



ISSN 1175-1584

MINISTRY OF FISHERIES

Te Tautiaki i nga tini a Tangaroa

**An evaluation of catchability assumptions in
New Zealand stock assessments**

R. I. C. C. Francis

R. J. Hurst

J. A. Renwick

An evaluation of catchability assumptions in
New Zealand stock assessments

R. I. C. C. Francis
R. J. Hurst
J. A. Renwick

NIWA
PO Box 14 901
Wellington

**Published by Ministry of Fisheries
Wellington
2001**

ISSN 1175-1584

**©
Ministry of Fisheries
2001**

Citation: Francis, R.I.C.C.; Hurst, R.J.; Renwick, J.A. (2001).
An evaluation of catchability assumptions in New Zealand stock assessments.
New Zealand Fisheries Assessment Report 2001/1. 37 p.

This series continues the informal
New Zealand Fisheries Assessment Research Document series
which ceased at the end of 1999.

EXECUTIVE SUMMARY

Francis, R.I.C.C.; Hurst, R.J.; Renwick, J.A. (2001). An evaluation of catchability assumptions in New Zealand stock assessments.

New Zealand Fisheries Assessment Report 2001/1. 37 p.

"Catchability" is used in two contexts in fisheries science. With regard to commercial catch per unit effort (CPUE) it means the proportion of a fish stock that is expected to be caught by a standard unit of fishing effort. For random trawl surveys it has a similar meaning, but can also be thought of as the product of three components: vulnerability, vertical availability, and areal availability.

All available data on New Zealand catchability were assembled. This comprised data from 48 stock assessments and 17 series of random trawl surveys. These data were examined for evidence supporting or rejecting several assumptions that underpin the use of catchability in stock assessments.

Coefficients of variation (c.v.s) assigned to CPUE indices were found, more often than not, to be larger than they should be, whereas those for trawl survey indices were usually too small. There is clear evidence of annual variation in catchability in the trawl survey data. In about one year in six catchability was extreme (either very high or very low) for many species.

Several instances were found of catchabilities that were implausible, either because they were too low, or because there was too much variation in catchability amongst stocks of the same species. There is also some evidence of positive autocorrelation in CPUE catchabilities but the interpretation of this is difficult.

The following recommendations follow from this work.

1. Default c.v.s for CPUE indices in stock assessments should be lower than is currently common in New Zealand: 0.15–0.2, rather than 0.3–0.35.
2. Default c.v.s for trawl survey indices should be calculated by "adding" an annual variability c.v. of 0.2 to the survey observation error c.v. (reminder: c.v.s "add" as squares, so 0.2 "added" to 0.3 is $\sqrt{0.2^2 + 0.3^2} = 0.36$).
3. Estimated trawl survey catchabilities that are very different from 1 should be treated as an error indication in stock assessments. Either the trawl survey series is not an appropriate method of tracking this particular stock or the survey indices are strongly in conflict with other model inputs.
4. Large within-species variation in catchability should also be treated as an error indication (but catchabilities should be compared only if they are of the same type).

This report fulfils Objective 1 of Ministry of Fisheries project SAM1999/01: *To investigate variability in catchability of trawl surveys* (although this objective refers only to trawl survey catchability, the *Rationale* section of the Ministry's tender document makes it clear that both types should be considered).

1. INTRODUCTION

Catchability is a key parameter that is estimated in many fish stock assessments (Arreguín-Sánchez 1996). However, it is usually thought of as a “nuisance” parameter: one that is not of intrinsic interest but which needs to be estimated so that other quantities, which are of interest (e.g., biomass), can be estimated. For this reason estimates of catchability are often not reported.

The use of catchability in stock assessment models is based on several assumptions. Much has been written about the most important of these — that catchability does not change with biomass — but the validity of other assumptions has received much less attention. It is these other assumptions that are the focus of this report.

We start by defining catchability, listing the assumptions to be considered, and reviewing the relevant literature. Two types of data sets (from stock assessments and trawl surveys), containing most of the information about New Zealand catchabilities, are assembled and apparent violations of the assumptions in them are listed and discussed. Finally, we make recommendations for future stock assessments, in particular with regard to the coefficients of variation (c.v.s) that should be used for biomass indices.

This report fulfils Objective 1 of Ministry of Fisheries project SAM1999/01: *To investigate variability in catchability of trawl surveys* (although this objective mentions only trawl survey catchability, the *Rationale* section of the Ministry’s tender document makes it clear that CPUE (catch per unit effort) catchability should also be considered.

2. DEFINITIONS, ASSUMPTIONS, AND LITERATURE REVIEW

2.1 Defining catchability

“Catchability” is used in several slightly different ways in the fisheries literature. In stock assessment models using biomass indices, catchability is a model parameter (conventionally denoted by q) which is usually defined by the equation

$$I_i = qB_i\epsilon_i \quad (1a)$$

or sometimes by

$$\log(I_i) = \log(qB_i)\epsilon_i \quad (1b)$$

where I_i is the index in year i , B_i is the corresponding biomass, and the error term, ϵ_i , is a random variable with expectation 1 and c.v. (coefficient of variation) c_i .

The interpretation of q in (1) depends on the source of the I_i . If they are derived from CPUE, q may be interpreted as the proportion of the population biomass that is caught by one unit of effort. Usually the CPUE is standardised, so the unit of effort is a standard one (e.g., if nationality and area are factors in the CPUE standardisation, then the standard unit of effort will be that for a vessel from the reference nation in the reference area). If the I_i are from a trawl survey series the interpretation of q is slightly different. Here, it is the product of the survey area and the proportion of the biomass that is caught per unit area swept.

Another common use of catchability is in either

$$C_i/f_i = q_iB_i \quad (2a)$$

or

$$F_i = qf_i \quad (2b)$$

where C_i , f_i , and F_i are, respectively, the catch, effort, and fishing mortality in year i . The F_i will usually be calculated from a stock assessment model, and the B_i may be either from the same model (which may or may not use the effort data) or some other more direct source (e.g., tagging or an acoustic survey).

Equations (1) and (2) clearly provide different definitions of catchability. However, the differences are subtle. We can think of the q_i of (2) as being the same as $q\epsilon_i$ in (1). Also, since C_i/f_i is just catch per unit effort (and so is analogous to I_i), and F_i is approximately equal to C_i/B_i , we can see that these equations are broadly equivalent.

Thus, although there are several different ways of defining catchability, all of these may be thought of as being the proportion of the biomass of a population that is caught by some specified unit of fishing effort. The effort may be from targetted commercial fishing or random research surveys. In the latter case it is conventional to multiply the proportion by the survey area. It is worth commenting that the unit of effort is much more closely specified in research surveys than in commercial fishing. Even if all commercial fishers were using essentially identical gear there would be many ways in which their effective effort could differ substantially because of subtle differences in the ways in which they rig and deploy the gear.

2.2 Components of trawl survey catchability

Trawl survey catchability may also be interpreted in terms of three quantities that may be thought of as components of catchability: vulnerability, v , vertical availability, u_v , and areal availability, u_a . The calculation of biomass indices from NIWA trawl surveys is based on the following equation from Francis (1989)

$$\text{Biomass} = [1/(vu_vu_a)] \sum_i a_i C_i$$

where a_i is the area, and C_i is the mean catch rate (weight per unit area swept), for the i th stratum in the survey, and the area swept at each station is calculated as a product of the distance towed and the width between the trawl doors (Vignaux 1994). If the catchability components were well known, trawl survey biomass estimates could be treated as absolute, rather than relative. This would make q equal to 1. In practice, very little is known about these components and they are usually set equal to their default value of 1. In this case, $q = vu_vu_a$. If an estimate of one component, say \hat{u}_v , was used, together with the default values of the other components then $q = vu_vu_a/\hat{u}_v$.

This interpretation restricts the range of plausible values for a trawl survey q . Because all three catchability components are defined as proportions (see Francis 1989) their product should be less than (or equal to) 1. [It is technically possible for v to exceed 1 (if, for instance, fish that are initially above the headline, and thus unavailable to the net, are herded downwards) but it is most unlikely that the product vu_v would be greater than 1; u_a cannot exceed 1.] Thus, if the default values of the catchability components have been used we would expect q to be less than 1. Also, very small values of q are implausible for any species that is assessed using trawl survey biomasses. Although there are species that are not well caught by trawls (e.g., because they are fast-swimming, high above the bottom, or burrow in the substrate) and thus have very low values of v and/or u_v , such species are not, for that reason, assessed using trawl survey biomasses. Similarly, a very low areal availability (implying that most of a fish stock is outside the survey area) would rule out the use of trawl surveys in assessing a stock.

There is also a limit to how much we would expect values of q for the same species to vary between surveys. For a given fishing vessel and trawl net, the components v and u_v are determined by fish behaviour (e.g., swimming speed, typical height above the bottom, reaction to an approaching net). This means that if the same vessel and gear are used in surveys in different areas we would expect the product vu_v not to vary very much for the same species (except, perhaps, between spawning and non-spawning periods, when there may be substantial behavioural differences). If different vessels, or gear, are used, we might get larger differences in vu_v .

2.3 Catchability assumptions

The main (and often implicit) assumption about catchability in Equation (1) is that it does not vary with abundance. That is, the proportion of a population that is caught by a unit of fishing effort does not change systematically as the population size increases or decreases. The literature relating to this assumption is reviewed briefly in Section 2.4 but the assumption is not otherwise addressed in this report.

A second assumption is that the c.v.s we use for the biomass indices in our assessments (the c_i) accurately represent the uncertainty in the indices. This uncertainty will be a mixture of observation error (which we can usually estimate directly) and annual variation in catchability (which we cannot estimate directly). (By “annual” variation we mean between-year, not within-year). For CPUE indices, the usual New Zealand practice is to allow for annual variation in catchability by making the c_i s much greater than the estimated observation error (Francis 1999b). For trawl indices, the method of setting the c_i s varies (see Section 3.1). In Section 4.1 we test the validity of this assumption for New Zealand stock assessments. That is, we ask the question “are our c_i s the right size?”.

A third assumption is that the ε_i are (statistically) independent. This assumption may not appear to concern catchability. However, if annual variation in catchability is caused by changes in an environmental variable, like temperature, which is often autocorrelated on an annual time scale (Zheng & Basher 1999) then this will imply autocorrelation in the ε_i . We look for such correlations in Section 4.2.

2.3.1 Variation within- and between-populations

It is important not to confuse within- and between-population variation in catchability. The annual variation discussed in the preceding section is “within-population”. That is, the catchability may vary from year to year, but it is always the catchability of the same population.

Some authors have investigated what we call “between-population” variation in catchability. That is, they have compared the catchabilities of different populations (or, sometimes, different segments of the same population). For example, Borgström (1992) compared catchabilities of brown trout in four Norwegian lakes, and Richards & Schnute (1986) compared catchabilities of quillback rockfish at various locations in the Strait of Georgia, British Columbia. The first example clearly concerns distinct populations. It is unclear whether this is true of the second example, but, because the measures of abundance used by Richards & Schnute were local (covering small areas), it seems clear that their various catchabilities applied to different groups of fish.

In this report we are interested in between-population variation when we look (in Section 4.3) at how much trawl survey catchability varies between different stocks of the same species. Too much variation would imply a failure of our assumptions about the components of catchability (Section 2.2).

2.4 Catchability vs abundance

Paloheimo & Dickie (1964) questioned the assumption that catchability is independent of abundance, giving theoretical reasons to expect an inverse relationship: as the abundance of a stock decreases, they expected its catchability to increase (and the area it occupies to decrease). This paper sparked a great deal of interest and many authors presented data confirming the Paloheimo & Dickie hypothesis. Most of the initial confirmations were for marine pelagic schooling species [Atlantic menhaden (Schaaf 1975), Pacific sardine (MacCall 1976), various herring stocks (Jakobsson 1980, Pope 1980, Ulltang 1980, Winters & Wheeler 1985), Atlantic mackerel (Pope 1980), and Peruvian anchoveta (Csirke 1989)], but similar results were soon found for marine demersal fish [various cod stocks (Pope & Garrod 1975, Houghton & Flatman 1981, Angelsen & Olsen 1987, Rose & Leggett 1991), Georges Bank haddock (Crecco & Overholz 1990), and Bering Sea pollock (Quinn & Collie 1990)] and for river and lake fisheries [Pacific salmon (Peterman & Steer 1981), lake whitefish (Henderson et al. 1983), and shad (Crecco & Savoy 1985)].

There are good reasons to be cautious about accepting this apparently large body of evidence in favour of an inverse relationship between catchability and abundance. The main reason is that CPUE was not standardised in most studies. This means that an increase in fishing power over time (a common pattern) would tend to make catchability increase over time. If, over the same time period, abundance has fluctuated about a stable mean or was trending downwards (which, unfortunately, is more common than an upwards trend) then there would appear to be an inverse relationship between catchability and abundance. Trends in the areal pattern of fishing can also cause an apparent increase in catchability in non-standardised CPUE (Cook & Armstrong 1985). A second reason for caution was pointed out by Shardlow & Hilborn (1985). If catchability is independent of abundance, but the latter is estimated with error, then a plot of catchability against abundance will appear to show an inverse relationship. This technical problem can be overcome (Peterman et al. 1985) but not all the authors cited above have clearly done so. A final reason for caution is that, although there *are* studies showing no relationship between catchability and abundance (Pinhorn 1988), such “negative” results are often considered unworthy of publication.

When considering catchability from non-research fishing (either commercial or recreational) we should be aware of the considerable ingenuity and tenacity of fishers faced with a declining resource. There is very strong pressure on fishers to maintain catch rates as abundance declines. This is a good reason to expect an *apparent* increase in catchability as abundance falls. If it were possible to correct for this change in fishing power (i.e., to adjust the unit of effort as fishing practice changes) then the true catchability may not be changing.

Might there be some fish stocks for which catchability and abundance are *positively* correlated? We have found only one (partial) example. Swain et al. (1994) examined catchability (from both commercial and research fishing) by age class for Atlantic cod and found a positive correlation with abundance for some age classes, and a negative one for others. Hilborn & Walters (1992, pp 190–191) describe circumstances which might lead to a positive correlation, but give only two examples, both from South Australia. However, both fisheries started in small areas of high fish density and subsequently spread to much larger areas of low density. Thus, the population covered by estimated catchability has changed over time. In other words, these are examples of between-population changes. This is also true of the study by Shardlow (1993).

Other studies of catchability vs abundance that used within-population comparisons, and thus are not relevant to the present question are those by Bannerot & Austin (1983), Richards & Schnute (1986), Richards (1987), and Borgstrøm (1992).

Dunin et al. (2000a) assembled biomass indices for more than 40 fish stocks to investigate the relationship between CPUE and abundance. They concluded that although the relationship may range from “hyperstability” (catchability increasing with decreasing abundance) to “hyperdepletion”

(catchability decreasing with decreasing abundance) “a proportional relationship may be an adequate approximation for some stocks”.

