*New Zealand Fisheries Assessment Research Document 93/14*

# Catch per unit of effort (CPUE) analysis of the hoki fishery, 1987–92

M. Vignaux
MAF Fisheries Greta Point
P.O. Box 297
Wellington

October 1993

MAF Fisheries, N.Z. Ministry of Agriculture and Fisheries

# Catch per unit of effort (CPUE) analysis of the hoki fishery, 1987–92

M. Vignaux

## 1 Executive summary

A standardised CPUE index was calculated for the hoki fishery based on winter spawning aggregations on the west coast of the South Island (WCSI). This index is calculated from commercial catch and effort data, and has been standardised for effects such as changing vessel or net sizes using a multiple regression technique. It should be used as an index of abundance for the western stock of hoki. The index shows a decline from 1987 (an index of 1) to 1991 (0.48) followed by an increase in 1992 (0.64).

# 2 Introduction

In 1992 a standardised CPUE index was calculated for the hoki fishery based on spawning aggregations on the west coast of the South Island (WCSI) in winter (Vignaux 1992). This CPUE index was based on commercial catch and effort data, and was standardised for effects such as changing vessel size or net sizes using a multiple regression technique. This paper describes a recalculation of that index to include the 1992 season, with a number of modifications and improvements to the method. The multiple regression technique was described in detail by Vignaux (1992).

# 3 The data

## 3.1 Source of data

The catch-effort data came from the fishery statistics database administered by the Information Technology Directorate. Records from all tows on the WCSI between 1987 and 1992 that targeted hoki in the area between longitudes 167° E to 172° E and north of latitude 46° S and after day 150 of the year (about the start of June) were selected from the database. This definition is slightly different from last year's analysis, and ensures that all tows on the WCSI are included, while tows on the Chatham Rise, in Cook Strait and at Puysegur are excluded. The change in definition affects only a small number of outliers.

## 3.2 Contents

For each tow, the following information was available.

- Vessel identification number
- Date
- Time of the end of the tow
- Estimated hoki catch (kg)
- Duration of tow (hours)
- Speed of tow (knots)
- Latitude (decimal degrees)
- Depth of bottom (m)
- Depth of groundrope (m)
- Headline height of net (m)

## 3.3 Error checks

These data have been carefully checked for anomalies. If the cause of the anomaly was clear, the record was altered so that it was consistent with surrounding records and other information (such as gear type, end position, etc). Otherwise, the original paper record was examined to determine the cause of the error. It this was unavailable, the record has been deleted.

Some examples of easily detected errors are:

- Sudden increases or decreases in headline height with no change in gear used – usually mistyped so that in error by a factor of 10.
- Speed very high or very low – usually a 4 misread as a 7
- Sudden changes in latitude or longitude – usually in error by 1 degree
- Duration very long – usually either the end time or the start time mistyped; occasionally the end time and the start time swapped over. Note that a long duration (over 15 hours) may not be an error, so where possible they were checked.
- Depth of net or depth of bottom very deep or very shallow – often mistyped so that in error by a factor of 10, or the headline height entered in the depth field

  More often, the depth of the net was missing or was greater than the depth of the bottom. Where the depth of the net was missing it was assumed that the net was on the bottom. Where the depth of the net was greater than the depth of the bottom it was assumed that the two depths had been swapped.
- Catch extremely large – usually mistyped so that in error by a factor of 10 (These catches were 250 000 kg, 500 000 kg, and 729 000 kg, which would have been impossible for these vessels to have brought on board)
- Duplicate records (only two cases)

## 3.4 Vessel data

For each vessel, this information was available:

- Nationality of vessel
- Overall length (m)
- Gross Tonnage
- Year built
- Power of engines (kW)
- Breadth of vessel (m)
- Draught of vessel (m)
- Processing capability of vessel
- Company vessel was registered to in each of the years 1987–1992

Most of this data came from the vessel registration database. Missing, or apparently anomalous values, were checked against Lloyds Register(Lloyds 1990). Of 210 vessels in the database, 75 were checked. Of these, 13 had incorrectly recorded tonnages, and 15 had incorrectly recorded lengths.

A plot of tonnage against the year a vessel was built for vessels of different nationality (Figure 1) suggests that there are identifiable groups of vessels built by different nations at different times. For example, the distinct set of U marks in the top right of the plot are all vessels in the Soviet Meridian class. These design classes may contain some of the explaining power attributed to the vessel nationality. However, another important component of the fishing power of vessels of different nations is likely to be fishing strategy or fishing skill, about which we have no data.

The processing capability of the vessel was obtained from data held by MAF Quality Management. Vessels that have no processing licences must land fish whole (green); those licenced to pack fish can remove heads and guts, but they must have a full packhouse licence to fillet the fish or produce surimi. Therefore, vessels were coded as I (landing the fish green), L (licensed to partially process), or P (licensed for full processing). Vessels which landed some surimi in the period were coded S. On advice from industry, three vessels capable of processing to surimi were recoded as P in 1992 as they had mainly processed to fillets due to the poor price of surimi (G. Patchell, Sealord Products Ltd. pers. comm.).

A plot of vessel engine power against tonnage for vessels with different processing capabilities (Figure 2) suggests that there is an association between vessel size and processing capability : the smallest vessels do not process at all, and the largest process mainly to surimi. This may mean that the variability in CPUE that could be explained by one variable, such as processing capability, can be explained equally well by another variable such as vessel engine power or tonnage. This correlation between the explanatory variables means that the choice of variable included in a particular model may be arbitrary. Once one of the variables is included, there will be little explanatory power in a second, highly correlated variable, so this other variable is unlikely to be included as well. However, whichever of the highly correlated variables is chosen, the year effect should not be greatly affected.

## 3.5 Effort definition

The original definitions of the "Start time" and "End time" of a tow were the times that the net reached fishing depths and that hauling commenced respectively. Between the 1990 and the 1991 seasons, MAF decided that the fields "Start Time" and "End Time" on the trawl catch effort and processing returns were to be interpreted as the times that the net went into the water and that it was retrieved from the water, respectively. No extra instruction sheets were distributed to fishers, so it is not clear which fishers have been informed of this decision, and which are still filling in the forms according to the old definition. It is not possible to contact each of the fishers individually, but we do know that the skippers are instructed in the completion of the forms by representatives of the

