day zero (Figure 4). Most of the exceptions to this can be readily explained. The double peak in 1990 represents two separate spawning events; the first on the northern Campbell Island Rise ground and the second a later spawning on the southern ground. In 1994 the rise after day zero is due to most of the vessels moving at day 3 to the southern ground, where the fish spawned at about day 6. Future analyses will include a re-examination of the available spawning data to attempt to remove this double peaked effect. This may improve the explanatory power of the season variable in the analysis.

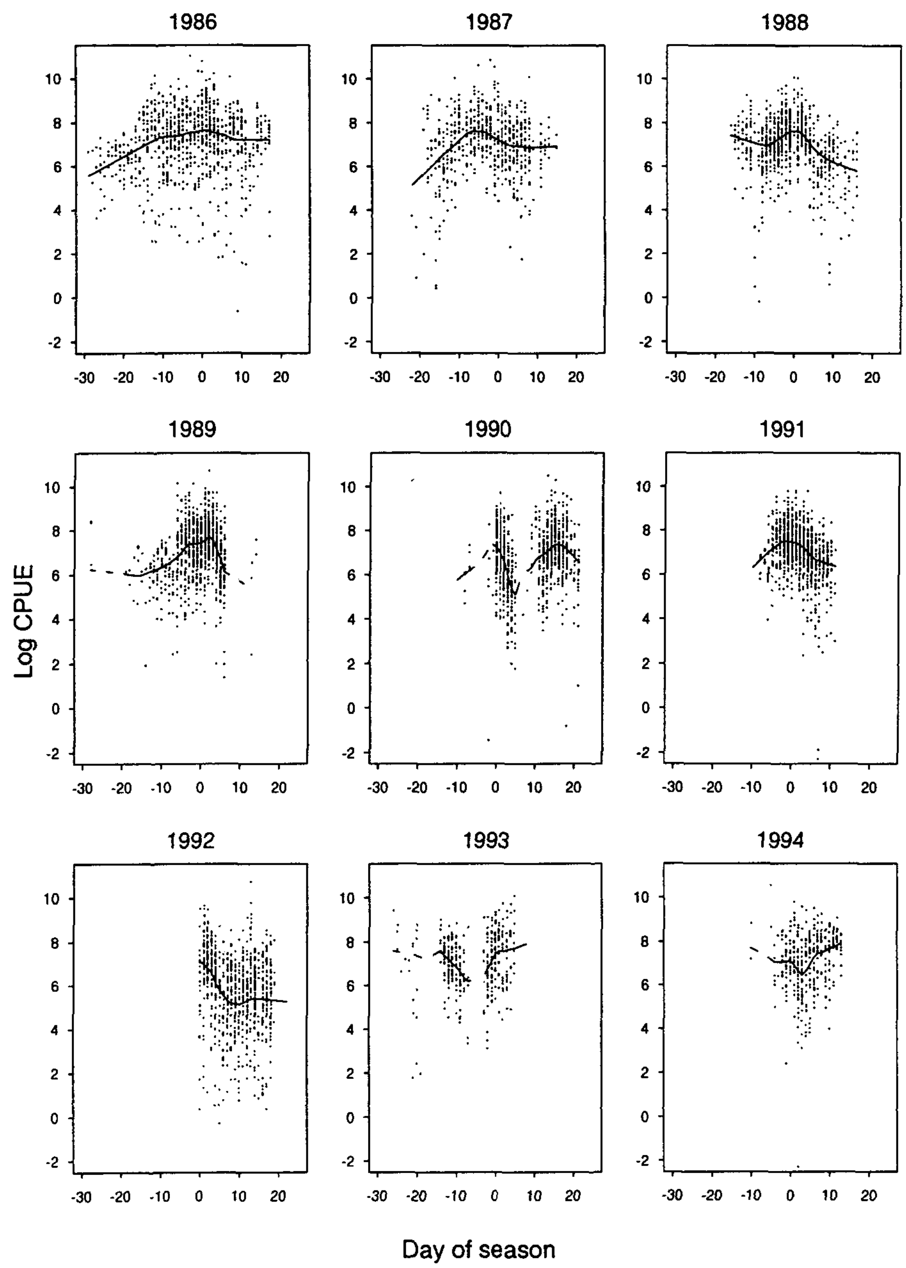

Figure 4: Log CPUE by day of season for each year. Lowess curves are fitted, dashed lines indicating uncertainty from insufficient data points.