In summary, we have found many examples where the catchability of a population appears to increase as its abundance declines and only one partial example of the contrary trend. As we have said above, there are good reasons to be suspicious of at least some of these examples. However, it is hard not to be struck by the extreme imbalance in the evidence.

Various models have been proposed in which catchability varies with abundance. The earliest, and simplest is $q = aB^b$ (Ulltang 1980), where a and b are constants. Other examples are given by Bannerot & Austin (1983), Richards & Schnute (1986, equation 1.57), Hilborn & Walters (1992, p. 144), Tanaka (1997), and Quinn & Deriso (1999, pp 28–29).

2.5 Environmental effects on catchability

It is a commonplace observation amongst fishers, both commercial and recreational, that there are times when fishing is good, and other times when it is bad. Sometimes this is caused by fluctuations in abundance. However, it is also widely accepted that environmental changes can cause a change in the catchability of fish.

Many different environmental factors have been found (or suspected) to affect catchability. In this report we are concerned with year-to-year changes in catchability and so are not interested in shorter-term variations (e.g., the day-night, lunar, and tidal cycles). Amongst longer-period environmental influences the most often cited is (water) temperature (Bagge & Munch-Petersen 1979, Morrissy & Caputi 1981, Cayre & Roy 1986, Danzmann et al. 1991, Yamakawa et al. 1994, Macpherson & Gordo 1992, Skog et al. 1994, Perry & Smith 1994, Brodziak & Hendrickson 1999). Other factors include water clarity (Nielsen 1983, Morrissy & Caputi 1981, Buijse et al. 1992), thermocline depth (Fonteneau 1991), oxygen concentration (Bagge & Munch-Petersen 1979, Bailey et al. 1985), salinity (Perry & Smith 1994), El Niño (Alvial Muñoz 1988), intensity of ocean waves (Yamakawa et al. 1994), rainfall (Evans et al. 1997), and the position of the ice margin (Fréchet 1990). Catchability may also be affected by changes in the abundance of co-occurring species (Polovina 1986).

3. DATA

Two types of New Zealand data were examined: from stock assessments and random trawl surveys.

3.1 Assessment data

We gathered data from all recent stock assessments that used biomass indices from either trawl surveys or CPUE. One data set was constructed for each separate series of biomass indices (so an assessment using two different series provided two data sets). Each data set consisted of:

- the biomass indices input to the assessment,
- the years associated with these indices,
- the c.v.(s) assumed for these indices,
- a description of the assumed error distribution type,
- the model estimates of (absolute) biomass that correspond to each biomass index, and
- the model estimate of catchability, q , for the indices.

For each stock the latest available assessment (usually carried out in 2000) was used. Data sets with fewer than four annual indices were discarded. For assessments carried out using MIAEL estimation (Cordue 1998) the model estimates are those obtained at the least-squares stage of estimation.

Amongst these data sets there were three different error-distribution assumptions; these determine how standardised residuals are calculated (Table 1).

A total of 48 such data sets were constructed (30 with CPUE indices and 18 with trawl survey indices), ranging in length from 4 years to 40 years (Table 2). The two rock lobster assessments (CRA.NSN and CRA.45) used a time step of 6 months; all other assessments used a 1 year time step.

We refer to the c.v.s in Table 2 as “assumed”, rather than “estimated”, because we can only estimate one component of these c.v.s: that due to observation error. By setting a value for one of these c.v.s we are implicitly making an assumption about the other component: annual variation in catchability. We did not solicit information on how these c.v.s were set. However, the most common way is based on a subjective assessment of the reliability of the associated biomass indices: the less reliable the indices are judged to be, the higher the assumed c.v.. This is presumably why almost half of the series with CPUE indices (13 out of 30) have assumed c.v.s of 0.35, and many of the series with survey indices (13 of 18) have c.v.s of 0.25. For at least one survey data sets (ORH NEChat) the c.v.s are equal to the estimated observation error (i.e., the c.v.s calculated from the trawl survey data) which implies that annual variability in catchability is assumed to be negligible. This was also true for the four data sets with oreo CPUE (SSO3A.SOV, SSO3A.DOM, BOE3A.SOV, BOE3A.DOM) but the observation error c.v.s for these series were calculated using a jackknife procedure (Doonan et al. 1995) rather than taking the conventional regression statistic (which typically produces very low c.v.s).

Table 1: Three alternative error-distribution assumptions for biomass indices in stock assessments and the associated form of the standardised residuals. Notation: I_i is the i th biomass index and B_i is the corresponding model estimate of (absolute) biomass; for assumption $lnorm$, $\sigma_i^2 = \log(c_i^2) + 1$.

Label	Description	Standardised residual
<i>norm</i>	I_i is normally distributed with mean qB_i and c.v. c_i	$\frac{1}{c_i} \left(\frac{I_i}{qB_i} - 1 \right)$
<i>lnorm</i>	I_i is lognormally distributed with mean qB_i and c.v. c_i	$\frac{1}{\sigma_i} \left[\log \left(\frac{I_i}{qB_i} \right) + 0.5\sigma_i^2 \right]$
<i>lognorm</i>	$\log(I_i)$ is normally distributed with mean $\log(qB_i)$ and s.d. c_i	$\frac{1}{c_i} \log \left(\frac{I_i}{qB_i} \right)$

Table 2: Details of the 48 New Zealand assessment data sets used in this report. Series labels start with a three-digit species code followed by a number, or other code, which defines a particular stock.

Label	Type ¹	Error distribution ²	Number of years	Assumed c.v.(s)	Reference
ORH MEC	trawl CPUE	norm	15	0.3	Francis & Field (2000)
ORH7A	trawl CPUE	norm	17	0.3	Field & Francis (unpubl. results)
ORH EC	trawl CPUE	norm	6	0.3	Anderson (2000)
ORH NEChat	Box ORH	norm	9	0.15-0.67	Francis (1999a)
SBW Camp	trawl CPUE	log(norm)	13	0.5	Hanchet (2000)
HOK	trawl CPUE	lnorm	12	0.35	Cordue (2000)
GUR1E.1	BoP	lnorm	5	0.25	Hanchet et al. (2000)
GUR1E.2	HG Kaharoa	lnorm	10	0.25	Hanchet et al. (2000)
GUR1W	FMA9	lnorm	6	0.25	Hanchet et al. (2000)
GUR2	ECNI	lnorm	4	0.25	Hanchet et al. (2000)
JDO1E.1	trawl CPUE	lnorm	8	0.35	Horn et al. (1999)
JDO1E.2	Danish seine CPUE	lnorm	8	0.35	Horn et al. (1999)
JDO1E.3	BoP	lnorm	5	0.25	Horn et al. (1999)
JDO1E.4	HG Kaharoa	lnorm	11	0.25	Horn et al. (1999)
JDO1W.1	trawl CPUE	lnorm	8	0.35	Horn et al. (1999)
JDO1W.2	FMA9	lnorm	6	0.25	Horn et al. (1999)
JDO2E	ECNI	lnorm	4	0.25	Horn et al. (1999)
JDO2W	FMA8	norm	4	0.25	Horn et al. (1999)
LIN34.1	longline CPUE	norm	9	0.35	Horn et al. (2000)
LIN34.2	Chat HOK	norm	8	0.25	Horn et al. (2000)
LIN56.1	longline CPUE	lnorm	8	0.35	Horn et al. (2000)
LIN56.2	S. Plat	lnorm	4	0.25	Horn et al. (2000)
LIN6B	longline CPUE	norm	7	0.35	Horn et al. (2000)
LIN7	longline CPUE	norm	9	0.35	Horn et al. (2000)
STA34	Chat HOK	norm	7	0.25	Paul et al. (1999)
STA4	Chat HOK	lnorm	7	0.25	Paul et al. (1999)
WAR3.1	Southland	lnorm	4	0.4	Horn & Bagley (1999)
WAR3.2	Shinkai Maru ³	lnorm	4	0.6	Horn & Bagley (1999)
TRE7.1	trawl CPUE	log(norm)	8	0.22	Hanchet (1999)
TRE7.2	trawl CPUE	log(norm)	20	0.32	Hanchet (1999)
PAU5B	CPUE	lnorm	16	0.5	Breen et al. (2000)
PAU7	CPUE	lnorm	17	0.5	Andrew et al. (unpubl. results)
CRA.NSN	CPUE	lnorm	20 ⁴	0.3	Breen et al. (unpubl. results)
CRA.45	CPUE	lnorm	20 ⁴	0.3	Breen et al. (unpubl. results)
HAK4	trawl CPUE	lnorm	7	0.35	Dunn et al. (2000b)
SNA.HGBoP	longline CPUE	log(norm)	9	0.35	Gilbert et al. (2000)
SKI1	trawl CPUE	lnorm	10	0.35	Hurst et al. (2000)
SKI2	trawl CPUE	lnorm	10	0.35	Hurst et al. (2000)
SNA8.pair	pair trawl CPUE	log(norm)	18	0-0.3	Davies & McKenzie (unpubl. results)
SNA8.sngl	single trawl CPUE	log(norm)	10	0.022-0.028	Davies & McKenzie (unpubl. results)
SNA8.age2	FMA9	log(norm)	5	0.06-0.096	Davies (unpubl. results)
SNA8.age3	FMA9	log(norm)	5	0.03-0.11	Davies (unpubl. results)
SNA.ENLD	longline CPUE	log(norm)	9	0.35	Gilbert et al. (2000)
SSO3A.SOV	trawl CPUE	lnorm	6	0.61	Doonan et al. (1999a)
SSO3A.DOM	trawl CPUE	lnorm	12	0.51	Doonan et al. (1999a)
BOE3A.SOV	trawl CPUE	lnorm	9	0.24	Doonan et al. (1999b)
BOE3A.DOM	trawl CPUE	lnorm	12	0.36	Doonan et al. (1999b)
ORH7B	trawl CPUE	norm	12	0.3	Annala et al. (1998)

¹ Specifies the source of the biomass indices used in the assessment: either CPUE (sometimes with the fishing method — e.g., trawl or long-line — specified) or one of the trawl survey series of Table 3 below

² As defined in Table 1

³ Southland surveys between 1981 and 1986; not included in Table 3 because they covered different months

⁴ 6-month time step in this assessment

3.2 Trawl survey data

Data from all stratified random trawl surveys in the research database *trawl* were considered. The surveys were grouped into series, each of which contained surveys covering (approximately) the same area at about the same time of year using the same (or similar) vessel(s) and gear. Where it seemed appropriate a series was split into two, by area. Series with fewer than four surveys were rejected. This left 17 series, with between 4 and 11 surveys per series (Table 3). This includes two instances of geographical splitting: the series covering the west coast of the North Island (Morrison & Stevenson in press) became our series FMA8 and FMA9; and the one for the west coast of the South Island (Stevenson 1998) which became WCSI and TBGB. The latter split was particularly important because of the different fish communities in the two subareas. In one series, HG Ikatere, the survey design was systematic (i.e., the same stations were occupied each year), rather than stratified random. The method of analysing these data (which involves defining strata, and treating the catches as if they came from randomly generated stations) is detailed in Appendix 2.

For each trawl survey series a list of “suitable” species was generated by listing all species caught in the series and then excluding species deemed to be “unsuitable” for any of the following reasons:

- species caught in only a small percentage of tows
- species caught in small quantities (low mean catch per tow)
- species not well caught by the net because they are too small, or too large or too hard down, or too high in the water column
- where species identification was poor, or inconsistent over time
- species whose range was poorly covered by the surveys (e.g., those occurring mostly on rough ground, or mostly in shallower/deeper water than covered by the series).

The idea was to include as many species as possible for each series. In considering a particular species in a specific trawl survey series, the following question was useful: “If this were a valuable commercial species would it be appropriate to use this series of trawl surveys to generate biomass indices to put into a stock assessment?”. An answer of “yes” (or even “maybe”) was a good reason to include this species. The c.v.s of biomass estimates were not considered in making this decision. For each series the list of suitable species was compiled by people with an intimate knowledge of that series and the associated species. No attempt was made to derive consistent objective criteria (e.g., exclude all species that occur in fewer than 30% of tows) for all series. The exclusion of a species from one series was no barrier to its inclusion with another. The number of acceptable species in a series varied between 4 and 25 (Table 3). Species lists, and other series details, are given in Appendix 1.

Relative biomass estimates, and c.v.s, were calculated for each suitable species in each survey. Some of these estimates differ from published values because either (a) stratum areas and gear widths have been revised, or (b) estimates were restricted to a subset of the survey strata to make them comparable across all surveys in a series. In all series but two, vulnerability, areal availability, and vertical availability were set to 1. In the 1989, 1990, and 1991 surveys in the Box ORH series vertical availability was assumed to differ from 1 (with values ranging from 1.38 to 1.48) at a total of nine stations. In the S. Chat. series the vulnerability was set equal to the net wingtip width divided by the trawl door spread (so the effective fishing width was assumed to be the wingtip width, rather than the door spread).

There was a wide range of estimated c.v.s. Even when the c.v.s for each species in a series is averaged over all surveys these averaged values spanned an order of magnitude, from 0.07 to 0.70, (Figure 1).

Table 3: Description of the 17 data sets associated with series of New Zealand trawl surveys. Target species sometimes varied from year to year; those listed below were targets (i.e., listed in the *trawl* database as main species) in most, if not all, years. The order of the series in this Table is approximately clockwise around the coast of New Zealand, starting at the Hauraki Gulf.