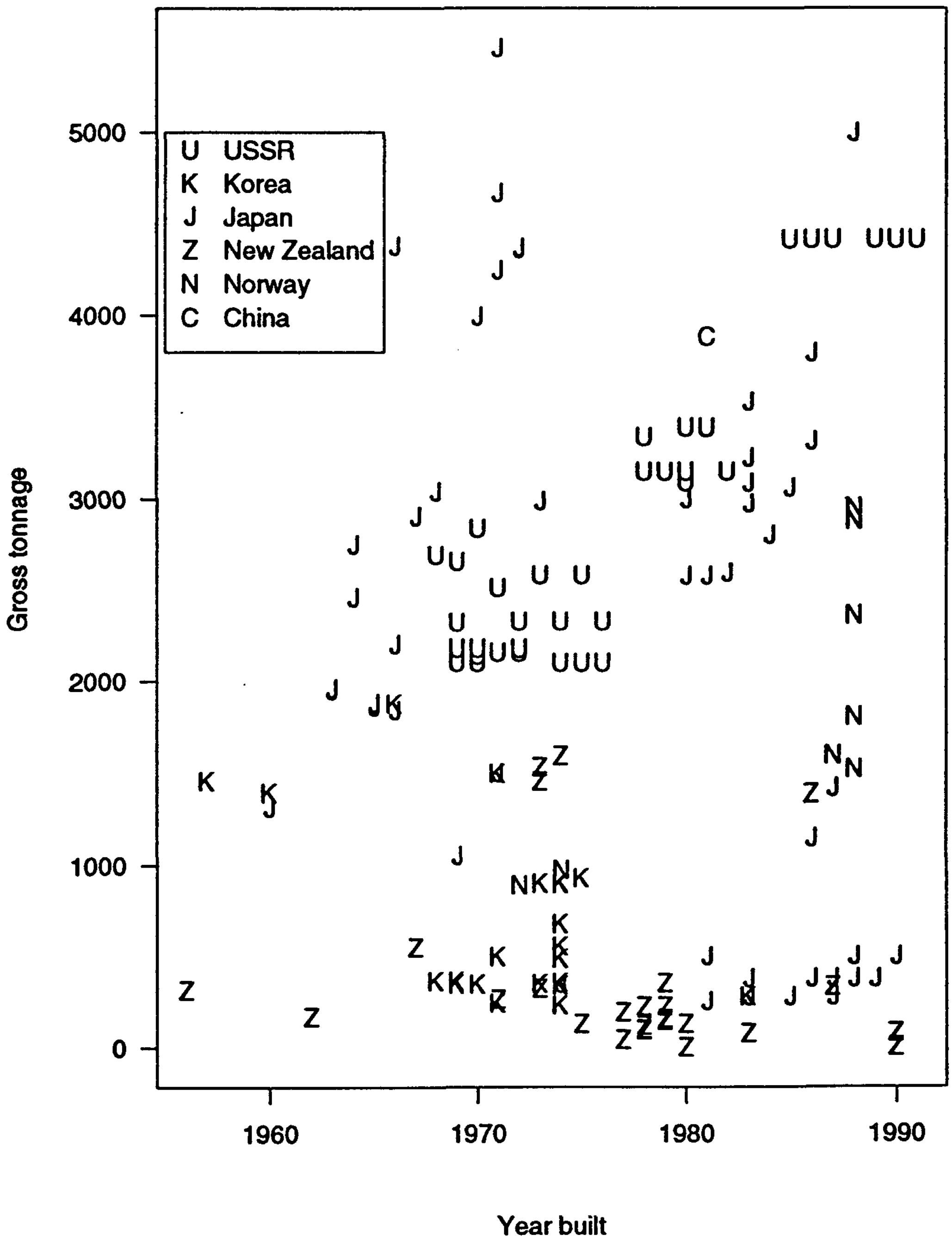


Figure 1: Plot of gross tonnage against year built for vessels from different nations

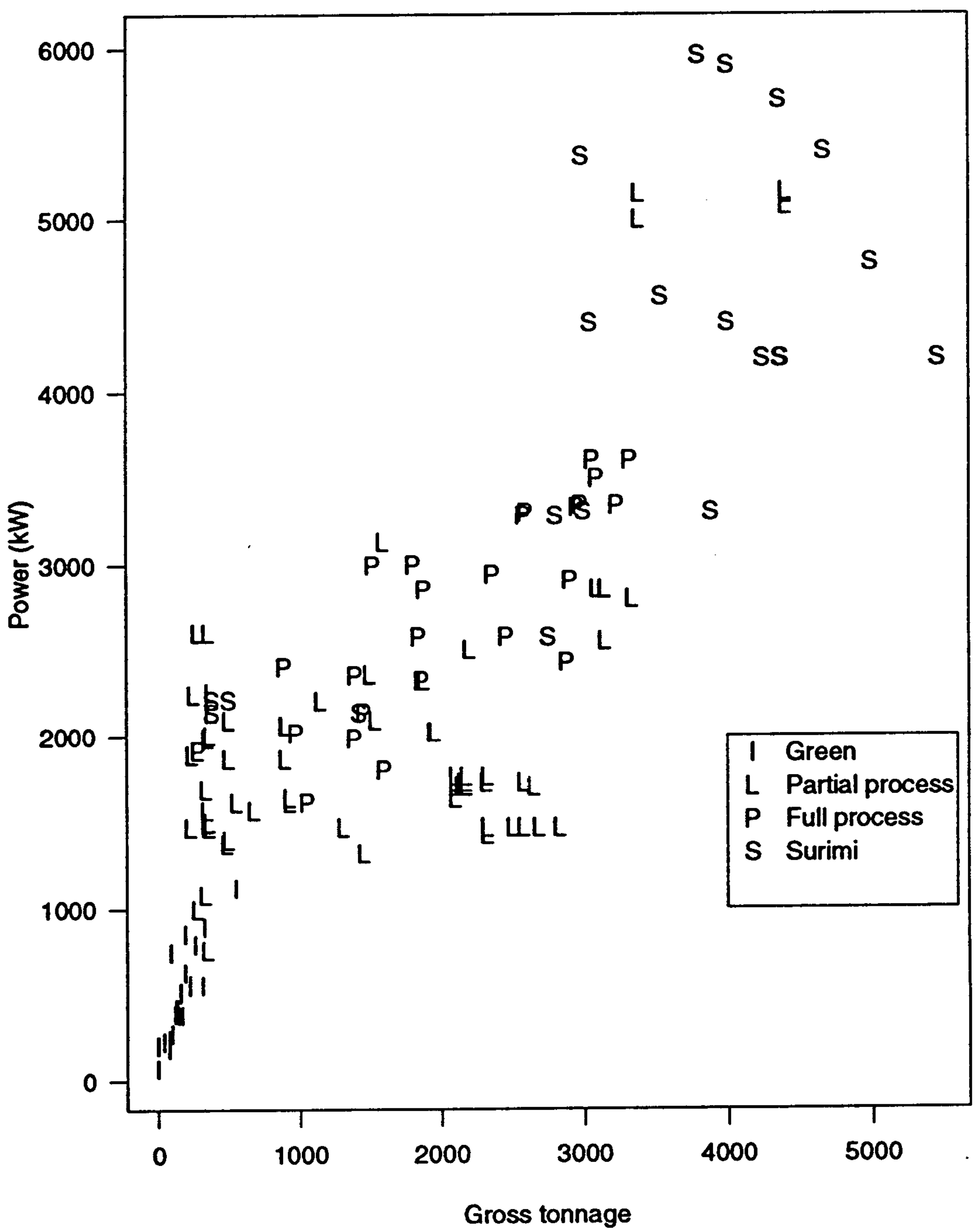


Figure 2: Plot of engine power (kW) against gross tonnage for vessels with different processing capabilities

company to which the vessel is registered. Each of these company representatives was interviewed to determine how they interpreted the forms. The result was that about 70% of the vessels were managed by companies that had continued to fill in the forms in the old style, and about 30% used the new definition.

This difference is likely to have the effect of adding about an hour of effort to the tows of all vessels in the companies using the new definition in the last two seasons. This could introduce a downwards bias to the last 2 years of CPUE. It was decided to eliminate all data from companies which used the new definition. This does not solve the problem entirely, as we do not know whether skippers in the other companies were following instructions.

Eliminating data from particular companies introduces the possibility of bias in the year indices. If the vessels managed by one company are on average larger, or have greater power, than those managed by another company, this is not a problem, as these effects should be standardised by the vessel variables (length, engine power, etc.) in the model. However, if the vessels managed by a company are on average significantly better at fishing (for their nationality, size, or engine power ) than those from another company, or if a company gives different instructions about fishing strategy or practices which lead to different average catch rates, then there may be a problem if the company effect is not evaluated as a variable.

To investigate such an effect, once the effects of vessel nation, tonnage, and breadth have been taken into account, a binary variable was created to distinguish between tows done by those companies which used the new definition in 1991 and 1992, and those companies that continued to use the old definition. Any systematic difference between vessels managed by different companies would be selected as a significant variable, but as described below, this was not the case.

## 3.6 Adjustment of zero catches

The index log(catch per n.mile) was chosen as the estimate of CPUE. Because the logarithm of zero is undefined, tows with zero catches of hoki cause problems for the analysis. Preliminary analyses showed that if all reported zero catches were replaced with some arbitrary low value, then the results were sensitive to the choice of this value.

Great effort was therefore made to eliminate errors in which the catch was incorrectly recorded as zero. Where possible, data were compared with equivalent information recorded by scientific observers. Nearly half of the tows recorded as zero tows on the trawl catch effort and processing returns that were also observed by scientific observers had non-zero reported catch on the observer's forms. The reported catches were between 0 and 171,680 kg of hoki: 6% of zero tows were eliminated using this information (47% of the tows which had observers on board).

Secondly, tows with bottom depth recorded as less than 275 m were eliminated. These records may have errors in the Target Species information rather than the hoki catch. A high proportion (30%) of these tows had zero catches, compared to the overall proportion of about 3%: 8% of zero tows were eliminated using this rule.

Thirdly, for every zero tow where it seemed that the same tow was repeated on the same day, the two tows were combined, in the sense that both CPUEs were replaced with a CPUE calculated from the sum of the catches divided by the sum of the distances covered. This is equivalent to the vessel having kept the net in the water for the entire time, and therefore not recording the zero catch as a separate tow. A repetition of the tow was defined as a tow by the same vessel on the same day within 50 m of the same bottom and net depth, in the same general area, with a headline height within 20% and with the speed within 0.2 knots. If there was more than one repetition of the tow, the one closest in time to the zero tow was used. This method eliminated about 25% of the zero tows.