The peaks in CPUE by day of season (Figure 4) correspond to the peaks in depth fished by day of season (see Figure 3), suggesting that CPUE increases when more tows are made in mid water. However, CPUE does not show a consistent increase with distance of the net off the bottom in all years (Figure 5). So the corresponding peaks may also be related to some other variable, or combination of variables. If aggregations of fish at peak spawning are denser than at other times, or if the fish are easy to catch, CPUE will increase at this time regardless of what depth is being fished. For example, in 1991 there were relatively few tows off the bottom (illustrated by the flat lowess curve in Figure 3), yet there was still a strong peak in CPUE at day zero (see Figure 4).

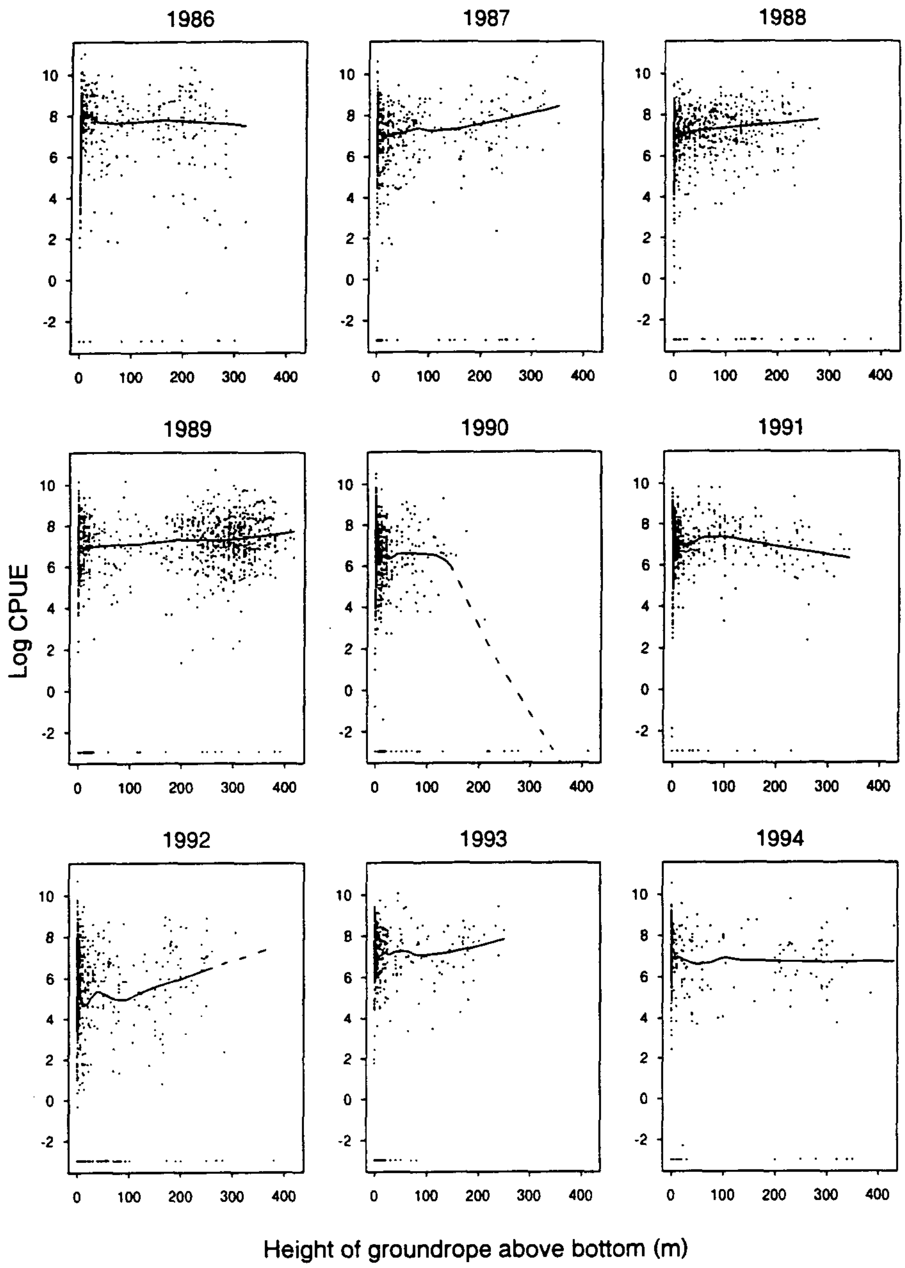

Figure 5: Log CPUE by depth of the net off the bottom for each year. Lowess curves are fitted, dashed lines indicating uncertainty from insufficient data points.