Series	Area	Main months	Target species	Depth range (m)	First & last years	No. of surveys	No. of species
HG Ikaterere	Hauraki Gulf	9,10	SNA	5–70	1976,1980	5	10
HG Kaharoa	Hauraki Gulf	10,11	SNA	10–150	1984,1997	11	5
BoP	Bay of Plenty	2,3	SNA	10–150	1983,1999	6	4
Scampi	Bay of Plenty	1	SCI	200–600	1993,1998	4	23
ECNI	east coast N.I.	2,3	TAR ¹	20–400	1993,1996	4	22
ECSI summ.	east coast S.I.	12,1	ELE,GUR,RCO	10–400	1997,2000 ²	4	25
ECSI wint.	east coast S.I.	5,6	RCO	30–400	1991,1996	5	15
Chat HOK	Chatham Rise	1	HOK,HAK,LIN	200–800	1992,2000	9	18
Box ORH	NE Chatham Rise	7	ORH	750–1250	1984,1994	9	25
S. Chat	South Chatham Rise	10,11	BOE,ORH,SSO	600–1500	1991,1995	4	16
Southland	Southland	2,3	BAR,SKI ³	50–600	1993,1996	4	15
S. Plat.	Southern Plateau	4,5	HOK,HAK,LIN	300–800	1992,1998	4	13
WCSI	west coast S.I.	3,4	GUR,RCO,STA	20–400	1992,2000	5	18
TBGB	Tasman Bay/Golden Bay	3,4	GUR,RCO,STA	20–60	1992,2000	5	5
Chall	Challenger Plateau	7	ORH	800–1200	1984,1990	5	25
FMA8	SW coast N.I.	10,11,12	SNA	10–100	1989,1996	4	5
FMA9	NW coast N.I.	10,11,12	SNA	10–100	1986,1996	6	5

¹ Originally also SNA, TRE, & SKI

² years in which each survey ended

³ also BGZ, WAR, STA, HAP, LIN, & SCH

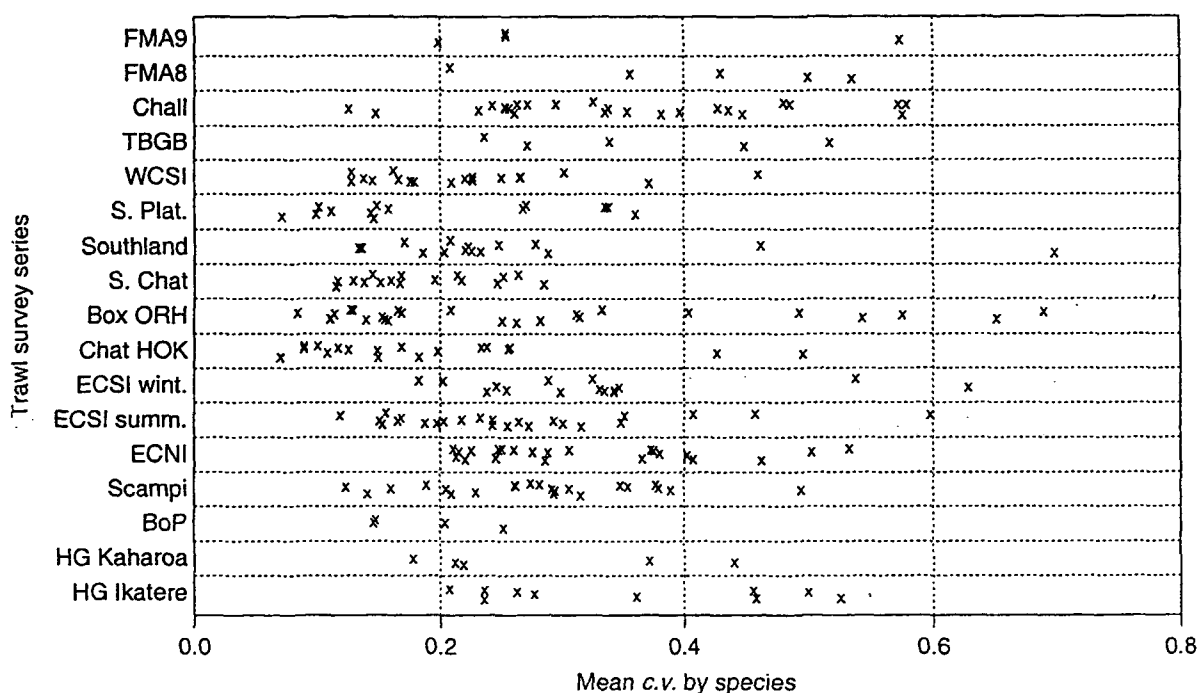


Figure 1: Mean c.v.s, by species, of biomass indices in the trawl survey data of Table 3. Each plotted point relates to one species in a trawl survey series, and indicates the mean of all estimated c.v.s for that species in that series.

4. FAILURES OF ASSUMPTIONS

We investigated three types of assumption failure in our data sets. These are discussed in Sections 4.1 to 4.3. In Section 4.4 we seek evidence of annual variability in trawl survey catchability.

4.1 Are our assessment c.v.s the right size?

If the stock assessment assumptions are correct, and the assumed c.v.s are accurate, the standardised residuals (calculated following Table 1) for each of the assessment data sets in Table 2 should have variance close to 1. A variance much greater than 1 suggests that the assumed c.v.s are too small, and vice versa.

We constructed a residual statistic, κ , that was designed to indicate whether the stock assessment c.v.s (the c_i) were too small or too large in each data set. A positive (or negative) value of κ suggests that the residuals were too large (too small), and thus c.v.s were too small (too large), and a value greater than 2 (less than -2) is statistically significant. The statistic is based on the median absolute standardised residual (*m.a.s.r.*), rather than the residual variance, because the latter is not very robust (it is easily inflated by outliers). We defined

$$\kappa = \begin{cases} 2(\mu - \mu_{0.5})/(\mu_{0.5} - \mu_{0.025}) & \text{if } \mu < \mu_{0.5} \\ 2(\mu - \mu_{0.5})/(\mu_{0.975} - \mu_{0.5}) & \text{if } \mu > \mu_{0.5} \end{cases}$$

where μ is the *m.a.s.r.* from the assessment data and μ_q is the q th quantile of the sampling distribution of μ . The standardised residuals were assumed to follow a Student's t -distribution with $n-2$ degrees of freedom, where n is the number of years in the data set. (We assumed $n-2$ degrees of freedom because, in an assessment with only a single series of relative biomass indices, only two parameters can be estimated, e.g., B_0 and q . When there are many data inputs there may be many more than two parameters estimated.) For each value of n the μ_q were estimated by: simulating 1000 data sets of size n from a t -distribution with $n-2$ degrees of freedom; calculating the median absolute value for each simulated data set; and taking the q th quantile of this set of 1000 medians.

Results were different for the two types of assessment data (Figure 2). For those with CPUE indices, there is a tendency for c.v.s to be too large: κ is negative for 21 of the 30 such data sets (this is significantly more than half, $P = 0.02$), and is less than -2 for 9 of them. In contrast, κ is positive for 13 out of the 18 data sets with survey indices (again, significantly more than half, $P = 0.02$), and is greater than 2 for 2 of them.

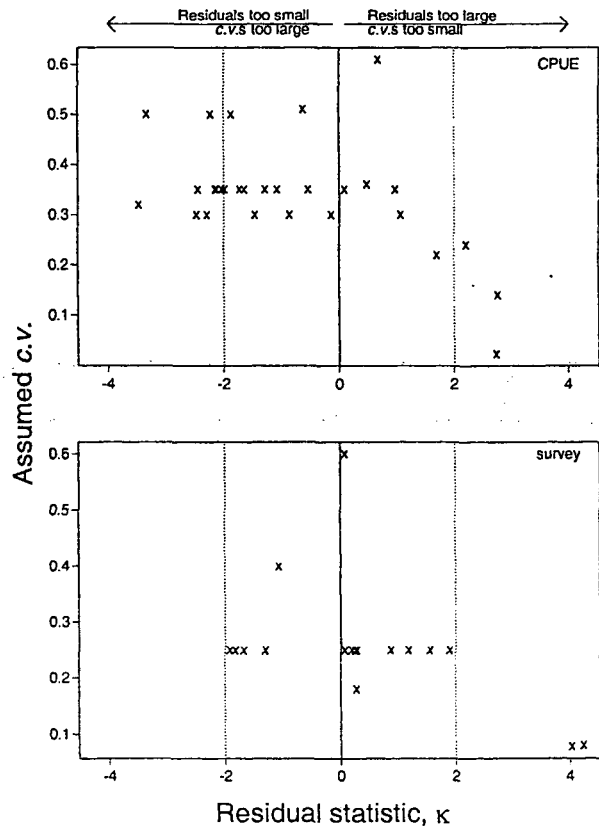


Figure 2: The residual statistic, κ , plotted against assumed c.v. for each of the assessment data sets of Table 2: those with CPUE indices in the upper panel, those with survey indices in the lower panel.

Median c.v.s for data sets for which the c.v.s were found to be significantly too large ranged from 0.3 to 0.5; where c.v.s were significantly too small the median c.v.s were between 0.02 and 0.24 (Table 4).

Table 4: Details of the assessment data sets of Table 2 for which the assumed c.v.s are significantly too large ($\kappa < -2$) or too small ($\kappa > 2$). “Type” refers to the type of biomass index: CPUE or trawl survey.

	Series label	Type	κ	Median c.v.
c.v.s too large	JDO1E.2	CPUE	-2.4	0.35
	LIN6B	CPUE	-2.1	0.35
	LIN7	CPUE	-2	0.35
	TRE7.2	CPUE	-3.5	0.32
	PAU5B	CPUE	-3.3	0.5
	PAU7	CPUE	-2.2	0.5
	CRA.NSN	CPUE	-2.3	0.3
	CRA.45	CPUE	-2.5	0.3
	SNA.ENLD	CPUE	-2.2	0.35
c.v.s too small	SNA8.pair	CPUE	2.8	0.14
	SNA8.sngl	CPUE	2.7	0.02
	SNA8.age2	survey	4	0.08
	SNA8.age3	survey	4.2	0.08
	BOE3A.SOV	CPUE	2.2	0.24

4.2 Is catchability autocorrelated?

Most of the assessment data sets are so short that a test for significant autocorrelation in the residuals of an individual data set would have little power. Instead we devised a simple test that could be applied to the entire data collection and thus have more power. The test statistic was calculated by counting, in each data set, the number of times that residuals from consecutive years had the same sign, and summing these counts over all data sets. This statistic is designed to test autocorrelations with lag 1 year. Thus, much of the assessment data from trawl surveys was ignored because adjacent residuals are often not from consecutive years. Where the assessment time step was 6 months (in the two rock lobster data sets, CRA.NSN and CRA.45) the i th and $(i+2)$ th residual were compared so that the time lag between them was 1 year.

The resulting statistic is 246, which is higher than its expected value (under the null hypothesis of no autocorrelation) of 202.5 ($= 0.5 \sum n_i$, where n_i is the number of pairs of consecutive years in the i th data set). This suggests positive autocorrelation. To test whether this correlation is statistically significant we generated 1000 simulated data sets having the same structure (i.e., the same n_i s) as the real data but with the residuals drawn at random from the set $\{-1, 1\}$, and calculated the same test statistic for each. All of these 1000 simulated statistics differed from the expected value by more than the actual statistic did so the positive autocorrelation in the assessment data is significant with $P < 0.001$. When the test was repeated for two data types separately we found a significant positive correlation for the data sets with CPUE indices (test statistic = 226, expected value = 176, $P < 0.001$) but a non-significant negative correlation for those with survey indices (test statistic = 20, expected value = 26.5, $P = 0.07$).

Our method for calculating statistical significance here is flawed. It assumes, for example, that all patterns of residual sign are equally likely, whereas in practice the sequence $+, +, +, +$ will be less likely than $+, -, +, -$. However, the relative likelihood of the various sequences depends on the biomass trajectory. A more robust procedure would involve rerunning the assessments with simulated uncorrelated data. However, it is not practical to do this for so many assessments.

4.3 Are our trawl survey catchabilities too extreme?

In Section 2.2 we said that trawl survey catchabilities would be implausible if they were too extreme: either very low or very high, or varying too widely for a single species. It is difficult to be very specific about what constitutes “too extreme” or “too widely” in this context. Nevertheless, the 18 estimates plotted in Figure 3 provide clear grounds for concern.

The overall range is very wide, covering more than two orders of magnitude, from 0.0035 to 1.6. Although the theoretical maximum value for q is 1, the two values that exceed this may not be of concern if we allow for estimation error. However, the lowest values *are* of concern. If these are accurate then it would seem inappropriate to use trawl surveys to assess these stocks. For example, a q that is less than 0.01 means that more than 99% of the stock is, in some sense, not available to the trawl survey — either because it is outside the survey area (low areal availability), does not encounter the trawl (low vertical availability), or easily avoids it (low vulnerability).

For two species the range of values is implausibly wide: for GUR the four estimates vary by a factor of 79 (0.0039 to 0.31); for JDO the factor is 49 (0.0035 to 0.17) (the next widest range is for STA, a factor of just 2.8).

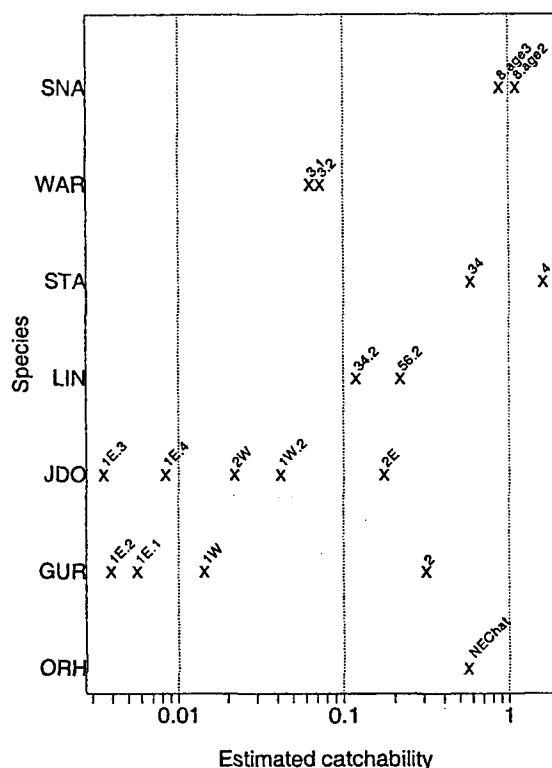


Figure 3: Estimates of trawl survey catchability by species. The labels above each point identify the corresponding data set in Table 2 (e.g., the two STA points labelled “34” and “4” correspond to data sets STA34 and STA4).

4.4 Years of extreme catchability

It is not possible to extract precise or definitive information about fluctuations in catchability from trawl survey data alone. However, for some trawl survey series, there appears to be evidence of “extreme” years — years in which catchability was either very low, or very high.

If we standardise our trawl survey biomass indices, to make between-species comparisons easier, it becomes immediately apparent that the biomass indices for different species in a series sometimes fluctuate synchronously (Figure 4). In other words there are sometimes strong between-species correlations in a series. We will argue below that the main cause of these correlations is that catchability was extreme (for many species) in some years. Before we do that it is worth pointing an obvious example. For ECSI summ., the biomass indices for many species follow the same down-up-down pattern. In fact, 17 of the 25 species in this series had their two highest biomass indices in the same years (1996 and 1998). This would be very unlikely to occur by chance alone if there were no between-species correlations. The probability (assuming no correlations) of 17 or more of the 25 species having their highest indices in the same two years is only 9.8×10^{-8} . This very low probability is clear evidence of correlations.

We used the following procedure to identify extreme years. For each species in a trawl survey data set, the survey years were ranked in order of increasing biomass index, and then these ranks were averaged across species to obtain a mean rank for each year. Then we calculated the rank deviations, $d_i = |r_i - 0.5(n + 1)|$, where r_i is the mean rank for year i , n = the number of survey years, and $0.5(n + 1)$ is the overall mean of the mean ranks. These calculations are illustrated in Table 5, and the mean ranks are plotted in Figure 5.

Table 5: The procedure for calculating a mean rank and rank deviation for each survey year, illustrated with the BoP data set.