However, there were still many zero tows which could not be eliminated logically. These were replaced with an arbitrary catch of 5 kg. These points can be seen as a small second peak to the left of the main peak in Figure 3, partially violating the assumption of normally distributed data. The results are only very slightly sensitive to the choice of this arbitrary catch.

An alternative approach is to use a poisson generalised linear model rather than the simple gaussian generalised linear model that the multiple regression model is equivalent to. Thus, instead of using the variable $\log(CPUE)$ (assumed to have constant variance), we would use the variable $\mu$ = CPUE, related to the explanatory variables by the link function $\log(\mu)$, and with variance proportional to $\mu$.

## 3.7 Comparability over the time series

Another possible problem is that the data from the FSU period (1983–88) were adjusted so that the total estimated catch for a day would equal the total processed catch for a day, while keeping the proportion in each tow the same as was originally estimated. This adjustment was not made to the later data. This change would be a problem if fishing masters systematically estimated catches as lower or higher than their true values, as there would be an inconsistency between the earlier and the later data. However, this does not seem to be a significant problem, as (in the 1992 fishery) the total estimated catch was 97.8% of the total processed (or landed) catch. Individual vessel proportions ranged from 82% to 110%, and the mean proportion was 97.6%.

# 4 The model

The model used was similar to the multiple regression model described by Vignaux(1992). Improved computing facilities meant that this year no sub-sampling was required. A summary of the variables used in the model is shown in Table 1, and a summary of the model itself is shown in Table 2.

In last year's analysis both log(catch per tow) and log(catch per n. mile) were considered, but it was found that since tow length had increased considerably over the period of the index, it was more likely that log(catch per n. mile) would index changes in abundance.

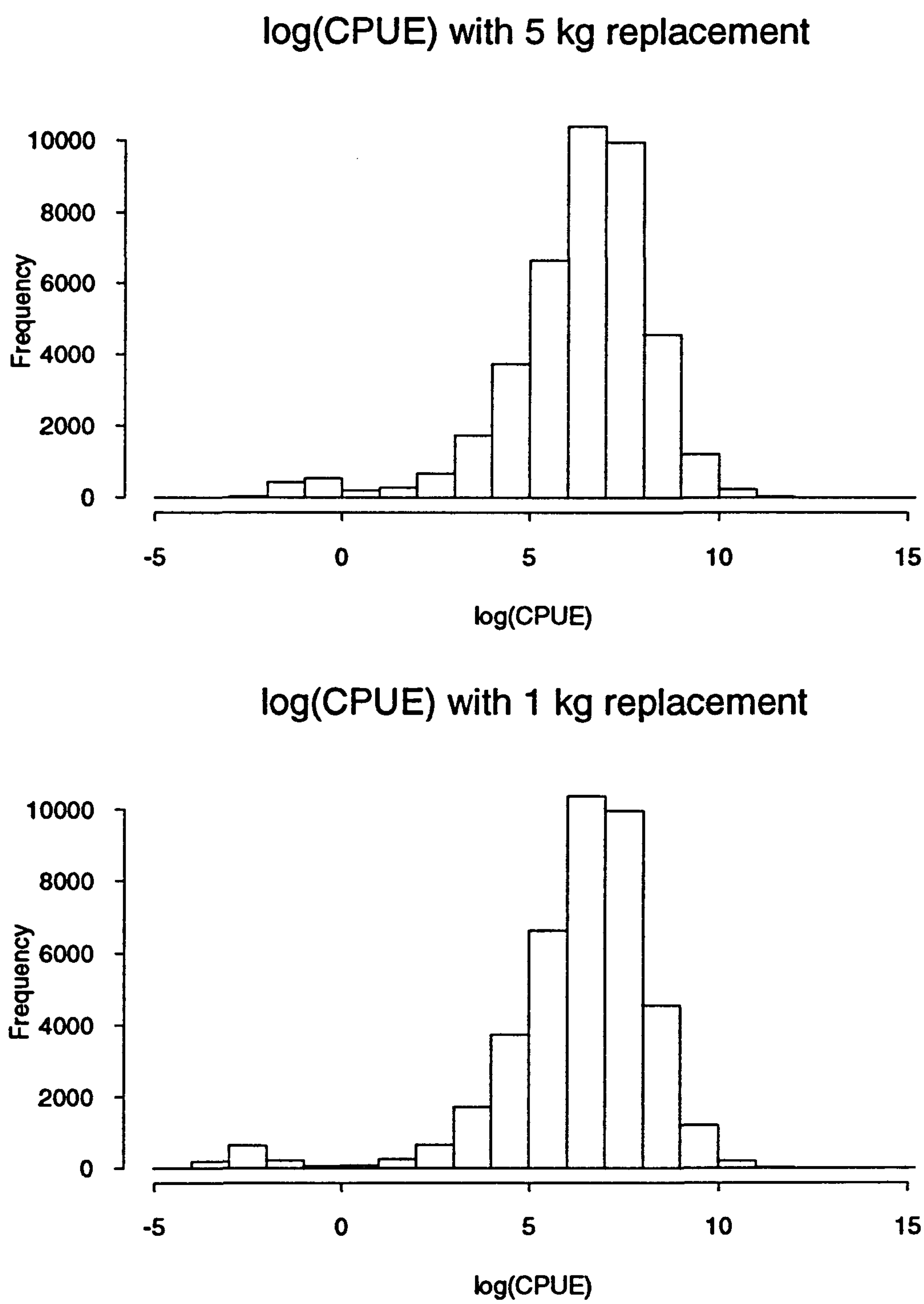


Figure 3: Histograms of log(CPUE) with the arbitrary zero replacement catch set to 5 kg or 1 kg

Table 1: Summary of variables in model. Cont. indicates a continuous variable, Cat. indicates a categorical variable with a given number of categories

| Variable | Type | Description |
|---|---|---|
| year | Cat 6 | Year that the tow occurred in |
| season | Cont | Day relative to the peak of season for that year |
| end time | Cont | Time in decimal hours of end of tow |
| latitude | Cont | Position in decimal degrees at start of tow |
| depth bottom | Cont | Depth in metres of bottom at start of tow |
| depth net | Cont | Depth in metres of groundrope of net at start of tow |
| net off bottom | Cont | Distance between groundrope and bottom at start of tow |
| headline height | Cont | Headline height of net at start of tow |
| vessel nation | Cat 6 | Country of origin of the vessel |
| vessel length | Cont | Length overall of the vessel, in metres |
| vessel tonnage | Cont | Gross tonnage of the vessel |
| vessel year built | Cont | Year the vessel was built |
| vessel power | Cont | Power of vessel engines in kilowatts |
| vessel breadth | Cont | Breadth of the vessel, in metres |
| vessel draught | Cont | Draught of the vessel, in metres |
| vessel processing | Cat 4 | Processing capability of vessel |
| vessel l*b*d | Cont | Vessel length * breadth * draught in cubic metres |
| vessel company | Cat 2 | Binary variable to identify companies dropped from analysis in 1991 |

The categorical variables (nation, year, and processing type) were included with the constraint that the sum of the regression coefficients should be zero. The continuous type variables were included up to a cubic power, but, if selected, only the terms which were significant were included in the model. In the last iteration, there was insufficient memory in the computer to evaluate the new variables to cubic power, so they were evaluated as linear variables.

The year variable was a categorical variable taking one of six values: 1987, 1988, 1989, 1990, 1991, 1992 since it was decided by the Hoki Working Group that the period up to 1987 had significantly different fishing patterns.

Season was a continuous variable. Figure 4 shows a smoothed line for the plot of CPUE against date for each year. Although the seasons all have a similar peaked shape, they are frequently offset from each other in time. So in this model, the seasonal variable is taken as the number of days before or after the peak of the season in that particular year. The peak of the season was found using smoothed lines as shown in Figure 4.