The second new variable, vessel class, reflects an attempt to group 'like' vessels into clusters. A plot of year a vessel was built against vessel length for vessels of different nationalities shows that there are identifiable groups (Figure 6).

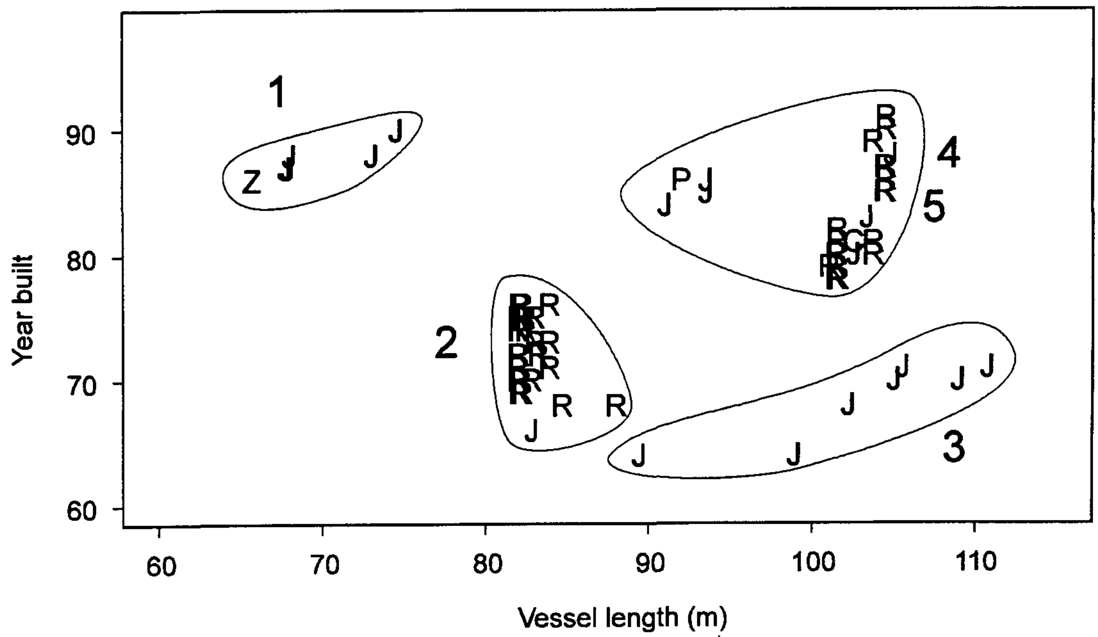

Figure 6: Year vessel built by vessel length. The clusters of vessels indicate the categories for the vessel class variable. Letters represent vessel nation; J = Japan, R = Russia + Soviet, P = Poland, Z = New Zealand, C = China. Categories 4 and 5 lie within the same cluster; 4 is Japanese vessels only and 5 is all others.

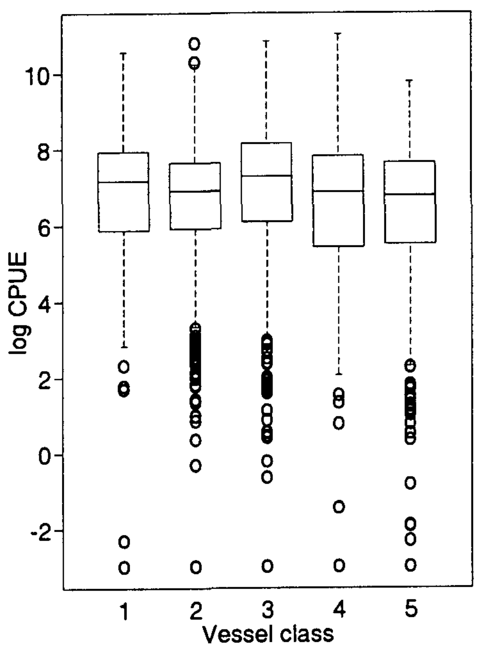

Figure 7: Boxplot of log CPUE for each vessel class category.