	Species	Survey year					
		1983	1985	1990	1992	1996	1999
Biomass indices	SNA	125	482	1565	1141	969	1644
	GUR	355	47	413	272	320	365
	LEA	63	48	131	257	118	89
	JDO	113	111	157	236	191	176
Ranks	SNA	1	2	5	4	3	6
	GUR	4	1	6	2	3	5
	LEA	2	1	5	6	4	3
	JDO	2	1	3	6	5	4
Mean ranks, r_i		2.25	1.25	4.75	4.5	3.75	4.5
Rank deviations, d_i		1.25	2.25	1.25	1.00	0.25	1.00

Extreme years are those for which the rank deviation, d_i is, large. The threshold value for large d_i will depend on the number of species and of survey years in the series. We used the following simulation procedure to determine the threshold for each series.

1. The actual biomass indices were replaced by randomly generated indices (using a uniform distribution).
2. Mean ranks, and rank deviations, were calculated for each survey year using these simulated biomass indices.
3. The largest of these rank deviations, $d_{\max,1}$, was stored.
4. Steps 1 to 3 were repeated 999 times, generating $d_{\max,j}$, for $j = 2, \dots, 1000$.
5. The threshold d was taken as the smallest number that is greater than at least 95% of the $d_{\max,j}$ (i.e., year i was labelled as extreme if d_i was greater than or equal to at least 95% of the $d_{\max,j}$).

In other words we are asking, for each rank deviation d_i , how likely we are to observe a deviation at least as large as this if there was no between-species correlations. If the probability is less than or equal to 0.05 we label the year as extreme.

A total of 16 survey years (out of 94) were found to be extreme (9 with high catchability, and 7 with low), and there were 8 survey series for which no years were extreme (Figure 5). Note how the threshold (which is half the vertical distance between the broken lines in Figure 5) decreases with increasing number of species, and decreasing number of survey years.

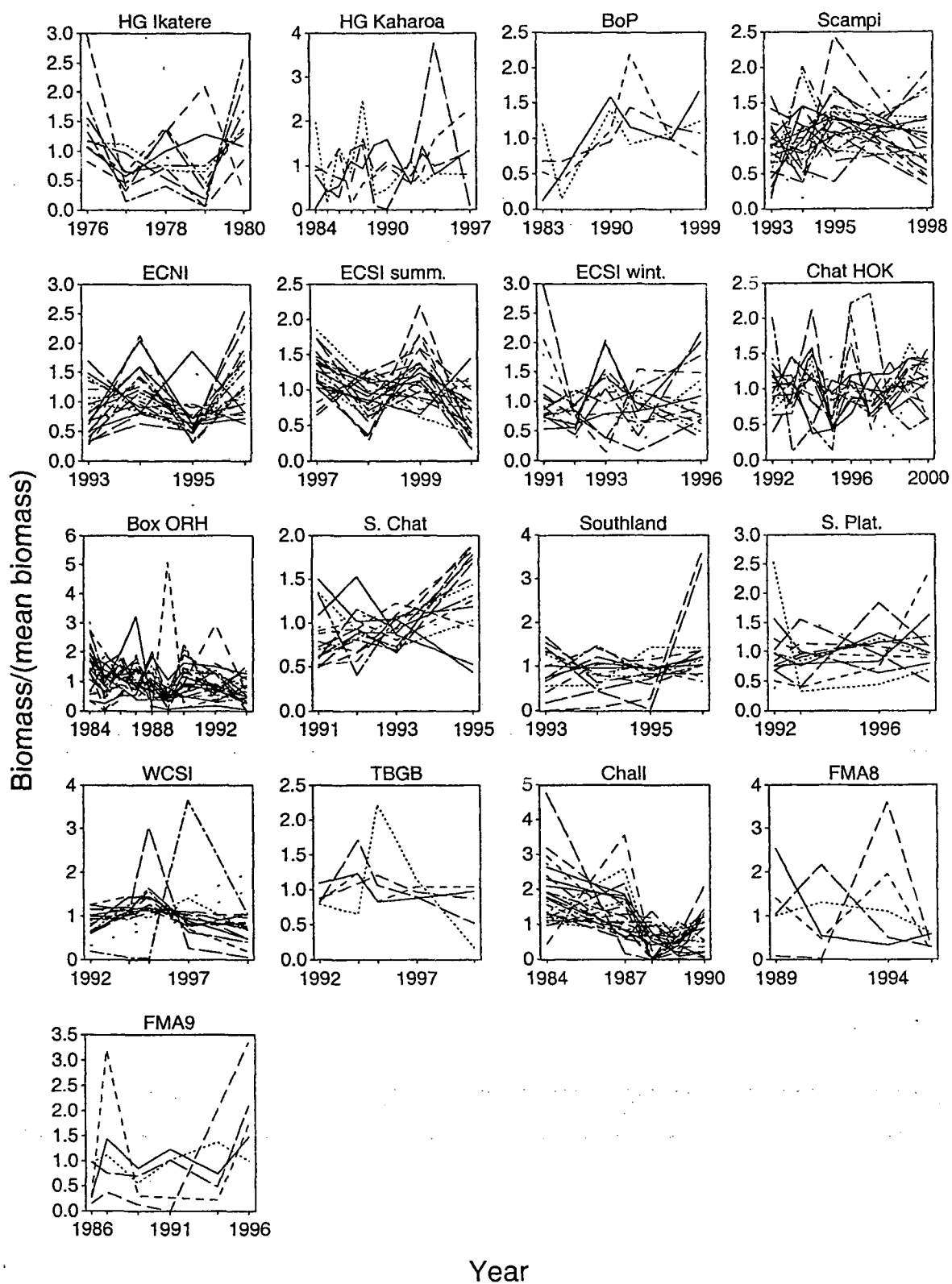


Figure 4: Standardised biomass indices, plotted against year, for each of the trawl survey data sets of Table 3. For each species in a data set, the biomass indices were standardised by dividing by the mean index for that species.

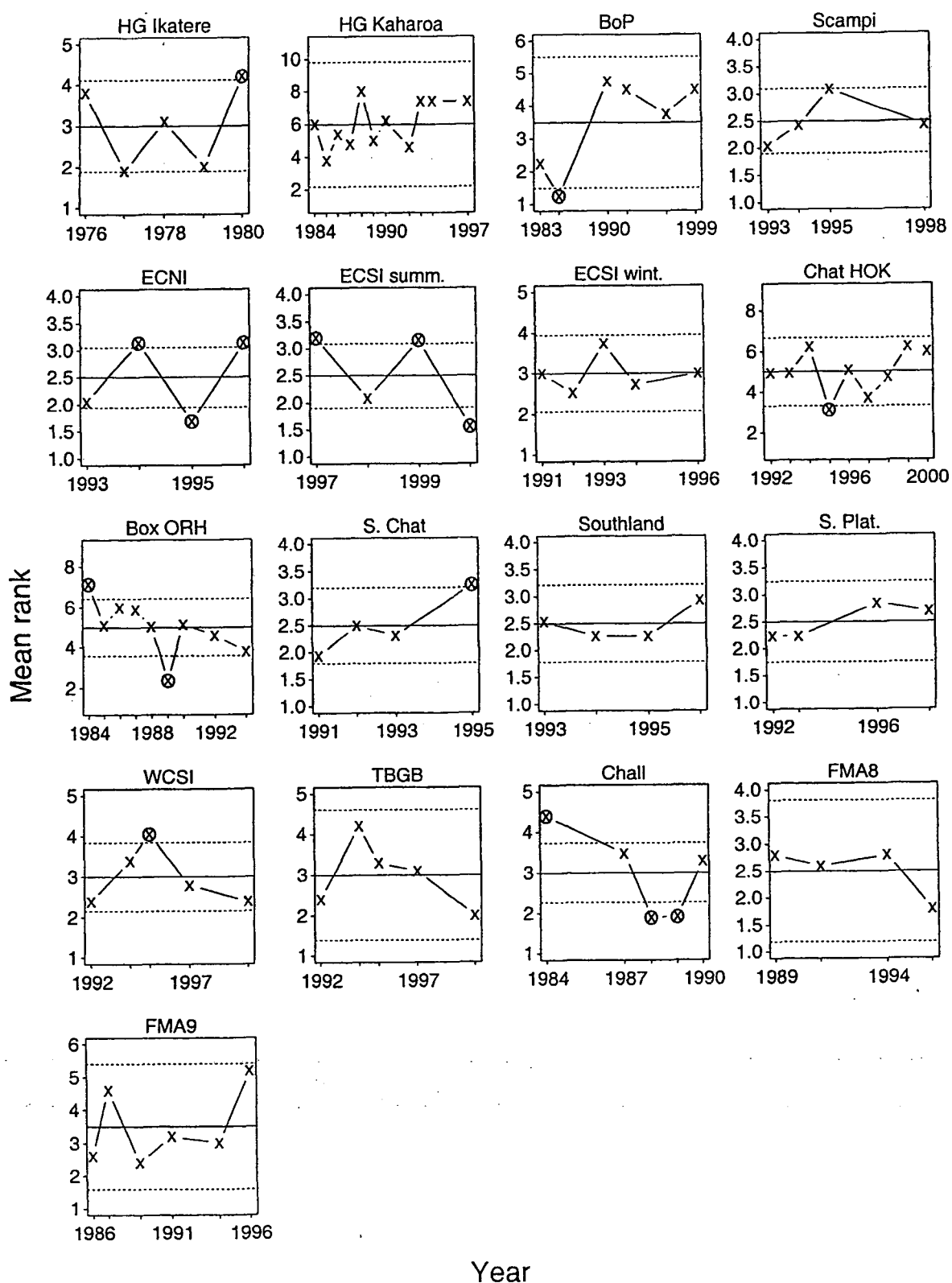


Figure 5: Mean ranks (calculated as illustrated in Table 5) for each trawl survey data set. The solid horizontal line in each panel indicates the overall mean rank; mean ranks outside the broken lines (circled points) indicate extreme survey years.

What causes these extreme years? We have said that they arise because of correlations between species. These correlations could be caused by any one of the three factors that contribute to year-to-year changes in the trawl survey biomass indices for a species: sampling error, changes in species abundance, and changes in catchability. For example, sampling error is associated with the element of chance involved in whether there happen to be many fish at a randomly chosen location at the time it is sampled by the trawl. Because some pairs of species co-occur, we can expect that if we are “lucky” with one species (i.e., we happen to hit dense concentrations of it) then we will tend to be “lucky” with its co-occurring species. Thus, the sampling errors of co-occurring species will be correlated. Further, if the abundance of one species is decreased by fishing then we can expect that this will be true for co-occurring species. Thus, changes in abundance of co-occurring species can be expected to be correlated. Finally, we don’t have a great deal of information about how environmental changes affect the catchability of fish. However, it is at least plausible that changes that, say, decrease the catchability of one species, will do the same for other species with similar behaviour. Thus, catchability changes can cause at least part of the between-species correlations we described above.

In deciding which of these three factors is most important in causing these extreme years it is useful to look at the magnitude of change of biomass indices between consecutive surveys. We expressed this change as a biomass ratio for each species, and we calculated these ratios for each pair of consecutive surveys which included one extreme year (Table 6). The value of these ratios is that they help us to rule out biomass change as an explanation for some of the extreme years. Consider, for example, the HG Ikateri series, in which the median change in biomass index between 1979 and 1980, calculated over 10 species, was a factor of more than 3. It is not plausible to say that the biomass of all these species changed by that much in just one year. A second example is years 1988 to 1990 for Box ORH. Here the median change (over 25 species) was a halving, from 1988 to 1989, followed by a doubling, from 1989 to 1990. Again, it is not plausible to say that the actual biomasses changed by this much.

What Table 6 allows is a subjective evaluation of the likelihood that these changes in biomass indices could be caused by actual biomass changes. This evaluation is complex, because the likelihood depends on the magnitude of the ratios, the number of years between surveys, the number of species involved, and any “adjacent” changes (e.g., for Box ORH, a large drop in biomass in 1989 is less plausible because it appears to be followed by a large rise in the next year). Thus it is not easy to provide a threshold and say that some changes are plausible but others are not. Nevertheless, there is a clear range of plausibility. For the examples in the previous paragraph it is extremely implausible that changes in actual biomass were responsible, but for 1995 in WCSI the biomass changes are not as implausible (but it is a matter of judgment as to whether one could call them plausible!).

It seems unlikely that the extreme mean ranks shown in Figure 5 (or the biomass ratios in Table 6) were caused solely by correlated sampling errors. In principle, we should be able to quantify this likelihood. From the survey tow-by-tow data we could infer the extent of between species correlations at the level of individual stations, from which we could calculate correlations for whole surveys. This information could then be used to calculate the probability of generating biomass ratios as big as those in Table 6. However, to do so would be a major multi-level simulation exercise which is beyond the scope of the present work. What we do know, from other studies, is that between-species correlations, when they exist, are not large. Values of 0.2 to 0.4 seem to be typical (for square-root-transformed catch rates in the same depth stratum) (B. Bull, NIWA, pers. comm.). It does not seem at all likely that such small correlations would cause the very substantial synchronous fluctuations we see in Figure 5 and Table 6.

We are left with the conclusion that the trawl survey data contain clear evidence of the failure of the assumption that research-vessel catchability does not change significantly from year to year. In most, if not all, of the circled years in Figure 5 the catchability of many species appears to have been either much higher or much lower than normal.

Table 6: Biomass ratio statistics for “extreme” years (as identified in Figure 5) in the trawl survey data. For each series, the table has one row for each pair of consecutive surveys that includes one extreme year (extreme years are underlined). A biomass ratio for the two years is calculated for each species; the table presents the median of these ratios, as well as the number of species for which this ratio exceeds 1.5. In each row, the order of the years is such that the expected biomass ratio is greater than 1.

Series	Years	Median ratio	Number of species where ratio exceeds 1.5
HG Ikatere	<u>1980</u> /1979	3.4	8/10
BoP	1983/ <u>1985</u>	1.2	1/4
	1990/ <u>1985</u>	3.0	3/4
ECNI	<u>1994</u> /1993	1.8	12/22
	<u>1994</u> /1995	1.6	14/22
	<u>1996</u> /1995	1.9	13/22
ECSI summ.	<u>1997</u> /1998	1.4	11/25
	<u>1999</u> /1998	1.5	14/25
	<u>1999</u> /2000	1.9	18/25
Chat HOK	1994/ <u>1995</u>	1.4	8/18
	1996/ <u>1995</u>	1.6	10/18
Box ORH	<u>1984</u> /1985	1.7	15/25
	1988/ <u>1989</u>	2.0	16/25
	1990/ <u>1989</u>	2.1	19/25
S. Chat	<u>1995</u> /1993	1.4	7/16
WCSI	<u>1995</u> /1994	1.1	1/18
	<u>1995</u> /1997	1.4	8/18
Chall	<u>1984</u> /1987	1.4	11/25
	1987/ <u>1988</u>	2.3	16/25
	1990/ <u>1989</u>	1.9	13/25

An obvious question to ask is whether there is any consistency, between series, in the years that are labelled as having high or low catchability. There is only limited scope for between-series comparisons because the years or seasons covered by different series may not overlap, and, in any case, only about one in six years is labelled as extreme. There are three years which are labelled as extreme for more than one series: 1984, 1989, and 1995 (Figure 6). In two of these three the labels are consistent: the two ORH surveys agree in finding catchability to be high in the winter of 1984 and low in July 1989. For 1995, two surveys found low catchability (Chat HOK in month 1 and ECNI in months 2 & 3) and two found high (S. Chat in months 10 & 11 and WCSI in months 3 & 4). Given the differences in depth ranges and months for these series it is unclear how much consistency in catchability could be expected.