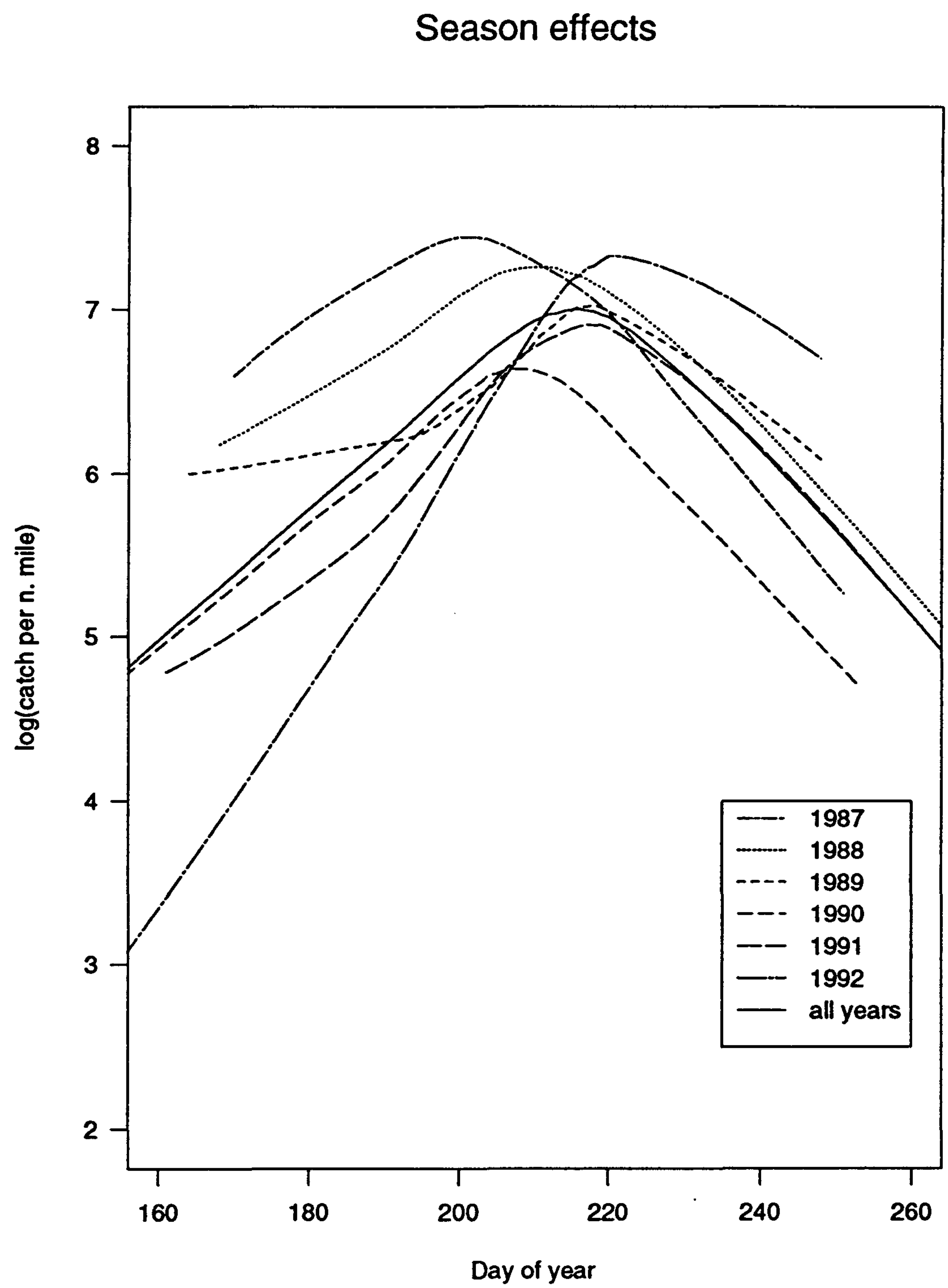


Figure 4: Plot of the seasonal effect on CPUE

# 5 Results

The first variable selected into the model was the seasonal effect, which entered the model as a quadratic variable, peaking during the season. This was followed by the nationality of vessel, then the year effect, followed by a quadratic variable for the end time of the tow. The headline of the net entered next, as a linear variable, followed by the gross tonnage and the breadth of the vessel (both cubic): 23% of the variability in log(CPUE) was explained by these seven variables. Coefficients and correlations of the variables are given in the appendix.

These variables were not identical to those selected in last year's model. Of the seven variables selected into the model last year, the five most important variables were selected again this year (season, time of day, year, vessel tonnage, and vessel breadth), while the two least important variables (vessel length and latitude) have been replaced. Headline height replaces latitude, and vessel nation replaces vessel length.

As discussed earlier, many of the variables are highly correlated, so that small changes in the dataset can change the selection variables without necessarily altering the year indices. There have also been improvements to the set of variables, and to the quality of the data which will influence their ability to explain catch rates. In particular, the generalisation of the seasonal effect seems to have increased its importance as an explanatory variable. Improvements in data quality in the vessel database can have a large influence, because an error in the parameters of a single vessel can be copied to many records, and may destroy any consistent relationship between catch rate and some measure of vessel size.

A possible index of abundance can be calculated (*see* Vignaux 1992) from the regression coefficients for the year effects (Table 4 and Figure 5). It is compared in Figure 5 with the index used for 1987–91 in the 1992 stock assessment. In that analysis the base year was 1989, but the index was recalculated relative to 1988 for comparison in this plot.

As described above, when this work was presented, computing limitations prevented completion of the eighth iteration of Table 2 with cubic powers of the variables, so the new variables were evaluated as linear variables in the last iteration. Table 3 shows that if it had been possible to evaluate the variables to cubic power, latitude would have been selected into the model as well, and the $R^2$ value would have been increased to 24.06%. Figure 8 shows the change in the index due to the introduction of the latitude variable.

Once the direct effects had been included, first order interactions were tested (for example between season and end-time). The only first order interaction to give a significant improvement was between season and nation, which improved the multiple $R^2$ value to 25%. The indices were not significantly altered by this variable (Figure 6).

The model was completely re-evaluated with the arbitrary zero replacement catch set at 1 kg instead of 5 kg (*see* Figure 3). This changed the selection of variables into the model slightly, with latitude being selected instead of headline height. Figure 7 shows the slight sensitivity in the indices to this factor of 5 change in the arbitrary zero replacement catch.

Table 2: Choice of variables in regression against log(catch per n. mile) for 1987–92 in order of importance. Numbers in the table are the multiple regression coefficient $R^2$. In the first iteration, log(cpue) was regressed against each of the variables in turn, to find the variable which explained the most variation in log(cpue) (had the highest multiple regression coefficient $R^2$). This variable was included in the model. In Iteration 2, log(cpue) was regressed against this model plus each of the other variables in turn to find the next most powerful variable. This process continued in a series of 8 iterations. Memory limitations meant that in the eighth iteration, the new variables were evaluated as linear variables only.

| | $R^2$ at iteration | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Season | 9.27 | | | | | | | |
| Vessel nation | 5.53 | 16.40 | | | | | | |
| End time | 1.97 | 11.15 | 18.59 | | | | | |
| Year | 2.40 | 11.26 | 18.34 | 20.55 | | | | |
| Headline height | 2.45 | 11.56 | 17.35 | 19.64 | 22.19 | | | |
| Vessel tonnage | 3.67 | 13.23 | 17.92 | 20.17 | 22.02 | 22.96 | | |
| Vessel breadth | 4.75 | 14.39 | 17.74 | 20.05 | 22.06 | 22.84 | 23.48 | |
| l*b*d | 5.31 | 15.29 | 17.80 | 20.08 | 21.99 | 22.82 | 23.32 | 23.62 |
| Vessel processing | 5.83 | 15.97 | 17.70 | 19.96 | 22.09 | 22.83 | 23.35 | 23.62 |
| Vessel draught | 4.52 | 14.77 | 17.45 | 19.69 | 21.60 | 22.65 | 23.15 | 23.61 |
| Year vessel built | 0.42 | 10.03 | 16.64 | 18.81 | 21.05 | 22.45 | 23.14 | 23.52 |
| Depth of bottom | 0.94 | 10.34 | 17.26 | 19.00 | 20.78 | 22.41 | 23.14 | 23.51 |
| Latitude | 0.61 | 9.96 | 17.27 | 19.49 | 21.17 | 22.71 | 23.46 | 23.50 |
| Net off bottom | 2.97 | 12.13 | 16.97 | 19.32 | 20.79 | 22.22 | 22.99 | 23.50 |
| Vessel length | 4.36 | 13.97 | 17.80 | 20.10 | 22.11 | 22.91 | 23.20 | 23.49 |
| Depth of net | 2.12 | 9.59 | 16.68 | 18.76 | 20.67 | 22.35 | 23.14 | 23.49 |
| Vessel company | 0.05 | 9.55 | 16.42 | 18.61 | 20.55 | 22.20 | 22.97 | 23.49 |
| vessel power | 4.59 | 14.47 | 17.77 | 20.04 | 22.02 | 22.84 | 23.05 | 23.48 |
| Polynomial | 2 | cat | 2 | cat | 2 | 3 | 3 | |
| Improvement | | 7.13 | 2.19 | 1.96 | 1.64 | 0.77 | 0.52 | |
| Degrees of freedom | 40520 | 40515 | 40513 | 40508 | 40506 | 40503 | 40500 | |