These vessel design class categories may replace other vessel variables such as tonnage, nation and year built when considered by the models.

There are no clear differences in CPUE between the five different vessel class categories, but categories 1 and 3 appear to be slightly higher than the others (Figure 7).

A summary of all the variables and their type is shown in Table 2.

Table 2: Summary of variables used in models. cont = continuous variables, cat = categorical variables with a given number of categories

Variable	Type	Description
year	cat 9	Year in which tow occurred
tow position	cat 4	Position at start of tow
start time	cat 6	Time at start of tow in 4 h periods (e.g. 0001–0400 h)
end time	cat 6	Time at end of tow in 4 h periods (e.g.0001–0400 h)
net depth	cont	Depth (m) of groundrope at start of tow
bottom depth	cont	Depth (m) of sea bed at start of tow
time-depth	cat 4	Relative depth of net at day and night
season	cont	Day relative to date at which 10% of fish are spawning in that year
headline height	cat 4	Height (m) of headline above groundrope
vessel class	cat 5	Design class of vessel (derived from length and year built)
vessel nation	cat 6	Country of origin of vessel
vessel year built	cont	Year the vessel was built
vessel tonnage	cont	Gross tonnage of the vessel
vessel length	cont	Length (m) overall of the vessel
vessel breadth	cont	Breadth (m) of the vessel
vessel draught	cont	Draught (m) of the vessel
vessel l*b*d	cont	Vessel length * breadth * draught in cubic metres
vessel company	cat 23	Company the vessel was fishing for in that year

5. THE MODELS

The two GLM models employed for the analysis were a gamma log-link (GLL) model and a combination model (Vignaux 1994). The lognormal linear model analysis of the data in 1994 (Chatterton in prep.) was not repeated as the GLL and combination models fit the data better and handle zeros in a more plausible manner (Chatterton 1995).

5.1 Gamma log-link (GLL) model

In the GLL model, CPUE was modelled using a log-link function and a gamma error distribution (producing a variance proportional to the square of the mean). Variables were selected to enter the model in a stepwise procedure until the addition of another variable led to a less than 1% improvement in the deviance explained (Chatterton in prep.). Seven variables were selected in this way (Table 3). Year was the most significant variable, followed by vessel length, time-depth, tow position, season, headline height, and end time. The model failed to converge when considering the bottom depth variable in iterations 2 to 4. Table 4 compares this year's variables to those selected in last year's analysis. Year is the first variable to enter the model in both analyses. Vessel length is the only vessel-related variable to enter the model in the 1994 analysis, and vessel nation and tonnage are no longer significant (Table 3). The new variable 'time-depth' enters the model this year, and its presence probably replaces in importance the component variables time and depth compared to 1993. The model explains 18.0% of the variability in CPUE.

Table 3: Gamma log-link (GLL) model. Choice of variables, in order of importance, in model of CPUE data from 1986 to 1994. The variables above the line were selected for use in the model

	Residual deviance at iteration							
Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
year	8 675							
vessel length	8 957	8 375						
time-depth	8 790	8 385	8 057					
tow position	8 982	8 494	8 224	7 904				
season	8 915	8 550	8 247	7 933	7 748			
headline height	9 018	8 520	8 236	7 913	7 754	7 608		
end time	8 944	8 473	8 159	7 918	7 762	7 609	7 471	
net depth	9 011	8 555	8 247	7 972	7 788	7 647	7 513	7 394
nation	8 875	8 480	8 318	8 000	7 846	7 695	7 545	7 413
vessel class	8 931	8 481	8 337	8 020	7 865	7 716	7 564	7 428
vessel tonnage	9 012	8 403	8 336	8 018	7 865	7 711	7 578	7 442
bottom depth	9 092	379 094	$33*10^{18}$	$72*10^{18}$	7 857	7 716	7 573	7 445
vessel breadth	8 931	8 399	8 332	8 018	7 861	7 712	7 584	7 449
start time	8 967	8 485	8 177	7 963	7 806	7 658	7 520	7 455
vessel draught	8 948	8 468	8 358	8 048	7 895	7 739	7 591	7 459
vessel l*b*d	8 934	8 433	8 357	8 050	7 897	7 741	7 592	7 459
vessel year built	9 108	8 602	8 357	8 044	7 894	7 741	7 594	7 460
vessel company	8 968	8 659	8 370	8 055	7 901	7 746	7 602	7 466
Percent improvement		**3.3**	**3.5**	**1.7**	**1.6**	**1.5**	**1.5**	**0.8**