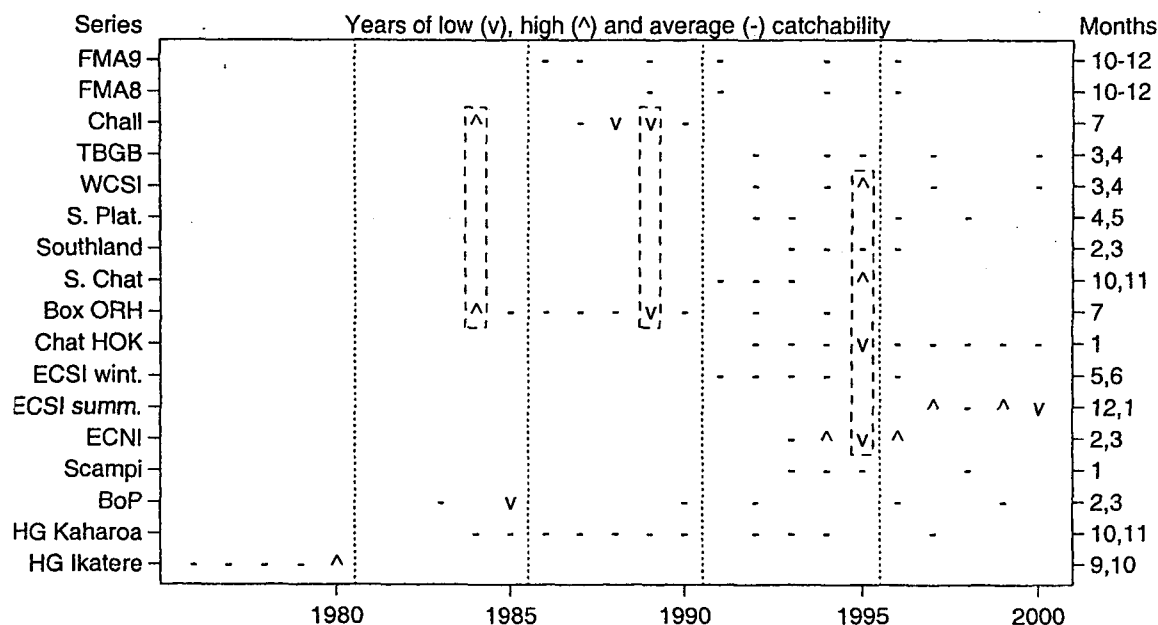


Figure 6: Estimated years of low (v), high (^), and average (-) catchability for each trawl survey series. The main survey months for each series are shown on the right axis; where there is more than one survey with extreme catchability in a year the associated symbols are circled.

We also looked to see whether there is consistency between the extreme years in the trawl survey data and the CPUE indices in the assessment data sets. Each person who provided CPUE indices was asked, for each CPUE series, which, if any, of the trawl data sets in Table 2 were “comparable” in that they relate to similar areas, depths, and seasons. For “comparability” it wasn’t necessary that the CPUE species was a target for the trawl survey. We were interested in knowing whether they thought that the fact that the catchability seemed to be extreme for many species in the trawl survey in a particular year would be reasonable grounds to believe that this would affect their CPUE index in a similar way (but they were not told which trawl survey years were considered extreme). For each match that was found between a CPUE index and a trawl survey extreme year we asked whether the two were consistent: that is, whether high (or low) trawl survey catchability corresponded to a positive (or negative) CPUE residual. We found only 12 matches, and the data were consistent at 67% of these (Table 7). This is not significantly different from the value of 50% we would expect if the data were uncorrelated ($p = 0.39$).

Table 7: Instances when residuals from CPUE assessment data are consistent, or inconsistent, with the extremes of catchability identified from the trawl survey data (see Figure 5).

Year	Trawl survey series	Trawl survey catchability	CPUE series	
			consistent	inconsistent
1984	Box ORH	High	ORH MEC	
1984	Chall	High	ORH7A	
1988	Chall	Low	ORH7A, ORH7B	
1989	Chall	Low	ORH7A, ORH7B	
1994	ECNI	High		SKI2
1995	ECNI	Low	SKI2	
1995	Chat HOK	Low		HAK4
1995	S. Chat	Low	BOE3A.DOM	SSO3A.DOM
1996	ECNI	High		SKI2

Another place one might look for consistency is between biomass trends for the same species in different survey series. We made plots of biomass trends for every instance where there were at least three survey series with that species and at least 3 years in common. Of the 16 such plots, the three showing the most consistency are given in Figure 7. The almost perfect agreement for the SPD trends in the ECSI wint. and Southland series is striking, as is the consistency of the three SDO trends. However, it's hard to judge the significance of these results because of the possibility of obtaining agreement by chance.

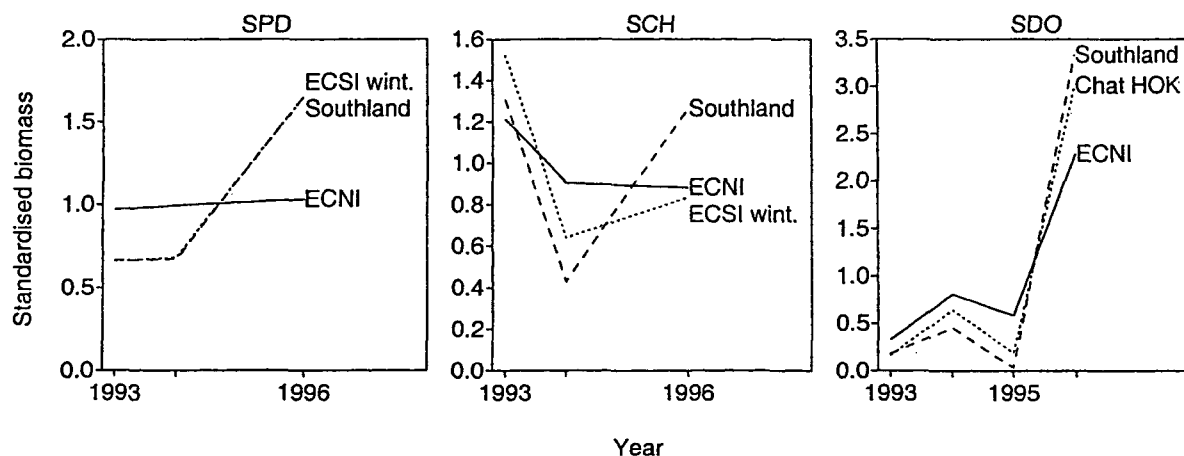


Figure 7: Examples of consistency in biomass trends for the same species in different trawl survey series. Each line shows the standardised biomass (biomass/mean biomass) for the given species in a survey series. The trend lines are plotted only for those years which were in common to all series in the plot.

4.4.1 Alternative mean rank calculations

We tried three variations on the above procedure for identifying extreme years. In each case we were evaluating an alternative hypothesis about the nature of between-species correlations. Each hypothesis leads to a different method of calculating mean ranks (or alternative statistics), and we applied the new method to both the survey data, and to simulated data (to calculate threshold values for the new statistics). If the hypothesis were true we would expect to see more extreme years. In fact, we saw fewer extreme years for all of these alternatives.

First, we repeated the above calculations after omitting species for which the mean c.v. (see Figure 1) exceeded 0.4. The idea here is that, for species with high c.v.s, there is little information in the year-to-year changes in their biomass indices. Thus these indices may mask the synchronous fluctuations in the other species, so that omitting them would produce more extreme years. In some cases it did make the most extreme rank deviations more extreme. However, it also had the effect of increasing the threshold (because the number of species decreased). The net effect was to produce slightly fewer extreme years. There was one additional extreme year — 1993 for ECSI wint.. However, the following years were no longer deemed extreme: 1980 for HG Ikateri; 1994 and 1996 for ECNI; and 1988 for Chall.

Our second alternative was based on the idea that environmental changes may affect different species differently. That is, an environmental extreme produces extreme catchability, but this may be high for some species and low for others. To test this we calculated rank deviations for each species and then averaged the rank deviations (rather than averaging the ranks and then calculating deviations). This method identified only four extreme years — three were as in Figure 5 (1984 and 1989 for Box ORH and 1995 for S. Chat) and one was new (1979 for HG Ikateri).

The third alternative was a variant on the second. We assumed that the species for each series fall into two groups: one group whose catchabilities are all affected in the same way by environmental

changes, and a second group for which the effect is opposite. That is, when catchability is high for the first group it will be low for the second, and vice versa. We calculated the mean ranks as above and then determined, for each species, the Euclidean distance between these mean ranks and: (a) the species ranks, and (b) the “inverse” of the species ranks (if a species ranks are, say, 1,4,2,3 then the inverse ranks are 4,1,3,2). When the latter distance was smaller the species was said to fall into the second group. The ranks for all group-two species were replaced by their inverse ranks and the mean ranks (and thus rank deviations) were recalculated. With this method only two extreme years were found, both of which are extreme in Figure 5 (1984 and 1989 for Box ORH). Often there was no clear separation between groups one and two. Sometimes (but only when there were few species) group two was empty. We also tried a cluster analysis approach to the identification of groups one and two but this produced no better results.

4.4.2 Environmental predictors of extreme years

We sought environmental variables that were correlated with the patterns of extreme years in Figure 5 and could, thus, have caused them. The straightforward way to do this is to assemble, for each survey series, a collection of likely local environmental variables and look for one that has highs and lows that correspond to the survey extreme years. However, this is too easy. Because the survey series are short and there are many different environmental variables to choose from (and each of these can be modified by averaging over slightly different time periods or geographical areas) it is highly likely that we could find a matching predictor for any pattern of extreme years. We could have little confidence that the selected variable was, in fact, the cause of the catchability extremes.

To avoid this problem the person who was charged with examining the environmental data (JAR) was given incomplete information. For each survey series that had at least one extreme year he was told the area and depth range covered, and the months and years of the surveys. He was also told how many years had high catchability, and how many had low, but not which years were which. His task was to compile a set of candidate environmental variables for the appropriate area and season, and then infer, from the pattern extremes in these variables, which were the extreme years. For example, for the Box ORH series he knew there was one year of high catchability and one year of low. He compiled environmental variables for the winters of the appropriate years (1984–1990, 1992, 1994) and then looked for one year in which some of the variables were extreme in one direction, and another year in which they were extreme in the other direction. In this case his task was to identify the two extreme years (but not to say which was associated with high catchability).

The following environmental variables were initially considered:

- satellite-derived sea surface temperature (SST) at between one and four locations depending on area (the SST grid resolution is $1^{\circ} \times 1^{\circ}$);
- local westerly and southerly winds, and “vorticity” (Holton 1992) at SST locations;
- the Southern Oscillation Index; and
- the leading Trenberth pressure indices (Trenberth 1976) (these are broader scale measures of wind speed and direction).

This proved to be too many variables so vorticity and the Trenberth indices were dropped. Sea surface and bottom temperature data from the trawl survey database were not used because there is doubt about the calibration of the sensors.

Of the 16 extreme years, only 7 were identified as such from the environmental data (Table 8). This is slightly less than the expected number correct (7.67) if the extreme years were chosen at random (see Appendix 4).

Table 8: Comparison between years labelled as extreme from the trawl survey data (as in Figure 5) and from environmental data (see text for details). Years correctly identified as extreme from the environmental data are underlined and the number of such is given as the actual number correct. The expected number correct was calculated assuming years were chosen at random (see Appendix A4).

Series	<u>From survey data</u>		<u>From environmental data</u>		<u>Number correct</u>	
	Low	High	"Low"	"High"	actual	expected
HG Ikatere		1980		1977	0	0.20
BoP	1985		1992		0	0.17
ECNI	1995	1994,1996	<u>1995</u>	<u>1993,1994</u>	2	2.25
ECSI summ.	2000	1997,1999	<u>1999</u>	<u>1997,2000</u>	3	2.25
Chat HOK	1995		1992		0	0.11
Box ORH	1989	1984	1988	1990	0	0.44
S. Chat		1995		<u>1995</u>	1	0.25
WCSI		1995		1992	0	0.20
Chall	1988,1989	1984	<u>1989,1990</u>	1987	<u>1</u>	<u>1.80</u>
Totals					<u>7</u>	<u>7.67</u>

It is disappointing that we were unable to demonstrate a clear connection between environmental variables and years of extreme trawl survey catchability. However, because our times series are short, we were compelled to use a technique that would be fruitful only if catchability was correlated with many environmental variables.

5. WHAT SIZE SHOULD OUR ASSESSMENT C.V.S BE?

In Section 4.1 we showed that the c.v.s used for CPUE indices in New Zealand stock assessments are, more often than not, too large, whereas those for trawl survey indices are often too small. So what is the right size for these c.v.s? We can get an approximate answer to this problem by changing the assumed c.v.s for some of the assessment data sets of Table 2 and recalculating the residual statistic κ . Our aim is to obtain sets of values of κ in which approximately half are negative and half are positive. This would suggest that our new assumed c.v.s are about the right size. We will do this separately for the CPUE and trawl survey indices.

Before proceeding we should acknowledge a deficiency in this procedure. Strictly speaking, when we change the assumed c.v.s we should rerun the assessment model and generate new model biomass estimates. However, it is not practical to do this for so many assessments. Thus, we have to assume that the change in c.v. will not change the model estimates too much. Our experience is that this is often true for assessments with only one series of biomass indices. It is least likely to be true when there are more than one series and the series show markedly different trends.

It would seem, from Table 2, that there are two common "default" values for c.v.s for CPUE indices used in New Zealand: 0.3 and 0.35. One or other of these values is used in 19 of the 30 data sets with CPUE indices. For these 19 data sets, we investigated the effect on κ of using various alternative default values. The best value lies between 0.15 and 0.2. Values in this range give approximately equal numbers of positive and negative values of κ (Table 9).

Table 9: Effect of using different default c.v.s for the 19 data sets of Table 2 for which the indices are from CPUE and the assumed c.v. is either 0.3 or 0.35. Each line of the table gives the number of these data sets for which κ falls in the given range for the given default c.v.

Default c.v.	Number of data sets			
	$\kappa < -2$	$-2 < \kappa < 0$	$0 < \kappa < 2$	$2 < \kappa$
0.35	7	9	3	0
0.3	6	10	3	0
0.25	3	10	5	1
0.2	0	11	5	3
0.15	0	8	6	5
0.1	0	1	9	9

We took a slightly different approach for the assessment data sets with trawl survey indices. We took the c.v.s calculated from the trawl survey data to be estimates of the observation error for the biomass indices. For some data sets (ORH NEChat, SNA8.age2, and SNA8.age3) these are what were assumed in the assessment (and thus are given in Table 2); the c.v.s for the other data sets are given in Table 10. We then assumed that the c.v. associated with annual variation in catchability was the same for all stocks and “added” this c.v. to the observation error c.v.s to get assumed c.v.s for the stock assessments. Note that c.v.s are “added” as squares, so that when we “add” c.v.s of 0.2 and 0.3 we get $0.36 [= \sqrt{(0.2^2 + 0.3^2)}]$.