Table 3: Choice of variables in regression against log(catch per n. mile) for 1987–92 in order of importance. Numbers in the table are the multiple regression coefficient $R^2$. All continuous variables were evaluated to cubic power.

| | $R^2$ at iteration | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Season | 9.27 | | | | | | | | |
| Vessel nation | 5.53 | 16.40 | | | | | | | |
| End time | 1.97 | 11.15 | 18.59 | | | | | | |
| Year | 2.40 | 11.26 | 18.34 | 20.55 | | | | | |
| Headline height | 2.45 | 11.56 | 17.35 | 19.64 | 22.19 | | | | |
| Vessel tonnage | 3.67 | 13.23 | 17.92 | 20.17 | 22.02 | 22.96 | | | |
| Vessel breadth | 4.75 | 14.39 | 17.74 | 20.05 | 22.06 | 22.84 | 23.48 | | |
| Latitude | 0.61 | 9.96 | 17.27 | 19.49 | 21.17 | 22.71 | 23.46 | 24.06 | |
| Depth of bottom | 0.94 | 10.34 | 17.26 | 19.00 | 20.78 | 22.41 | 23.14 | 23.69 | 24.24 |
| Depth of net | 2.12 | 9.59 | 16.68 | 18.76 | 20.67 | 22.35 | 23.14 | 23.66 | 24.22 |
| l*b*d | 5.31 | 15.29 | 17.80 | 20.08 | 21.99 | 22.82 | 23.32 | 23.64 | 24.21 |
| Vessel processing | 5.83 | 15.97 | 17.70 | 19.96 | 22.09 | 22.83 | 23.35 | 23.62 | 24.17 |
| Vessel draught | 4.52 | 14.77 | 17.45 | 19.69 | 21.60 | 22.65 | 23.15 | 23.62 | 24.20 |
| Vessel power | 4.59 | 14.47 | 17.77 | 20.04 | 22.02 | 22.84 | 23.05 | 23.62 | 24.19 |
| Year vessel built | 0.42 | 10.03 | 16.64 | 18.81 | 21.05 | 22.45 | 23.14 | 23.56 | 24.15 |
| Vessel length | 4.36 | 13.97 | 17.80 | 20.10 | 22.11 | 22.91 | 23.20 | 23.53 | 24.12 |
| Net off bottom | 2.97 | 12.13 | 16.97 | 19.32 | 20.79 | 22.22 | 22.99 | 23.51 | 24.09 |
| Vessel company | 0.05 | 9.55 | 16.42 | 18.61 | 20.55 | 22.20 | 22.97 | 23.49 | 24.07 |
| Polynomial | 2 | cat | 2 | cat | 2 | 3 | 3 | 3 | |
| Improvement | | 7.13 | 2.19 | 1.96 | 1.64 | 0.77 | 0.52 | 0.58 | |
| Degrees of freedom | 40520 | 40515 | 40513 | 40508 | 40506 | 40503 | 40500 | 40497 | |

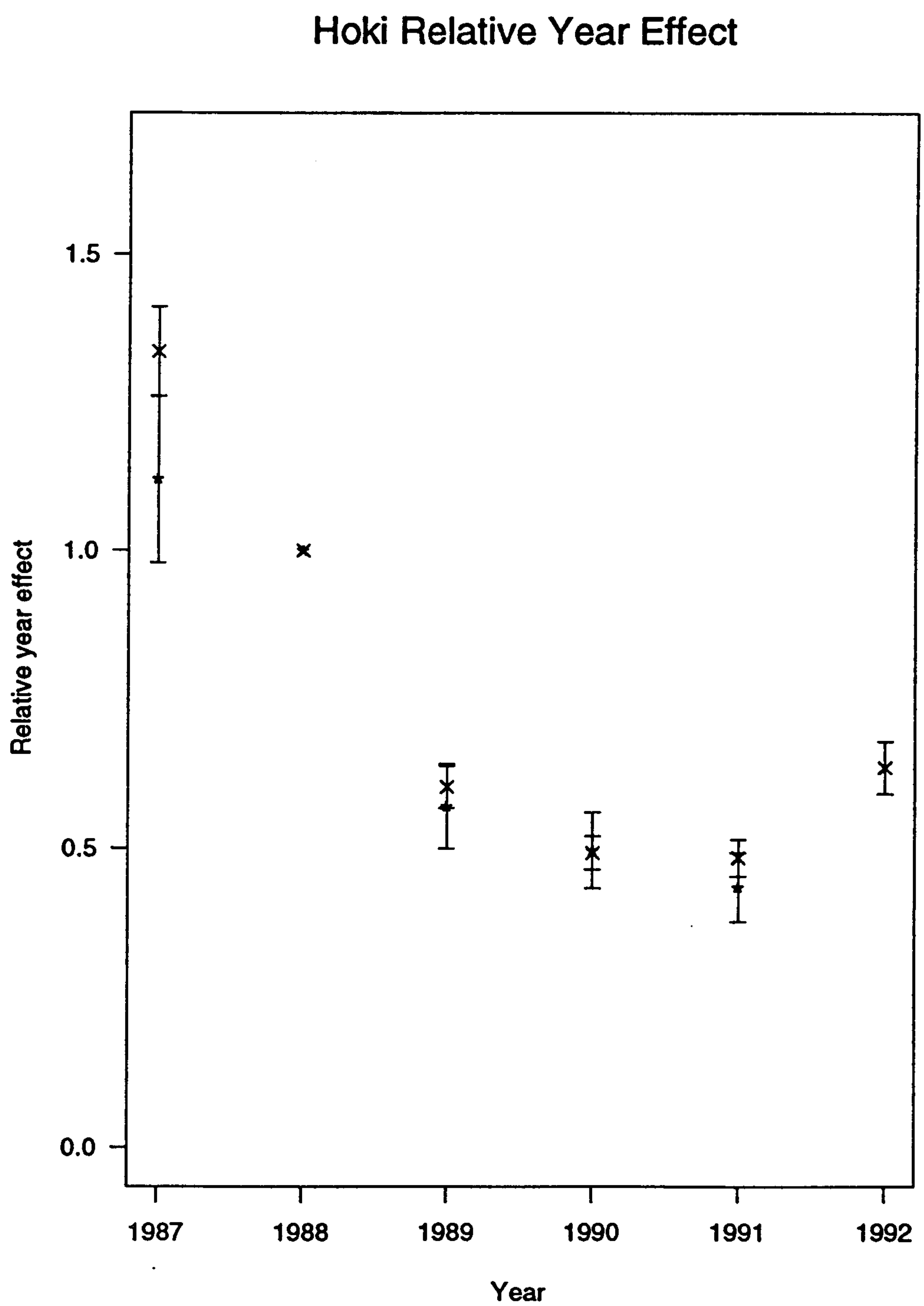


Figure 5: Relative year effects (×) estimated from log(catch per n. mile) compared with indices from last year's assessment (*). Lines are $\pm 2s_{\hat{A}_j}$