Table 4: Comparison of variables and their order of entry into the GLL model for analyses carried out on data 1986 to 1993 and data 1986 to 1994

Variables	1993	1994
1	year	year
2	vessel nation	vessel length
3	tow position	time-depth
4	end time	tow position
5	net depth	season
6	vessel tonnage	headline height
7	season	end time

The relative year effect indices calculated by the model showed a steady decline from 1.0 in 1986 to 0.22 in 1992, rose sharply to 0.57 in 1993, then dropped slightly to 0.52 in 1994 (Figure 8). The indices from this model are very similar to those from last year's analysis (Chatterton in prep.), apart from a noticeable drop in 1993 (Figure 9). When the variables used in 1993 were applied to the entire dataset up to 1994, the indices were almost identical to those calculated last year apart from a slight drop in the 1993 index.

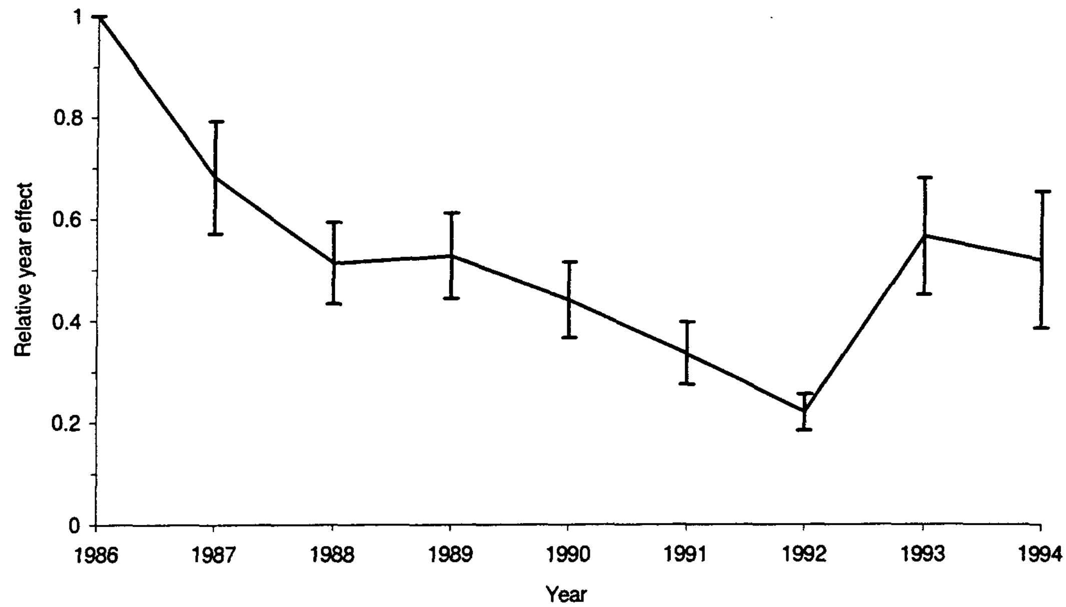

Figure 8: Gamma log link (GLL) model relative year effects ± 2 standard errors.