Table 10: Observation error c.v.s (calculated from the trawl survey data) for the assessment data sets of Table 2 with survey indices (excluding those data sets where these c.v.s are the same as those in Table 2).

Label	c.v.s
GUR1E.1	0.09–0.23
GUR1E.2	0.09–0.36
GUR1W	0.12–0.30
GUR2	0.16–0.44
JDO1E.3	0.12–0.44
JDO1E.4	0.12–0.35
JDO1W.2	0.15–0.47
JDO2E	0.17–0.48
JDO2W	0.25–0.62
LIN34.2	0.06–0.16
LIN56.2	0.05–0.09
STA34	0.11–0.23
STA4	0.13–0.26
WAR3.1	0.36–0.40
WAR3.2	0.43–0.72

The best default value for a trawl survey annual variation c.v. appears to be about 0.2. This is the value that gives approximately equal numbers of positive and negative values of κ (Table 11).

Table 11: Effect of using different default annual variation c.v.s for the 18 assessment data sets of Table 2 with trawl survey indices. Each line of the table gives the number of these data sets for which κ falls in the given range when the assumed c.v. in the assessment is calculated by “adding” the given default c.v. to the observation error c.v.s.

Default c.v. for annual variation	Number of data sets			
	$\kappa < -2$	$-2 < \kappa < 0$	$0 < \kappa < 2$	$2 < \kappa$
0	0	4	10	4
0.1	0	6	9	3
0.15	0	6	11	1
0.2	0	9	8	1
0.25	1	9	8	0
0.3	1	11	6	0

6. DISCUSSION

It is difficult to make inferences about catchability because we cannot measure it directly. Instead, we must estimate it indirectly via stock assessment models. These estimates are compromised by the weakness of our models, which provide only crude representations of population dynamics (because the data to develop more complex models are not available). With trawl survey data alone we cannot estimate catchability but can only detect years when it was extreme for many species. Nevertheless, the large data sets we have assembled (see Tables 2 & 3) do allow us to draw some conclusions about New Zealand catchabilities.

For CPUE, the common (but usually tacit) assumption that catchability varies from year to year is supported. It is clear from Figure 2 and Table 4 that, had the CPUE c.v.s been set equal to the observational error (typically less than 0.1, Francis 1999b), the associated stock assessment residuals would have been much too large. However, in allowing for annual variability we seem to have gone too far. The c.v.s that are used in stock assessments are, more often than not, too large. In other words, the annual variability in CPUE catchability is not as large as is implied by these c.v.s. Where the use of a default c.v. is appropriate it would seem that values around 0.15–0.2 would be better than the values of 0.3–0.35 that are currently used (see Section 5).

We found two reasons to believe that trawl survey catchability varies from year to year. First, if we use observation error c.v.s (as routinely calculated from trawl survey data) in stock assessments, we get residuals which are, more often than not, larger than they ought to be. Second, in about one year in six the (trawl survey) catchability of many species is extreme enough to be detectable (Figure 6). The results of Table 11 suggest that 0.2 is a reasonable default c.v. for annual variability in trawl survey catchability. This c.v. should be “added” to the observation error c.v.s to obtain a c.v. for use in stock assessments.

For many trawl survey series it may be sensible to use the same observation error c.v. for all years. For example, in the JDO1E.3 series the estimated c.v.s for the five years are 0.24, 0.12, 0.16, 0.12, and 0.44. If the design of all surveys was approximately the same, and there has been no major change in fish distribution, then it is reasonable to assume that the true c.v. was the same for all surveys and our five values are all estimates of this. We could then use the mean of these five values as the observation error c.v. for all five years (a mean is better than a median here because the sampling distribution of an estimated c.v. is typically strongly skewed). If, on the other hand, much of the variation between the five c.v. estimates is attributable to changes in survey design (in particular, the number of tows), or to a change in fish distribution, as is the case for Box ORH (Francis 1996), then it would be better to have a different observation error c.v. for each year.

The blanket use of default c.v.s is clearly undesirable. It is obviously wrong to assume that all CPUE indices have the same c.v., regardless of which species or fishery they describe, or the quality and quantity of data from which they are calculated. Similarly, we should expect that annual variability in trawl survey catchability will vary from stock to stock. However, we have little choice in this matter. In most stock assessments we will not have the information to depart from a default value. Two obvious exceptions are the SSO data set in Table 2, for which Doonan et al. (1999a) showed that the CPUE data were unusually weak. The above default values imply smaller c.v.s for CPUE than for trawl surveys. This is surprising, and contrary to the prevailing view that the latter are more reliable than the former. Nevertheless, it is clearly indicated by the data sets examined here.

The other important result is that some estimates of trawl survey catchability are not credible. For both JDO and GUR some estimates are implausibly low, and the variation amongst stocks of these same species is implausibly high. Where possible, this variability should also be examined for CPUE catchabilities (as long as they are comparable — we should not compare trawl and long-line catchabilities). This was not done here because it involves adjusting for different reference units of effort in different CPUE series, which requires specialist knowledge about the individual fisheries and data sets.

The significance of the autocorrelations we found in Section 4.2 is dubious. The positive autocorrelations are not, perhaps, unexpected. If annual changes in catchability are driven by temperature then this is likely to produce such correlations. However, these correlations could be artifacts. Our assessment models are, of necessity, rather simplistic. If they are insufficiently flexible to track actual changes in biomass they will tend to produce autocorrelated residuals. Also, as pointed out above, the simulation procedure we used to judge statistical significance is flawed.

Our analyses have not been able to take account of the practice, in some stock assessments, of using c.v.s as a measure of the “quality” of a biomass index, rather than its precision. This happens when a high c.v. is assigned to a series (often of CPUE) that is believed not to index biomass well. The intention is to lessen the contribution of the series to the assessment. A problem with this is that the judgment of quality is subjective, as is the decision as to how high a c.v. to assign to represent poor quality. It would be very rare that we had sufficient information to determine whether the judgment of poor quality was justified, and if the assigned c.v.s were appropriate. It may be that some of the assessments analysed above produced a biomass trajectory that was a very good fit to a CPUE series (suggesting that a low c.v. should have been used) but that the trajectory was wrong because, in this case, CPUE was not proportional to abundance. We cannot distinguish such an outcome from one in which a precise CPUE series indexed abundance well. The practice of assigning c.v.s subjectively is not desirable. Ideally, we should change the model assumption of proportionality between biomass and index rather than inflating c.v.s. However, we acknowledge that stock assessment is a very pragmatic discipline in which many compromises are necessary, and we hope that the above results will provide practitioners with empirical evidence to support some of their subjective decisions.

7. RECOMMENDATIONS

1. Default c.v.s for CPUE indices in stock assessments should be lower than is currently common in New Zealand: 0.15–0.2, rather than 0.3–0.35.
2. Default c.v.s for trawl survey indices should be calculated by “adding” an annual variability c.v. of 0.2 to the survey observation error c.v. (reminder: c.v.s “add” as squares, so 0.2 “added” to 0.3 is $\sqrt{(0.2^2 + 0.3^2)} = 0.36$).
3. Estimated trawl survey catchabilities that are very different from 1 should be treated as an error indication in stock assessments. Either the trawl survey series is not an appropriate method of tracking this particular stock or the survey indices are strongly in conflict with other model inputs.
4. Large within-species variation in catchability should also be treated as an error indication (but catchabilities should be compared only if they are of the same type).

8. ACKNOWLEDGMENTS

We are indebted to Neil Bagley, for providing *Ikatere* data from the *fish_comm* database, to Dave Gilbert, for explaining the strata he defined for the *Ikatere* surveys, to Richard O’Driscoll, who found a blatant error in our data, and to the many people who provided trawl survey and assessment data for this work. Dave Gilbert, Larry Paul, Paul Breen, and Stuart Hanchet offered useful comments on a draft of this report. This project was funded by the Ministry of Fisheries under project SAM1999/01. The construction of *fish_comm* was funded under contract C01830 from the New Zealand Foundation for Research, Science and Technology.

9. REFERENCES

- Alvial Muñoz, A. (1988). Impactos de "El Niño" sobre el manejo de las pesquerías pelágicas en el Pacífico oriental. *Medio Ambiente* 9(1): 35–41.
- Anderson, O. F. (2000). Assessment of the East Cape hills (ORH 2A North) orange roughy fishery for the 2000–01 fishing year. *New Zealand Fisheries Assessment Report 2000/19*. 29 p.
- Anderson O.F.; Fenaughty, J.F. (1996). Trawl surveys of orange roughy on the Chatham Rise 1984–1992. *New Zealand Fisheries Data Report 81*. 116 p.
- Angelsen, K.K.; Olsen, S. (1987). Impact of fish density and effort level on catching efficiency of fishing gear. *Fisheries Research* 5: 271–278.
- Annala, J.H.; Sullivan, K.J.; O'Brien, C.J.; Iball, S.D. (1998). Report from the Fishery Assessment Plenary, May 1998: stock assessments and yield estimates. (Unpublished report held in the NIWA library, Wellington.)
- Arreguín-Sánchez, F. (1996). Catchability: a key parameter for fish stock assessment. *Reviews in Fish Biology and Fisheries* 6(2): 221–242.
- Bagge, O.; Munch-Petersen, S. (1979). Some possible factors governing the catchability of Norway lobster in the Kattegat. *Rapport et proces-verbaux des réunions du Conseil Permanent International pour l'Exploration de la Mer* 175: 143–146.
- Bagley, N.W.; Hurst, R.J. (1996). Trawl survey of middle depth and inshore bottom species off Southland, February–March 1996 (TAN9604). *New Zealand Fisheries Data Report 77*. 11 p.
- Bagley, N.W.; Livingston, M.E. (2000). Trawl survey of hoki and middle depth species on the Chatham Rise, January 1999 (TAN9901). *NIWA Technical Report 81*. 52 p.
- Bagley, N.W.; McMillan, P.J. (1999). Trawl survey of hake and middle depth species in the Southland and Sub-Antarctic areas, April–May 1998 (TAN9805). *NIWA Technical Report 52*. 48 p.
- Bailey, G.W.; Beyers, C.J.de B.; Lipschitz, S.R. (1985). Seasonal variation of oxygen deficiency in waters off southern South West Africa in 1975 and 1976 and its relation to the catchability and distribution of the Cape rock lobster *Jasus lalandii*. *South African Journal of Marine Science* 3: 197–214.
- Bannerot, S.P.; Austin, C.B. (1983). Using frequency distributions of catch per unit effort to measure fish-stock abundance. *Transactions of the American Fisheries Society* 112(5): 608–617.
- Beentjes, M.P.; Stevenson, M.L. (2000). Review of the east coast South Island winter trawl survey time series, 1991–96. *NIWA Technical Report 86*. 64 p.
- Borgstrøm, R. (1992). Effect of population density on gillnet catchability in four allopatric populations of brown trout (*Salmo trutta*). *Canadian Journal of Fisheries and Aquatic Sciences* 49(8): 1539–1545.
- Breen, P.A.; Andrew, N.L.; Kendrick, T.H. (2000). Stock assessment of paua (*Haliotis iris*) in PAU 5B and PAU 5D using a new length-based model. *New Zealand Fisheries Assessment Report 2000/33*. 37 p.
- Brodziak, J.; Hendrickson, L. (1999). An analysis of environmental effects on survey catches of squids *Loligo pealei* and *Illex illecebrosus* in the northwest Atlantic. *Fishery Bulletin* 97(1): 9–24.
- Buijse, A.D.; Schaap, L.A.; Bult, T.P. (1992). Influence of water clarity on the catchability of six freshwater fish species in bottom trawls. *Canadian Journal of Fisheries and Aquatic Sciences* 49(5): 885–893.
- Cayre, P.; Roy, C. (1986). Variations in yield of yellowfin (*Thunnus albacares*) and skipjack (*Katsuwonus pelamis*) in relation to interannual anomalies of surface temperature. *Collective Volume of Scientific Papers, International Commission for the Conservation of Atlantic Tunas* 25: 67–76.
- Clark, M.R.; Tracey, D.M.; Baird, G.G. (1991). Trawl survey of orange roughy on the Challenger Plateau, July 1990. *New Zealand Fisheries Technical Report 26*. 20 p.
- Cook, R.M.; Armstrong, D.W. (1985). Changes in catchability of cod, haddock and whiting associated with the Scottish seine-net fleet. *Journal du Conseil International pour l'Exploration de la Mer* 42(2): 171–178.