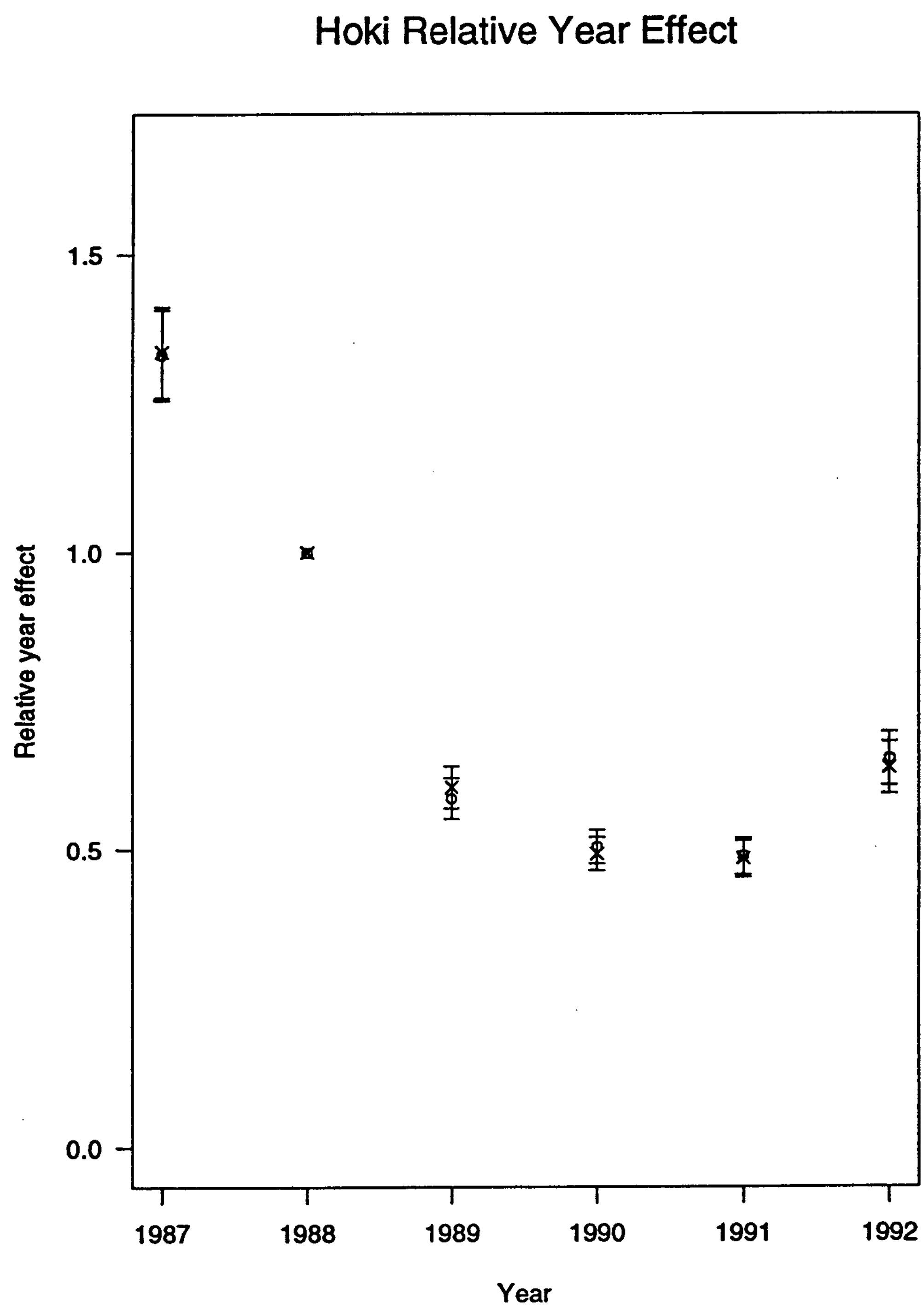


Figure 6: Relative year effects (×) estimated from log(catch per n. mile) compared with indices from model including interaction effect (○). Lines are $\pm 2s_{\hat{A}_j}$

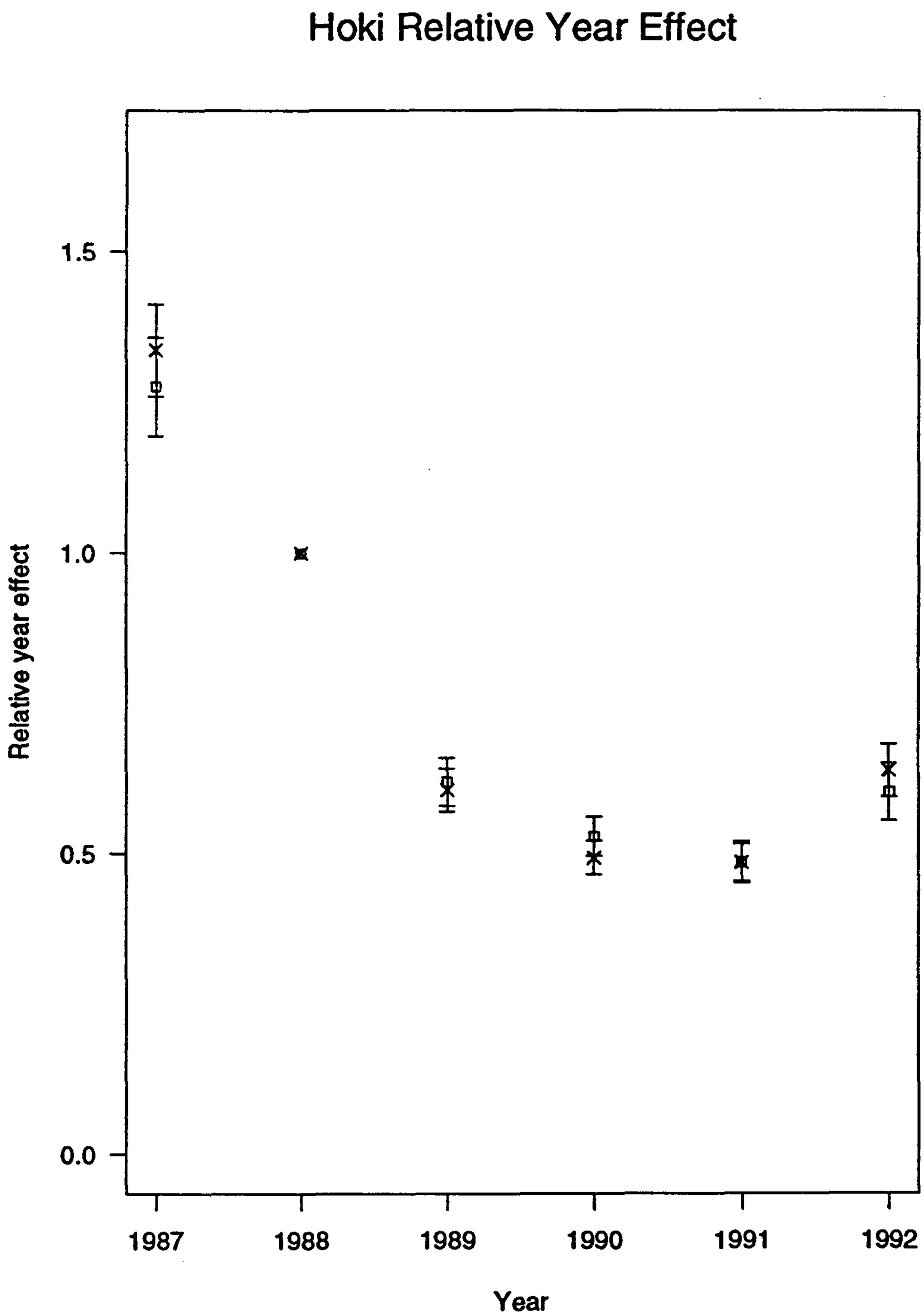


Figure 7: Relative year effects ($\times$) estimated from log(catch per n. mile) compared with indices with zero replacement catch of 1 kg ($\square$). Lines are $\pm 2 s_{\hat{A}_j}$