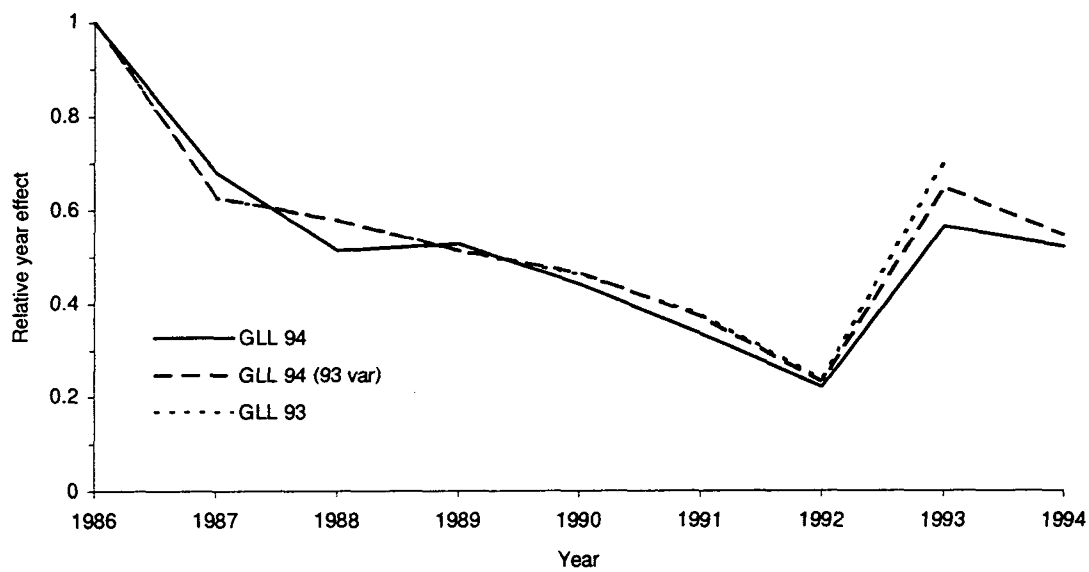

Figure 9: Relative year effects from GLL model for 1986 to 1994 data (GLL 94), 1986 to 1994 data using variables from last year's analysis (GLL 94 (93 var)) and 1986 to 1993 data (GLL 93).

There are concerns over the accuracy of the estimate of peak spawning in 1992. When the vessels arrived on the grounds more than 10% of fish were already running ripe (Hanchet 1993), so it is likely that day zero was 2–3 days earlier than that used in the model. The 1992 data also have the highest proportion of unsuccessful tows, and the lowest CPUE index. The reasons for this are still not known. A sensitivity analysis was carried out using the GLL model to see the effect on the indices if the day of peak spawning in 1992 was actually 3 days earlier. The effect was only very slight, with the same variables entering in the same order and the relative year effect index in 1992 increasing from 0.22 to 0.23.

5.2 Combination Model

The combination model (Delta lognormal model in Chatterton in prep.), is made up of a GLL model of the catch rate of successful (non zero) tows and a binomial model of the proportion of successful tows. The results of the two models are then combined and relative year effects calculated (Vignaux 1994).

5.2.1 GLL model of successful tows

Seven variables were selected to enter the model, and these were the same and entered in the same order (Table 6) as in the GLL model detailed in Section 5.1 above. The relative year effect indices are also quite similar to those calculated in the GLL model of all tows (Table 7).

Table 6: GLL model run on successful tows as part of combination model. Choice of variables in order of importance in model of CPUE data from 1986 to 1994. The variables above the line were selected for use in the model

	Residual deviance at iteration							
Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
year	9 736							
vessel length	10 036	9 435						
time-depth	9 869	9 445	9 115					
tow position	10 058	9 552	9 280	8 960				
season	9 994	9 609	9 306	8 990	8 803			
headline height	10 097	9 579	9 294	8 969	8 808	8 661		
end time	10 023	9 532	9 216	8 975	8 816	8 661	8 521	
net depth	10 091	9 616	9 306	9 030	8 843	8 699	8 564	8 442
vessel nation	9 951	9 537	9 374	9 057	8 898	8 744	8 595	8 460
vessel class	10 006	9 539	9 394	9 076	8 918	8 767	8 615	8 476
bottom depth	10 165	9 729	9 420	9 082	8 908	8 765	8 621	8 490
vessel tonnage	10 092	9 463	9 395	9 076	8 921	8 765	8 631	8 492
vessel breadth	10 010	9 458	9 391	9 075	8 916	8 764	8 635	8 498
start time	10 046	9 544	9 235	9 020	8 860	8 710	8 570	8 505
vessel year built	10 186	9 661	9 416	9 102	8 949	8 793	8 645	8 507
vessel draught	10 026	9 527	9 415	9 105	8 950	8 792	8 643	8 508
vessel l*b*d	10 013	9 492	9 415	9 108	8 952	8 794	8 643	8 508
vessel company	10 046	9 720	9 429	9 112	8 957	8 799	8 655	8 516
Percent improvement		**3.0**	**3.1**	**1.5**	**1.5**	**1.4**	**1.4**	**0.8**

Table 7: Proportion of zero tows and the relative year effect indices derived from the various models. The second column of proportion of zero tows indicates the proportion relative to 1986