- Cordue, P.L. (1998). Designing optimal estimators for fish stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 376–386.
- Cordue, P.L. (2000). MIAEL estimation of biomass and fishery indicators for the 1999 assessment of hoki stocks. *New Zealand Fisheries Assessment Report 2000/10*. 69 p.
- Crecco, V.; Overholtz, W.J. (1990). Causes of density-dependent catchability for Georges Bank haddock *Melanogrammus aeglefinus*. *Canadian Journal of Fisheries and Aquatic Sciences* 47(2): 385–394.
- Crecco, V.A.; Savoy, T.F. (1985). Density-dependent catchability and its potential causes and consequences on Connecticut River American shad, *Alosa sapidissima*. *Canadian Journal of Fisheries and Aquatic Sciences* 42(10): 1649–1657.
- Cryer, M. (1996). Scampi stock assessment for 1996. New Zealand Fisheries Assessment Research Document 96/10. 25 p. (Unpublished report held in NIWA library, Wellington.)
- Csirke, J. (1989). Changes in the catchability coefficient in the Peruvian anchoveta (*Engraulis ringens*) fishery. In: Pauly, D.; Muck, P.; Mendo, J.; Tsukayama, I. (eds) The Peruvian upwelling ecosystem: dynamics and interactions. *ICLARM Conference Proceedings* 18: 207–219.
- Danzmann, R.G.; MacLennan, D.S.; Hector, D.G.; Hebert, P.D.N.; Kolasa, J. (1991). Acute and final temperature preferenda as predictors of Lake St. Clair fish catchability. *Canadian Journal of Fisheries and Aquatic Sciences* 48(8): 1408–1418.
- Doonan, I.J.; McMillan, P.J.; Coburn, R.P.; Hart, A.C.; Cordue, P.L. (1995). Assessment of smooth oreo for 1995. New Zealand Fisheries Assessment Research Document 95/12 31 p. (Unpublished report held in NIWA library, Wellington.)
- Doonan, I.J.; McMillan, P.J.; Coburn, R.P.; Hart, A.C. (1999a). Assessment of OEO 3A smooth oreo for 1999–2000. New Zealand Fisheries Assessment Research Document 99/45 21 p. (Unpublished report held in NIWA library, Wellington.)
- Doonan, I.J.; McMillan, P.J.; Coburn, R.P.; Hart, A.C. (1999b). Assessment of OEO 3A black oreo for 1999–2000. New Zealand Fisheries Assessment Research Document 99/52 30 p. (Unpublished report held in NIWA library, Wellington.)
- Dunn, A.; Harley, S.J.; Doonan, I.J.; Bull, B. (2000a). Calculation and interpretation of catch-per-unit-effort (CPUE) indices. *New Zealand Fisheries Assessment Report 2000/1*. 44 p.
- Dunn, A.; Horn, P.L.; Cordue, P.L.; Kendrick, T.H. (2000b). Stock assessment of hake (*Merluccius australis*) for the 1999–00 fishing year. *New Zealand Fisheries Assessment Report 2000/50*. 50 p.
- Evans, C.R.; Opnai, L.J.; Kare, B.D. (1997). Fishery ecology and oceanography of the prawn *Penaeus merguensis* (de Man) in the Gulf of Papua: estimation of maximum sustainable yield and modelling of yield, effort and rainfall. *Marine and Freshwater Research* 48(3): 219–228.
- Fonteneau, A. (1991). Anomaly of the environment in 1984 in the Gulf of Guinea. Possible effects on yellowfin catchability. *Collective Volume of Scientific Papers, International Commission for the Conservation of Atlantic Tunas* 36: 380–408.
- Francis, R.I.C.C. (1981). Stratified random trawl surveys of deep-water demersal fish stocks around New Zealand. *Fisheries Research Division Occasional Publication No. 32*. 28 p.
- Francis, R.I.C.C. (1989). A standard approach to biomass estimation from bottom trawl surveys. New Zealand Fisheries Assessment Research Document 89/3. (Unpublished report held in NIWA library, Wellington.)
- Francis, R.I.C.C. (1996). Orange roughy sex ratios and catchrate distributions in the Chatham Rise Spawning Box. New Zealand Fisheries Assessment Research Document 96/13. 28 p. (Unpublished report held in NIWA library, Wellington.)
- Francis, R.I.C.C. (1999a). Stock assessment of orange roughy on the northeast Chatham Rise. New Zealand Fisheries Assessment Research Document 99/41. 16 p. (Unpublished report held in NIWA library, Wellington.)
- Francis, R.I.C.C. (1999b). The impact of correlations in standardised CPUE indices. New Zealand Fisheries Assessment Research Document 99/42. 30 p. (Unpublished report held in NIWA library, Wellington.)

- Francis, R.I.C.C.; Field, K.D. (2000). CPUE analysis and assessment of the Mid-East Coast orange roughy stock (ORH 2A South, 2B, 3A). *New Zealand Fisheries Assessment Report 2000/29*. 20 p.
- Fréchet, A. (1990). Catchability variations of cod in the marginal ice zone. *Canadian Journal of Fisheries and Aquatic Sciences* 47(9): 1678–1683.
- Gilbert, D.J.; McKenzie, J.R.; Davies, N.M.; Field, K.D. (2000). Assessment of the SNA 1 stocks for the 1999–2000 fishing year. *New Zealand Fisheries Assessment Report 2000/38*. 52 p.
- Hanchet, S.M. (1999). Stock assessment of trevally (*Caranx georgianus*) in TRE 7. New Zealand Fisheries Assessment Research Document 99/55. 27 p. (Unpublished report held in NIWA library, Wellington.)
- Hanchet, S.M. (2000). Southern blue whiting (*Micromesistius australis*) stock assessment for the Campbell Island Rise for 1999 and 2000. *New Zealand Fisheries Assessment Report 2000/15*. 36 p.
- Hanchet, S.M.; Coburn, R.P.; Field, K.D.; Horn, P.L.; Ó Maolagáin, C.; Paul, L.J.; Stevenson, M.L. (2000). Assessment of red gurnard (*Chelidonichthys kumu*) stocks GUR 1 and GUR 2. *New Zealand Fisheries Assessment Report 2000/40*. 51 p.
- Hart, A.C.; McMillan, P.J. (1998). Trawl survey of oreos and orange roughy on the south Chatham Rise, October–November 1995 (TAN9511). *NIWA Technical Report* 27. 48 p.
- Henderson, B.A.; Collins, J.J.; Reckahn, J.A. (1983). Dynamics of an exploited population of lake whitefish (*Coregonus clupeaformis*) in Lake Huron. *Canadian Journal of Fisheries and Aquatic Sciences* 40: 1556–1567.
- Hilborn, R.; Walters, C.J. (1992). Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York. 570 p.
- Holton, J.R. (1992). An Introduction to Dynamic Meteorology. International Geophysics Series, Dmowska, R.; Holton, J.R. (eds), Academic Press, San Diego. 511 p.
- Horn, P.L.; Bagley, N.W. (1999). Stock assessment of blue warehou (*Seriola lalandi*) in New Zealand waters. New Zealand Fisheries Assessment Research Document 99/26. 18 p. (Unpublished report held in NIWA library, Wellington.)
- Horn, P.L.; Hanchet, S.M.; Stevenson, M.L.; Kendrick, T.H.; Paul, L.J. (1999). Catch history, CPUE analysis, and stock assessment of John dory (*Zeus faber*) around the North Island (Fishstocks JDO 1 and JDO 2). New Zealand Fisheries Assessment Research Document 99/33. 58 p. (Unpublished report held in NIWA library, Wellington.)
- Horn, P.L.; Harley, S.J.; Ballara, S.L.; Dean, H. (2000). Stock assessment of ling (*Genypterus blacodes*) around the South Island, New Zealand (Fishstocks LIN 3, 4, 5, 6, and 7). *New Zealand Fisheries Assessment Report 2000/37*. 70 p.
- Houghton, R.G.; Flatman, S. (1981). The exploitation pattern, density-dependent catchability, and growth of cod (*Gadus morhua*) in the west-central North Sea. *Journal du Conseil International pour l'Exploration de la Mer* 39(3): 271–287.
- Hurst, R.J.; Coburn, R.P.; Horn, P.L. (2000). Assessment of northern gemfish stocks (SKI 1 and SKI 2) for 2000. *New Zealand Fisheries Assessment Report 2000/18*. 41 p.
- Jakobsson, J. (1980). Exploitation of the Icelandic spring- and summer-spawning herring in relation to fisheries management, 1947–1977. *Rapport et proces-verbaux des réunions du Conseil Permanent International pour l'Exploration de la Mer* 177: 23–42.
- MacCall, A.D. (1976). Density dependence of catchability coefficient in the California Pacific sardine, *Sardinops sagax caerulea*, purse seine fishery. *Reports of the California Cooperative Oceanic Fisheries Investigations* 18: 136–148.
- Macpherson, E.; Gordoa, A. (1992). Trends in the demersal fish community off Namibia from 1983 to 1990. *South African Journal of Marine Science* 12: 653–649.
- Morrison, M.A. (1997). Trawl survey of snapper and associated species in the Bay of Plenty, February 1996 (KAH9601). *NIWA Technical Report* 2. 32 p.
- Morrison, M.A.; Francis, M.P. (1999). Trawl survey of snapper and associated species in the Hauraki Gulf, October–November 1997 (KAH9720). *NIWA Technical Report* 58. 37 p.
- Morrison, M.A.; Stevenson, M.L. (in press). Review of West Coast North Island trawl survey time series (1986–96). *NIWA Technical Report*.

- Morriissy, N.M.; Caputi, N. (1981). Use of catchability equations for population estimation of marron, *Cherax tenuimanus* (Smith) (Decapoda: Parastacidae). *Australian Journal of Marine and Freshwater Research* 32(2): 213–225.
- Nielsen, L.A. (1983). Variation in the catchability of yellow perch in an otter trawl. *Transactions of the American Fisheries Society* 112(1): 53–59.
- Paloheimo, J.E.; Dickie, L.M. (1964). Abundance and fishing success. *Rapport et proces-verbaux des réunions du Conseil Permanent International pour l'Exploration de la Mer* 155: 152–163.
- Paul, L.J. (1992). *Ikatere* research cruises, 1959–81: a record of coastal fisheries research, mainly in northern waters. MAF Fisheries Greta Point Internal Report 185. 72 p. (Unpublished report held in NIWA library, Wellington.)
- Paul, L.J.; Hanchet, S.M.; Horn, P.L.; Stevenson, M.L. (1999). Catch history and stock assessment of stargazer (*Kathetostoma giganteum*) on the Chatham Rise (Fishstock STA 4). New Zealand Fisheries Assessment Research Document 99/48. 35 p. (Unpublished report held in NIWA library, Wellington.)
- Perry, R.I.; Smith, S.J. (1994). Identifying habitat associations of marine fishes using survey data: an application to the northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences* 51(3): 589–602.
- Peterman, R.M.; Steer, G.J. (1981). Relation between sport-fishing catchability coefficients and salmon abundance. *Transactions of the American Fisheries Society* 110(5): 585–593.
- Peterman, R.M.; Steer, G.J.; Bradford, M.J. (1985). Density-dependent catchability coefficients. *Transactions of the American Fisheries Society* 114: 438–440.
- Pinhorn, A.T. (1988). Catchability in some of the major groundfish fisheries off the east coast of Canada. *Journal of Northwest Atlantic Fishery Science* 8: 15–23.
- Polovina, J.J. (1986). A variable catchability version of the Leslie model with application to an intensive fishing experiment on a multispecies stock. *Fishery Bulletin* 84(2): 423–428.
- Pope, J.G. (1980). Some consequences for fisheries management of aspects of the behaviour of pelagic fish. *Rapport et Proces-verbaux des Réunions du Conseil Permanent International pour l'Exploration de la Mer* 177: 466–476.
- Pope, J.G.; Garrod, D.J. (1975). Sources of error in catch and effort quota regulations with particular reference to variations in the catchability coefficient. *Research Bulletin, International Commission Northwest Atlantic Fisheries* 11: 17–30.
- Quinn, T.J., II; Collie, J.S. (1990). Alternative population models for eastern Bering Sea pollock. *International North Pacific Fisheries Commission Bulletin* 50: 243–257.
- Quinn, T.J., II; Deriso, R.B. (1999). Quantitative fish dynamics. Oxford University Press, New York. 542 p.
- Richards, L.J. (1987). Comparing imprecise abundance indices with a symmetric model. *Canadian Journal of Fishery and Aquatic Sciences* 44: 793–802.
- Richards, L.J.; Schnute, J.T. (1986). An experimental and statistical approach to the question: Is CPUE an index of abundance? *Canadian Journal of Fishery and Aquatic Sciences* 43: 1214–1227.
- Rose, G.A.; Leggett, W.C. (1991). Effects of biomass-range interactions on catchability of migratory demersal fish by mobile fisheries: An example of Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences* 48(5): 843–848.
- Schaaf, W.E. (1975). Fish population models: potential and actual links to ecological models. In: Russell, C.S. (ed.) *Ecological modeling in a resource management framework*. pp. 211–239. Resources for the Future, Inc., Washington, D.C.
- Shardlow, T.F. (1993). Components analysis of a density-dependent catchability coefficient in a salmon hook and line fishery. *Canadian Journal of Fisheries and Aquatic Sciences* 50(3): 513–520.
- Shardlow, T.; Hilborn, R. (1985). Density-dependent catchability coefficients. *Transactions of the American Fisheries Society* 114: 436–438.
- Skog, K.; Mikkelsen, K.O.; Bjordal, Å. (1994). Wrasse — availability and catchability in the period February-June. *Fisken og Havet* 4. 34 p.

- Stevenson, M.L. (1998). Inshore trawl survey of the west coast South Island and Tasman and Golden Bays, March-April 1997 (KAH9701). *NIWA Technical Report 12*. 70 p.
- Stevenson, M.L.; Beentjes, M.P. (in press). Inshore trawl survey of the Canterbury Bight and Pegasus Bay, December 1999-January 2000 (KAH9917 and CMP9901). *NIWA Technical Report*.
- Stevenson, M.; Hanchet, S. (2000). Review of the inshore trawl survey series of the east coast of the North Island, 1993-96. *NIWA Technical Report 85*. 58 p.
- Swain, D.P.; Nielsen, G.A.; Sinclair, A.F.; Chouinard, G.A. (1994). Changes in catchability of Atlantic cod (*Gadus morhua*) to an otter-trawl fishery and research survey in the southern Gulf of St Lawrence. *ICES Journal of Marine Science* 51(4): 493-504.
- Tanaka, E. (1997). A model of catchability of developing fishery. *Fisheries Science* 63(4): 496-502.
- Tracey, D. M.; Fenaughty, J. M. (1997). Trawl survey of orange roughy on the Chatham Rise, May-July 1994 (TAN9406). *New Zealand Fisheries Technical Report 44*. 43 p.
- Trenberth, K.E. (1976). Fluctuations and trends in indices of the Southern Hemisphere circulation. *Quarterly Journal of the Royal Meteorological Society*, 102: 65-75.
- Ulltang, Ø. (1980). Factors affecting the reaction of pelagic fish stocks to exploitation and requiring a new approach to assessment and management. *Rapport et Proces-verbaux des Réunions du Conseil Permanent International pour l'Exploration de la Mer* 177: 489-504.
- Vignaux, M. (1994). Documentation of trawlsurvey analysis program. MAF Fisheries Greta Point Internal Report 225. 44 p. (Unpublished report held in NIWA library, Wellington.)
- Winters, G.H.; Wheeler, J.P. (1985). Interaction between stock area, stock abundance, and catchability coefficient. *Canadian Journal of Fisheries and Aquatic Sciences* 42(5): 989-998.
- Yamakawa, T.; Matsumiya, Y.; Nishimura, M.; Ohnishi, S. (1994). Expanded Delury's method with variable catchability and its application to catch-effort data from spiny lobster gillnet fishery. *Fisheries Science* 60(1): 59-63.
- Zheng, X.; Basher, R.E. (1999). Structural time series models and trend detection in global and regional temperature series. *Journal of Climate* 12: 2347-2358.

Appendix 1 Details of the trawl survey data

Table A1: Voyage codes, species included, and references for the trawl survey series listed in Table 3

Voyage codes are of the form abc9313, which describes the 13th voyage in 1993 with vessel whose code is abc. Vessel codes: aex = *Amaltal Explorer*, arr = *Arrow*, buc = *Buccaneer*, cor = *Cordella*, ika = *Ikatere*, kah = *Kaharoa*, tan = *Tangaroa*, wil = *Will Watch*. Where possible, references are for the whole series; otherwise they are for the latest documented survey.