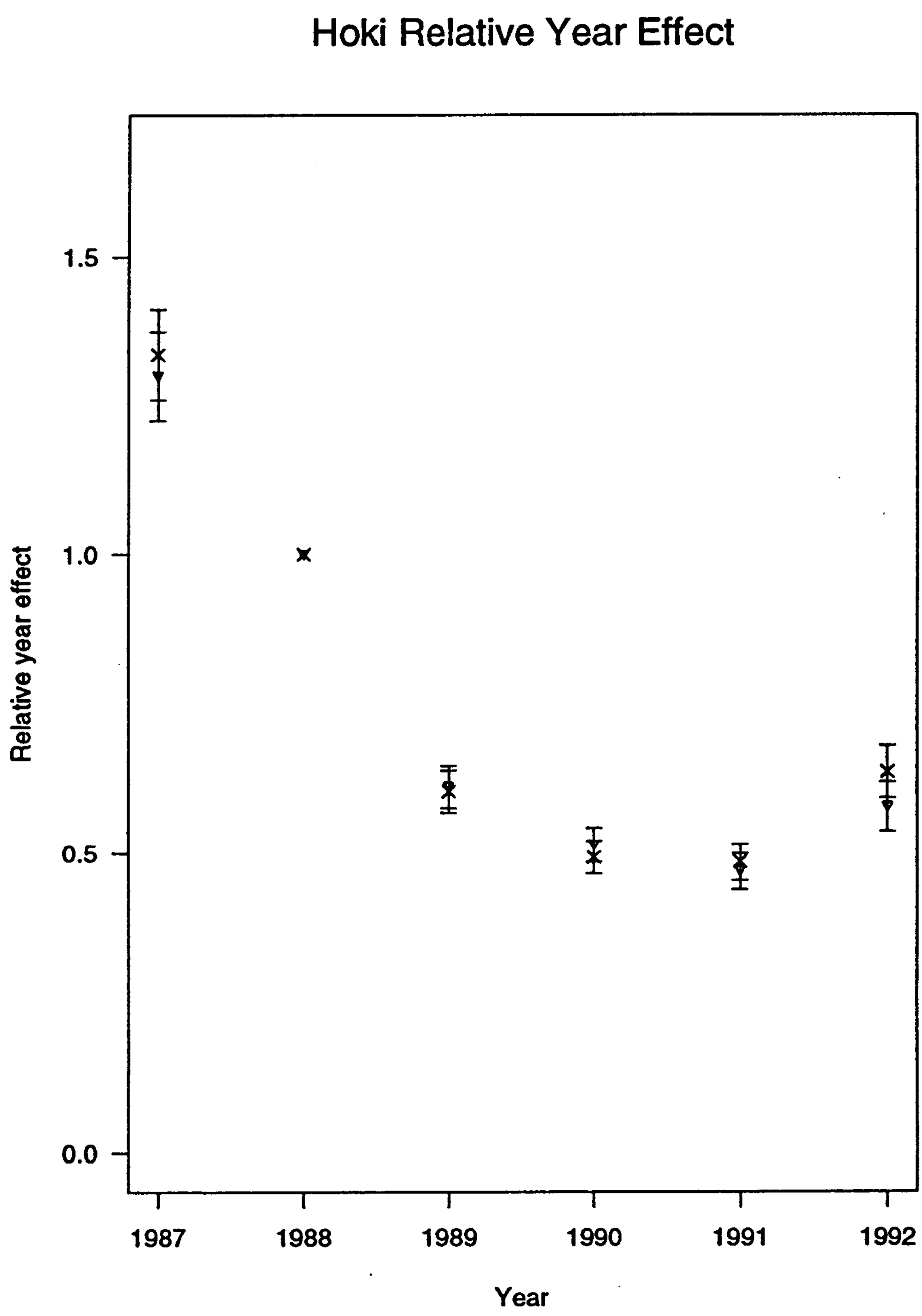


Figure 8: Relative year effects (×) estimated from log(catch per n. mile) compared with indices from model including latitude (▽). Lines are $\pm 2s_{\hat{A}_j}$

Table 4: Relative year effects ($\hat{A}_j$) for regression against log(catch per n. mile) for 1987–92. Reg. Coeff. are the regression coefficients of the year variable. Std. Err. are the standard errors of these coefficients in the regression, and Cov. is the covariance between the coefficients and the coefficient of the base year (1988). $\hat{A}_j$ is the relative year effect and $s_{\hat{A}_j}$ is the standard error of the year effect

| Year | No. tows | Reg. Coeff. | Std. Err. | Cov. | $\hat{A}_j$ | $s_{\hat{A}_j}$ |
|---|---|---|---|---|---|---|
| 1987 | 7110 | 0.639 | 0.021 | 3.47e-6 | 1.34 | 0.04 |
| 1988 | 8227 | 0.350 | 0.019 | 3.57e-4 | 1.00 | 0.00 |
| 1989 | 6525 | -0.155 | 0.020 | -5.96e-5 | 0.60 | 0.02 |
| 1990 | 8058 | -0.358 | 0.018 | -5.63e-5 | 0.49 | 0.01 |
| 1991 | 5920 | -0.375 | 0.021 | -1.04e-4 | 0.48 | 0.02 |
| 1992 | 4683 | -0.102 | 0.024 | -1.40e-4 | 0.64 | 0.02 |

# 6 Discussion

This analysis is an improvement over that described by Vignaux(1992) in a number of respects. First, the dataset used was the complete set of data available, rather than a subsample. Second, more error checking was done in the preparation of the data for analysis, as described above. Third, a greater range of variables was available for the regression (such as processing capability), although they were not used in the final model. Fourth, the effect of zero catches has been investigated, and shown to be insignificant (once all zeroes that can be logically removed, have been). Fifth, an error in the reporting of data (causing about a third of the vessels to report under a different definition of start and end time of tows) has been detected and (as far as possible) eliminated. Sixth, the seasonal effect has been generalised, so that an early or late season does not influence the year effect.

As can be seen from Figure 5 these changes have altered the indices (mainly the first point), but the indices for the years 1988–91 are similar to those presented in 1992. The elimination of data from companies using the new definition of start and end times has, as expected, raised the 1991 index relative to the 1990 index. The 1992 index is higher than the 1991 index, which is consistent with what was considered to be a late but very successful fishing season.(G Patchell pers. comm.)

However, to use a CPUE index as a stock index in a model such as the stock reduction model requires not only that changes in CPUE reflect changes in catch rates, but also that CPUE declines linearly with abundance with an intercept at zero. There are many reasons why CPUE may not be a linear index of abundance in this fishery. It may be that catch rates are determined not by the total abundance of fish, but by some feature of the density distribution of fish. For example, if the fishers can successfully detect and target aggregations of fish, catch rates may be determined by the density of the densest aggregations. If as abundance decreases, the fish reaggregate, the density of the densest aggregations may decrease much more slowly than the total abundance. If fishing is not targeted, catch rates may be determined by the total volume of aggregations which (if the fish reaggregate) may decrease faster than the total abundance. If because of changing availability of fish, fishing patterns change from well targeted fishing on aggregations to fishing on low density volumes of fish, or towards better targeting as fish finding equipment improves, then the picture is complicated further. Therefore, we must use any index of CPUE with caution.

# 7 Acknowledgments

I thank Elizabeth Bradford and members of the Hoki Working Group for useful discussions.

# 8 References

Lloyds 1990: Lloyds Register of Shipping. Lloyds, London 3 vols.

Vignaux 1992: Catch per unit of effort (CPUE) analysis of the hoki fishery. N.Z. Fisheries Assessment Research Document 92/14.

# 9 Appendix: Model parameters

For completeness, this appendix contains all the regression coefficients from the model and their correlation matrix. Where the variable is a continuous variable, the coefficients are for the linear, quadratic, and cubic terms in the variable. For categorical variables, the coefficients are for each of the categories labelled.

Table 1: Regression coefficients

| Variable | Model Parameters | | | | | |
|---|---|---|---|---|---|---|
| Season | 0.004 | -0.001 | | | | |
| Vessel nation | -0.553 | 0.602 | 0.175 | -0.170 | -0.150 | 0.095 |
| | CHINA | JAPAN | KOREA | NOR | NZ | USSR |
| End time | 0.002 | -7e-07 | | | | |
| Year | 0.639 | 0.350 | -0.155 | -0.358 | -0.375 | -0.102 |
| | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 |
| Headline height | 0.009 | 1.7e-05 | | | | |
| Vessel tonnage | -0.002 | 6e-07 | -5e-11 | | | |
| Vessel breadth | 0.200 | 0.023 | -0.001 | | | |