	Proportion of zero tows		GLL model		Binomial	Combination
Year	**Raw**	**Relative**	**All tows**	**non zero tows**	**model**	**model**
1986	0.046	1.00	1.00	1.00	1.00	1.00
1987	0.053	1.15	0.68	0.70	1.39	0.68
1988	0.071	1.54	0.51	0.53	1.52	0.52
1989	0.047	1.02	0.53	0.54	1.08	0.54
1990	0.078	1.70	0.44	0.48	1.60	0.46
1991	0.037	0.80	0.34	0.33	0.80	0.33
1992	0.187	4.07	0.22	0.27	4.63	0.23
1993	0.107	2.33	0.57	0.60	2.50	0.57
1994	0.068	1.48	0.52	0.53	1.27	0.52

5.2.2 Binomial model of proportion of successful tows

A binomial model was used to model the proportion of successful tows. Three variables were included in the model (Table 8). Year was again the most significant variable, followed by bottom depth, and finally the new vessel class variable. The first two variables are the same as last year's analysis, and the third and fourth variables from last year (vessel nation and season) are replaced by vessel class. The presence of vessel class makes the nation variable very non-significant (second to last), perhaps indicating that the two variables are closely linked.

Table 8: Binomial model as part of combination model. Choice of variables (above the line), in order of importance, in model of CPUE data from 1986 to 1994.

	Residual deviance at iteration			
Variable	**1**	**2**	**3**	**4**
year	3 818			
bottom depth	3 956	3 764		
vessel class	3 950	3 784	3 720	
season	3 992	3 786	3 734	3 682
end time	3 994	3 790	3 734	3 689
vessel year built	3 994	3 797	3 742	3 698
tow position	3 980	3 783	3 749	3 705
vessel length	4 008	3 802	3 743	3 709
vessel breadth	3 991	3 789	3 728	3 710
vessel company	3 993	3 814	3 758	3 711
vessel draught	3 984	3 786	3 724	3 711
headline height	4 009	3 805	3 752	3 712
vessel l*b*d	3 979	3 784	3 722	3 712
vessel tonnage	4 015	3 809	3 753	3 714
time-depth	4 006	3 812	3 761	3 716
net depth	4 014	3 808	3 763	3 717
nation	3 980	3 789	3 727	3 717
start time	4 009	3 803	3 748	3 704
Percent improvement		**1.3**	**1.1**	**0.8**

5.2.3 Combining the GLL and binomial models

The method for combining the two models is fully documented in appendix 3 of Vignaux (1994). The relative year effect indices derived from the combined model (Figure 10) differ only very slightly from the GLL model detailed in section 5.1 above (Figure 11 and Table 7).

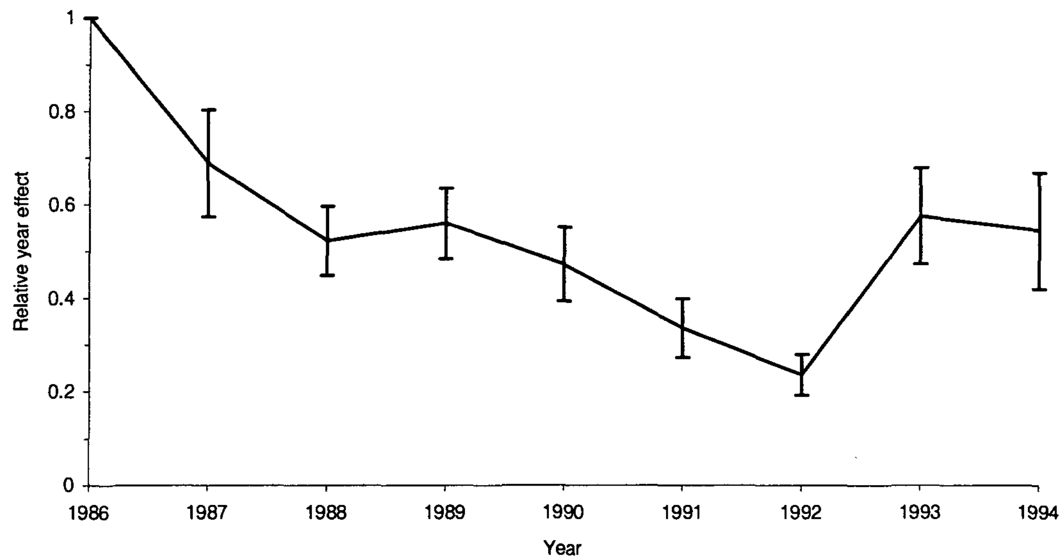

Figure 10: Combination model relative year effects ± 2 standard errors.