Label	Voyage codes	Species included	Reference
HG Ikatere	ika7612,ika7711,ika7810, ika7910,ika8008	SNA,JDO,JMA,GUR,SPO,EGR, SFL,LEA,TRE,LSO	Paul (1992)
HG Kaharoa	kah8421,kah8517,kah8613, kah8716,kah8810, kah8917, kah9016,kah9212,kah9311, kah9411,kah9720	SNA,GUR,LEA,JDO,SFL	Morrison & Francis (1999)
BoP	kah8303,kah8506,kah9004, kah9202,kah9601,kah9902	SNA,GUR,LEA,JDO	Morrison (1997)
Scampi	kah9301,kah9401,kah9501, kah9801	SCI,LIN,HOK,SKI,LDO,MDO, TAR,RIB,NSD,BRZ,SPE,JAV, SRH,CDO,FHD,BBE,PRK,SSI, YBO,ETL,CUC,RCO,GSH	Cryer (1996)
ECNI	kah9304,kah9402,kah9502, kah9602	BAR,SPD,RCO,TAR,STA,SPE, GUR,RSK,SSK,CAR,FRO,SNA, SCH,SPO,JMN,TRE,LIN,HOK, JDO,SDO,MDO,ERA	Stevenson & Hanchet (2000)
ECSI summ.	kah9618,kah9704,kah9809, kah9917	SPD,BAR,RCO,GSH,SPE,SQU, LIN,TAR,HOK,ELE,CAR,STA, SWA,WAR,GUR,SSK,SCG,WIT, SCH,RSK,PIG,JAV,SPO,LSO, ESO	Stevenson & Beentjes (in press)
ECSI wint.	kah9105,kah9205,kah9306, kah9406,kah9606	SPD,BAR,RCO,TAR,ELE,CAR, WAR,GUR,SCG,WIT,SCH,PIG, SPO,LSO,ESO	Beentjes & Stevenson (2000)
Chat HOK	tan9106,tan9212,tan9401, tan9501,tan9601,tan9701, tan9801,tan9901,tan0001	HOK,LDO,LIN,JAV,SPE,CBO, GSP,HAK,BBE,COL,CAS,GSH, CFA,RIB,FHD,SND,SDO,CBI	Bagley & Livingston (2000)
Box ORH	buc8401,buc8501,buc8601, buc8701,cor8801,cor8901, cor9002,tan9206,tan9406	ORH,SND,BEE,ETB,SSO,WHX, RIB,SBI,HJO,GSP,CYP,RCH, CYO,LCH,SOR,CMA,BSH,PLS, BSL,CSQ,SMC,MCA,SSM,SBK, TRS	Anderson & Fenaughty (1996), Tracey & Fenaughty (1997)
S. Chat	tan9104,tan9210,tan9309, tan9511	SSO,BOE,SND,ETB,SBI,HJO, SSM,CYP GSP,BEE,BJA,MIQ, RCH,MCA,LCH,WHX	Hart & McMillan (1998)
Southland	tan9301,tan9402,tan9502, tan9604	NOS,SPD,STA,BAR,RSK,SCH, CAR,TAR,WIT,SKI,SPE, SSK, BGZ,RCO,SDO	Bagley & Hurst (1996)
S. Plat.	tan9204,tan9304,tan9605, tan9805	HOK,JAV,LIN,GSP,CFA,HAK, CAS,COL,RIB,LDO,TOP,CBO, SND	Bagley & McMillan (1999)
WCSI	kah9204,kah9404,kah9504, kah9701,kah0004	BAR,SPD,RCO,TAR,STA,SPE, GUR,GSH,RSK,SSK,CAR,SCH, SPO,JMD,JMM,LIN,SDO,NSD	Stevenson (1998)
TBGB	kah9204,kah9404,kah9504, kah9701,kah0004	LEA,BCO,ESO,LSO,SFL	Stevenson (1998)
Chall	arr8401,aex8701,aex8801, aex8901,wil9001	ORH,RIB,SOR,HJO,CYO,CMA, SND,BEE,CYP,RCH,BSL,SBI, VSQ,ETB,TRS,CSQ,SSM,GSP, BSH,NBU,RSQ,LCH,CYL,WSQ, TOP	Clark et al. (1991)
FMA8	kah8918,kah9111,kah9410, kah9615	SNA,GUR,BAR,SPD,JDO	Morrison & Stevenson (in press)
FMA9	kah8612,kah8715,kah8918, kah9111,kah9410,kah9615	SNA,GUR,BAR,SPD,JDO	Morrison & Stevenson (in press)

Appendix 2 Biomass calculations for *Ikatere* HG Data

The calculation of biomass estimates for the *Ikatere* HG trawl survey did not use the standard trawl survey program (which was used for all other surveys) because of differences in the survey design (station positions were fixed — the same position was re-occupied for each survey — not randomly generated, and not related to a formal stratification) and in data recording practices (catches were recorded by number, rather than weight, and some species were not measured). This appendix briefly describes the calculations that were made.

Catch rates (kg km^{-2}) were taken from database *fish_comm*, which is a copy of part of the Ministry of Fisheries *trawl* database, except that the data have been groomed and some fields have been added. Where possible, catch weights were estimated from a length frequency and a length-weight relationship; otherwise they were based on numbers caught multiplied by a mean weight per individual (derived from other surveys in the area). The area covered by each tow was taken as the product of the distance towed and an assumed door spread of 44 m (same for all surveys).

The following stations were dropped as unreliable: station 366 in IKA7711 and station 309 in IKA7910 (D.J. Gilbert, NIWA, pers. comm.).

The strata used to calculate biomass are those in the *trawl* database. These were defined (by D.J. Gilbert) many years after the surveys by combining strata used in KAH8421 (Table A2). Standard formulae (see, for example, Francis 1981) were used to calculate biomasses and c.v.s.

Table A2: Strata used in calculating biomasses for the *Ikatere* HG surveys (areas include untrawlable ground).

Stratum	Area (km^2)	Strata combined from KAH8421
410	2754	41
310	1312	11,31,32,42
130	1081	12,13,14,15
230	310	23
270	1498	21,22,24,25
260	886	26,80

Appendix 3 Species list

Table A3 Common and scientific names for species codes used in this report

Code	Common name	Scientific name
BAR	BARRACOUTA	<i>Thyrsites atun</i>
BBE	BANDED BELLOWSFISH	<i>Centriscoops humerosus</i>
BCO	BLUE COD	<i>Parapercis colias</i>
BEE	BASKETWORK EEL	<i>Diastobranchus capensis</i>
BGZ	BANDED GIANT STARGAZER	<i>Kathetostoma spp</i>
BJA	BLACK JAVELINFISH	<i>Mesobius antipodum</i>
BOE	BLACK OREO	<i>Alloctytus niger</i>
BRZ	BROWN STARGAZER	<i>Xenocephalus armatus</i>
BSH	SEAL SHARK	<i>Dalatias licha</i>
BSL	BLACK SLICKHEAD	<i>Xenodermichthys spp</i>
CAR	CARPET SHARK	<i>Cephaloscyllium isabellum</i>
CAS	OBLIQUE BANDED RATTAIL	<i>Caelorinchus aspercephalus</i>
CBI	TWO SADDLE RATTAIL	<i>Caelorinchus biclinozonalis</i>
CBO	BOLLONS RATTAIL	<i>Caelorinchus bollonsi</i>
CDO	CAPRO DORY	<i>Capromimus abbreviatus</i>
CFA	BANDED RATTAIL	<i>Caelorinchus fasciatus</i>
CMA	MAHIA RATTAIL	<i>Caelorinchus matamua</i>
COL	OLIVERS RATTAIL	<i>Caelorinchus oliverianus</i>
CSQ	CENTROPHORUS SQUAMOSUS	<i>Centrophorus squamosus</i>
CUC	CUCUMBER FISH	<i>Chlorophthalmus nigripinnis</i>
CYL	CENTROSCYMNUS COELOLEPIS	<i>Centroscymnus coelolepis</i>
CYO	SMOOTH SKIN DOGFISH	<i>Centroscymnus owstoni</i>
CYP	CENTROSCYMNUS CREPIDATER	<i>Centroscymnus crepidater</i>
EGR	EAGLE RAY	<i>Myliobatis tenuicaudatus</i>
ELE	ELEPHANTFISH	<i>Callorhynchus milii</i>
ERA	ELECTRIC RAY	<i>Torpedo fairchildi</i>
ESO	N.Z. SOLE	<i>Peltorhamphus novaezeelandiae</i>
ETB	BAXTERS LANTERN DOGFISH	<i>Etmopterus baxteri</i>
ETL	LUCIFER DOGFISH	<i>Etmopterus lucifer</i>
FHD	DEEPSEA FLATHEAD	<i>Hoplichthys haswelli</i>
FRO	FROSTFISH	<i>Lepidopus caudatus</i>
GSH	GHOST SHARK	<i>Hydrolagus novaezeelandiae</i>
GSP	PALE GHOST SHARK	<i>Hydrolagus sp b2</i>
GUR	GURNARD	<i>Chelidonichthys kumu</i>
HAK	HAKE	<i>Merluccius australis</i>
HJO	JOHNSON'S COD	<i>Halargyreus johnsonii</i>
HOK	HOKI	<i>Macruronus novaezeelandiae</i>
JAV	JAVELIN FISH	<i>Lepidorhynchus denticulatus</i>
JDO	JOHN DORY	<i>Zeus faber</i>
JMA	JACK MACKEREL	<i>Trachurus declivis, t.m., t.nz.</i>
JMD	TRACHURUS DECLIVIS	<i>Trachurus declivis</i>
JMM	TRACHURUS MURPHYI	<i>Trachurus murphyi</i>
JMN	TRACHURUS NOVAEZELANDIAE	<i>Trachurus novaezeelandiae</i>
LCH	LONG-NOSED CHIMAERA	<i>Harriotta raleighana</i>
LDO	LOOKDOWN DORY	<i>Cyttus traversi</i>
LEA	LEATHERJACKET	<i>Parika scaber</i>
LIN	LING	<i>Genypterus blacodes</i>
LSO	LEMON SOLE	<i>Pelotretis flavilatus</i>
MCA	RIDGE SCALED RATTAIL	<i>Macrourus carinatus</i>
MDO	MIRROR DORY	<i>Zenopsis nebulosus</i>
MIQ	WARTY SQUID	<i>Moroteuthis ingens</i>
NBU	BULBOUS RATTAIL	<i>Kuronezumia bubonis</i>
NOS	NOTOTODARUS SLOANII	<i>Nototodarus sloanii</i>
NSD	NORTHERN SPINY DOGFISH	<i>Squalus mitsukurii</i>
ORH	ORANGE ROUGHY	<i>Hoplostethus atlanticus</i>

Table A3 continued Common and scientific names for species codes used in this report

Code	Common name	Scientific name
PIG	PIGFISH	<i>Congiopodus leucopaecilus</i>
PLS	PLUNKETS SHARK	<i>Centroscyminus plunketi</i>
PRK	PRAWN KILLER	<i>Ibacus alticrenatus</i>
RCH	WIDENOSED CHIMAERA	<i>Rhinochimaera pacifica</i>
RCO	RED COD	<i>Pseudophycis bachus</i>
RIB	RIBALDO	<i>Mora moro</i>
RSK	ROUGH SKATE	<i>Raja nasuta</i>
RSQ	OMMASTREPHEs BARTRAMI	<i>Ommastrephes bartrami</i>
SBI	SLICKHEAD, BIGSCALED BROWN	<i>Alepocephalus sp</i>
SBK	SPINEBACK	<i>Notacanthus sexspinis</i>
SCG	SCALY GURNARD	<i>Lepidotrigla brachyoptera</i>
SCH	SCHOOL SHARK	<i>Galeorhinus galeus</i>
SCI	SCAMPI	<i>Metanephrops challenger</i>
SDO	SILVER DORY	<i>Cyttus novaezealandiae</i>
SFL	SAND FLOUNDER	<i>Rhombosolea plebeia</i>
SKI	GEMFISH	<i>Rexea solandri</i>
SMC	SMALL-HEADED COD	<i>Lepidion microcephalus</i>
SNA	SNAPPER	<i>Pagrus auratus</i>
SND	SHOVELNOSE SPINY DOGFISH	<i>Deania calcea</i>
SOR	SPIKY OREO	<i>Neocyttus rhomboidalis</i>
SPD	SPINY DOGFISH	<i>Squalus acanthias</i>
SPE	SEA PERCH	<i>Helicolenus spp.</i>
SPO	RIG	<i>Mustelus lenticulatus</i>
SQU	ARROW SQUID	<i>Nototodarus sloanii & n gouldi</i>
SRH	SILVER ROUGHY	<i>Hoplostethus mediterraneus</i>
SSI	SILVERSIDE	<i>Argentina elongata</i>
SSK	SMOOTH SKATE	<i>Raja innominata</i>
SSM	SLICKHEAD, SMALLSCALED BROWN	<i>Alepocephalus australis</i>
SSO	SMOOTH OREO	<i>Pseudocyttus maculatus</i>
STA	GIANT STARGAZER	<i>Kathetostoma giganteum</i>
SWA	SILVER WAREHOU	<i>Seriola punctata</i>
TAR	TARAKIHI	<i>Nemadactylus macropterus</i>
TOP	PALE TOADFISH	<i>Neophrynichthys angustus</i>
TRE	TREVALLY	<i>Pseudocaranx dentex</i>
TRS	TRACHYSCORPIA CAPENSIS	<i>Trachyscorpia capensis</i>
VSQ	VIOLET SQUID	<i>Histioteuthis spp</i>
WAR	COMMON WAREHOU	<i>Seriola brama</i>
WHX	WHITE RATTAIl	<i>Trachyrincus aphyodes</i>
WIT	WITCH	<i>Arnoglossus scapha</i>
WSQ	WARTY SQUID	<i>Moroteuthis spp</i>
YBO	YELLOW BOARFISH	<i>Pentaceros decacanthus</i>

Appendix 4 Expected number right

In the experiment described in Section 4.4.2, 7 of the 17 survey-year combinations that were labelled as extreme on the basis of environmental data were “correct” (i.e., were also labelled extreme in Figure 5). In this appendix we show that if, instead of using the environmental data, we had simply picked the 17 survey-year combinations at random, then, on average, we would expect to get 8.4 of these correct.

Suppose we have a set of N years, and we know that n of these are extreme, but we don't know which. We want to calculate how many extreme years we would get right, on average, if we chose n years at random. We can calculate this using combinations. The probability of being correct for exactly k of n randomly chosen years is $P_k = \frac{{}_nC_k \times {}_{N-n}C_{n-k}}{{}_NC_n}$ (the top line is the number of ways of picking k of the n extreme years and $n-k$ of the $N-n$ non-extreme years; the bottom line is the number of ways of picking n years out of N). Thus the expected number of correct years is simply $\sum_k kP_k$. Applying this formula to each survey with extreme years we obtain the results given in Table A4.

Table A4: Number of extreme years (as identified in Figure 5), total number of years, and expected number correct if extreme years were chosen at random, for each series with extreme years.

Series	Number of extreme years	Total number of years	Expected number correct
HG Ikaterere	1	5	0.20
BoP	1	6	0.17
ECNI	3	4	2.25
ECSI summ.	3	4	2.25
Chat HOK	1	9	0.11
Box ORH	2	9	0.44
S. Chat	1	4	0.25
WCSI	1	5	0.20
Chall	3	5	<u>1.80</u>
		Total	<u>7.67</u>