Table 2: Correlation matrix

| Variable | Int | yr1 | yr2 | yr3 | yr4 | yr5 | sea1 | sea2 | nat1 | nat2 | nat3 | nat4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| int | 1.00 | -0.07 | 0.03 | -0.00 | 0.02 | -0.00 | 0.01 | 0.02 | 0.03 | -0.07 | 0.24 | 0.08 |
| year1 | -0.07 | 1.00 | 0.00 | -0.21 | -0.20 | -0.28 | -0.40 | -0.13 | 0.03 | -0.09 | -0.08 | 0.08 |
| year2 | 0.03 | 0.00 | 1.00 | -0.16 | -0.16 | -0.26 | -0.14 | 0.03 | 0.03 | -0.09 | -0.09 | 0.06 |
| year3 | -0.00 | -0.21 | -0.16 | 1.00 | -0.14 | -0.19 | 0.12 | 0.03 | 0.00 | -0.05 | -0.01 | 0.03 |
| year4 | 0.02 | -0.20 | -0.16 | -0.14 | 1.00 | -0.15 | 0.06 | 0.05 | 0.02 | 0.01 | 0.04 | -0.12 |
| year5 | -0.00 | -0.28 | -0.26 | -0.19 | -0.15 | 1.00 | 0.14 | 0.03 | 0.03 | 0.02 | 0.02 | -0.04 |
| season1 | 0.01 | -0.40 | -0.14 | 0.12 | 0.06 | 0.14 | 1.00 | 0.12 | 0.00 | 0.07 | -0.01 | -0.05 |
| season2 | 0.02 | -0.13 | 0.03 | 0.03 | 0.05 | 0.03 | 0.12 | 1.00 | 0.00 | -0.06 | 0.03 | 0.03 |
| nation1 | 0.03 | 0.03 | 0.03 | 0.00 | 0.02 | 0.03 | 0.00 | 0.00 | 1.00 | -0.62 | -0.59 | -0.45 |
| nation2 | -0.07 | -0.09 | -0.09 | -0.05 | 0.01 | 0.02 | 0.07 | -0.06 | -0.62 | 1.00 | 0.22 | -0.01 |
| nation3 | 0.24 | -0.08 | -0.09 | -0.01 | 0.04 | 0.02 | -0.01 | 0.03 | -0.59 | 0.22 | 1.00 | -0.00 |
| nation4 | 0.08 | 0.08 | 0.06 | 0.03 | -0.12 | -0.04 | -0.05 | 0.03 | -0.45 | -0.01 | -0.00 | 1.00 |
| nation5 | -0.31 | -0.05 | -0.01 | 0.00 | 0.04 | -0.01 | -0.02 | -0.00 | -0.54 | 0.14 | 0.31 | -0.07 |
| end1 | -0.04 | 0.00 | -0.00 | -0.00 | 0.00 | -0.02 | -0.01 | -0.00 | -0.00 | 0.01 | -0.00 | 0.02 |
| end2 | 0.03 | 0.00 | -0.00 | -0.00 | -0.00 | 0.02 | 0.02 | 0.00 | 0.00 | -0.01 | 0.00 | -0.02 |
| head1 | 0.07 | -0.01 | -0.07 | 0.03 | 0.07 | 0.00 | -0.10 | 0.06 | -0.09 | -0.17 | 0.12 | 0.27 |
| head2 | -0.05 | 0.06 | 0.10 | -0.05 | -0.09 | -0.01 | 0.07 | -0.04 | 0.11 | 0.10 | -0.09 | -0.21 |
| ton1 | -0.10 | -0.12 | -0.15 | -0.05 | -0.00 | 0.15 | -0.00 | 0.04 | 0.08 | 0.21 | -0.10 | -0.04 |
| ton2 | -0.23 | 0.11 | 0.14 | 0.09 | 0.02 | -0.15 | 0.01 | -0.03 | -0.11 | -0.06 | -0.00 | 0.06 |
| ton3 | 0.38 | -0.10 | -0.11 | -0.09 | -0.01 | 0.13 | -0.01 | 0.02 | 0.10 | 0.01 | 0.06 | -0.06 |
| bre1 | -0.98 | 0.05 | -0.04 | 0.00 | -0.03 | 0.02 | -0.00 | -0.02 | 0.00 | 0.09 | -0.30 | -0.07 |
| bre2 | 0.95 | -0.03 | 0.06 | 0.00 | 0.02 | -0.04 | 0.00 | 0.01 | -0.02 | -0.12 | 0.32 | 0.06 |
| bre3 | -0.91 | 0.02 | -0.08 | -0.01 | -0.03 | 0.05 | -0.00 | -0.01 | 0.03 | 0.12 | -0.33 | -0.05 |

| Variable | nat5 | end1 | end2 | hea1 | hea2 | ton1 | ton2 | ton3 | bre1 | bre2 | bre3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| int | -0.31 | -0.04 | 0.03 | 0.07 | -0.05 | -0.10 | -0.23 | 0.38 | -0.98 | 0.95 | -0.91 |
| year1 | -0.05 | 0.00 | 0.00 | -0.01 | 0.06 | -0.12 | 0.11 | -0.10 | 0.05 | -0.03 | 0.02 |
| year2 | -0.01 | -0.00 | -0.00 | -0.07 | 0.10 | -0.15 | 0.14 | -0.11 | -0.04 | 0.06 | -0.08 |
| year3 | 0.00 | -0.00 | -0.00 | 0.03 | -0.05 | -0.05 | 0.09 | -0.09 | 0.00 | 0.00 | -0.01 |
| year4 | 0.04 | 0.00 | -0.00 | 0.07 | -0.09 | -0.00 | 0.02 | -0.01 | -0.03 | 0.02 | -0.03 |
| year5 | -0.01 | -0.02 | 0.02 | 0.00 | -0.01 | 0.15 | -0.15 | 0.13 | 0.02 | -0.04 | 0.05 |
| season1 | -0.02 | -0.01 | 0.02 | -0.10 | 0.07 | -0.00 | 0.01 | -0.01 | -0.00 | 0.00 | -0.00 |
| season2 | -0.00 | -0.00 | 0.00 | 0.06 | -0.04 | 0.04 | -0.03 | 0.02 | -0.02 | 0.01 | -0.01 |
| nation1 | -0.54 | -0.00 | 0.00 | -0.09 | 0.11 | 0.08 | -0.11 | 0.10 | 0.00 | -0.02 | 0.03 |
| nation2 | 0.14 | 0.01 | -0.01 | -0.17 | 0.10 | 0.21 | -0.06 | 0.01 | 0.09 | -0.12 | 0.12 |
| nation3 | 0.31 | -0.00 | 0.00 | 0.12 | -0.09 | -0.10 | -0.00 | 0.06 | -0.30 | 0.32 | -0.33 |
| nation4 | -0.07 | 0.02 | -0.02 | 0.27 | -0.21 | -0.04 | 0.06 | -0.06 | -0.07 | 0.06 | -0.05 |
| nation5 | 1.00 | -0.02 | 0.01 | 0.05 | -0.05 | -0.16 | 0.19 | -0.17 | 0.22 | -0.15 | 0.10 |
| end1 | -0.02 | 1.00 | -0.96 | -0.01 | 0.02 | 0.00 | -0.01 | 0.01 | 0.00 | -0.00 | -0.00 |
| end2 | 0.01 | -0.96 | 1.00 | 0.00 | -0.01 | -0.00 | 0.00 | -0.00 | -0.00 | -0.00 | 0.00 |
| head1 | 0.05 | -0.01 | 0.00 | 1.00 | -0.93 | -0.02 | 0.06 | -0.06 | -0.08 | 0.05 | -0.04 |
| head2 | -0.05 | 0.02 | -0.01 | -0.93 | 1.00 | 0.12 | -0.15 | 0.13 | 0.06 | -0.05 | 0.04 |
| ton1 | -0.16 | 0.00 | -0.00 | -0.02 | 0.12 | 1.00 | -0.84 | 0.67 | 0.19 | -0.29 | 0.34 |
| ton2 | 0.19 | -0.01 | 0.00 | 0.06 | -0.15 | -0.84 | 1.00 | -0.95 | 0.17 | -0.08 | 0.02 |
| ton3 | -0.17 | 0.01 | -0.00 | -0.06 | 0.13 | 0.67 | -0.95 | 1.00 | -0.34 | 0.28 | -0.23 |
| bre1 | 0.22 | 0.00 | -0.00 | -0.08 | 0.06 | 0.19 | 0.17 | -0.34 | 1.00 | -0.98 | 0.96 |
| bre2 | -0.15 | -0.00 | -0.00 | 0.05 | -0.05 | -0.29 | -0.08 | 0.28 | -0.98 | 1.00 | -0.99 |
| bre3 | 0.10 | -0.00 | 0.00 | -0.04 | 0.04 | 0.34 | 0.02 | -0.23 | 0.96 | -0.99 | 1.00 |