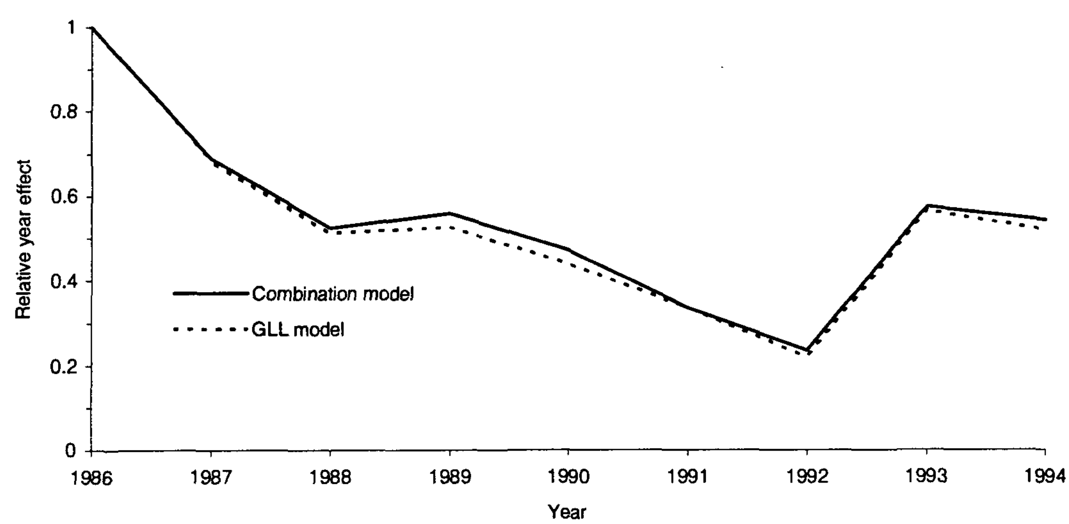

Figure 11: Relative year effect indices of combination and GLL models (CPUE data 1986 to 1994)

6. DISCUSSION

A number of minor changes have been made in this analysis compared to last year. The data have been checked for a previously overlooked error in recorded net depth and two more variables have been added to the analyses.

The time-depth variable attempts to classify tows into distinct groups reflecting likely changes in fish distribution in the water column. It is also likely that the distribution (in time and depth) of tows reflects this change in fish distribution throughout the season. Tows in time-depth category 4 (tows off the bottom during the day) occur almost exclusively at peak spawning, so the high CPUE values for this category could be due to the time and depth of the tows, but are more likely to be due to the time of season. This is also supported by evidence suggesting that CPUE remains high at peak spawning regardless of the depth fished.

There was very little difference in the CPUE between the five vessel class categories, but its presence in the binomial model suggests that it replaces other vessel related variables.

Even though two new variables entered the models, their presence did not improve the explanatory power of the models compared to the variables used in the models last year. The similarity in results when models are run with different variables suggests that the actual choice of variables makes little difference to either the relative year effects or the explanatory power of the model.

This year's analysis has lowered the 1993 index, making the increase from 1992 to 1993 less of a contrast from the steady decline in indices from 1986 to 1992.

A number of additions and improvements could be made to this analysis if it is repeated in future years. Variables which try to model skipper experience or search time would be useful, but are almost impossible to define. Tests for interaction terms in the models could be made and the fits of the models checked.

Chatterton (in prep.) suggested that the combination model (called DLN in his paper) would be more appropriate for this fishery given the increasing proportion of zero tows. However, the proportion of zero tows in 1994 did not show this continuing increase. Further investigation of zero catches is needed before the most appropriate model can be determined.

7. ACKNOWLEDGMENTS

Thanks to Marianne Vignaux for her mathematical and computing advice.

8. REFERENCES

Hanchet, S.M. 1993: Southern blue whiting (*Micromesistius australis*) fishery assessment for the 1993–94 fishing year. New Zealand Fisheries Assessment Research Document 93/17. 56p.

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