Available online at www.sciencedirect.com

SCIENCE DIRECT®

Fisheries Research 70 (2004) 251–264

FISHERIES RESEARCH

www.elsevier.com/locate/fishres

A generalized linear mixed model analysis of a multi-vessel fishery resource survey

Thomas E. Helser[a,*], André E. Punt[b], Richard D. Methot[a]

[a] *National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northwest Fisheries Science Center, 2725 Montlake Blvd., East, Seattle, WA 98112, USA*

[b] *School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195-5020, USA*

Abstract

A generalized linear mixed model (GLMM) that treats year and spatial cell as fixed effects while treating vessel as a random effect is used to examine fishing power among chartered industry-based vessels and a research trawler, the *FRV Miller Freeman*, for bottom trawl surveys on the upper continental slope of U.S. West coast. A Bernoulli distribution is used to model the probability of a non-zero haul and the gamma distribution to model the non-zero catch rates of four groundfish species. The use of vessel as a random effect allows the data for the various vessels to be combined and a single continuous time-series of biomass indices to be developed for stock assessment purposes. The GLMMs fit the data reasonably well. Among the different models examined, the GLMM incorporating a random vessel × year effect had the smallest ΔAIC and was thus chosen as the best model. Also, estimated random effects coefficients associated with the industry-based vessels and the *FRV Miller Freeman* for each year suggests that these vessels can be assumed to be from a common random effects distribution. These results suggest that combining data from the chartered industry-based vessels and from the research trawler may be appropriate to develop indices of abundance for stock assessment purposes. Finally, an evaluation of variances associated with abundance indices from the different models indicate that analyzing these data as a fixed effect GLM may underestimate the level of variability due to ignoring the grouped nature of tows within vessels. As such, use of a mixed model approach with vessel as a random effect is a reasonable approach to developing abundance indices and their variances.

Keywords: Trawl survey; Generalized linear mixed model; Delta distribution; Gamma distribution

1. Introduction

Stock assessment of slope groundfish resources off the west coast of the United States is generally based on analyses of data collected through fishery resource surveys. The primary index of abundance included in stock assessments for deepwater slope species such as Dover sole (*Microstomus pacifcus*), sablefish (*Anoplopoma fimbria*), shortspine thornyhead (*Sebastolobus alascanus*), and longspine thornyhead (*Sebastolobus altivelis*) has been developed from the Alaska Fisheries Science Center's (AFSC) survey of the upper-

* Corresponding author. Tel.: +1 206 302 2436; fax: +1 206 302 2435.

E-mail address: thomas.helser@noaa.gov (T.E. Helser).

0165-7836/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.fishres.2004.08.007

continental slope (Lauth et al., 1998). This resource survey, which started in 1988, involves the use of a single fishery research trawler, the *FRV Miller Freeman*, that performs standardized hauls using a bottom trawl. One of the major advantages of using a single vessel to monitor fishery resources is that inter-haul variability is not confounded by operational differences among different survey vessels, as might be the case for a multi-vessel resource survey (Sissenwine and Bowman, 1978; Byrne et al., 1981). However, the constraints of expense and availability of ship time associated with surveys conducted by a single research vessel may necessitate the use of surveys involving multiple vessels. The vessels used in such surveys are typically chartered from the fishing industry by management agencies. This type of resource survey has become more routine in recent years both in the United States and elsewhere (Wilkins et al., 1998; Fiorentini et al., 1999; Cotter, 2001).

The Northwest Fisheries Science Center (NWFSC) initiated a multi-vessel slope survey in 1998 using vessels chartered from the groundfish fishery to develop a new time series of information on the abundance and distribution of slope groundfish, as well as to collect biological data (Methot et al., 2000). The NWFSC survey was, in many ways, designed similarly to the AFSC survey in that the spatial coverage of both surveys is largely consistent in terms of depth (183–1280 m) and latitude ($35^{\circ}00'$–$48^{\circ}10'$N), particularly in recent years. A major factor driving the introduction of the multi-vessel survey was the anticipation that the AFSC slope survey would be terminated in 2001, in essence ending the time series of abundance data for slope groundfish species. Given that future stock assessments of these species are dependent on an index of abundance, an obvious question was whether the biomass indices based on the *FRV Miller Freeman* and those based on the chartered survey vessels could be combined into a single time series. If so, the NWFSC slope survey could be used to continue the slope survey historically conducted by the *FRV Miller Freeman*.

Fishery catch and effort (CPUE) data have often been used to develop indices of relative abundance. Catch rates are standardized to remove the effect of factors such as area, depth and changes in fleet composition through time (Gavaris, 1980; Hilborn and Walters, 1992; Punt et al., 2000). Generalized linear models (GLMs – McCullaugh and Nelder, 1989) are most commonly applied to standardizing CPUE data. Vessel is often treated as a fixed model effect in the analysis and estimated relative to a standard 'hypothetical' vessel. In this study, GLMs are used to standardize the data collected during the AFSC and NWFSC surveys. However, the analyses of this paper treat year and spatial cell as fixed model effects and vessel as a random effect (Wolfinger and O'Connell, 1993) using generalized linear mixed models (GLMMs). Factor effects are random if the levels of the factor that are used represent only a random sample of a larger set of potential levels, and that the larger set of levels constitutes a population of effects with a probability distribution. The use of four (five including the *FRV Miller Freeman*) vessels each year, in essence, mimics this idea. Furthermore, the procedure whereby industry vessels are chosen from a larger pool of vessels via a bid process, supports treating vessel as a random effect in the model. The analysis also accounts for two distinct processes associated with the capture of fish in a bottom trawl: the probability of obtaining a non-zero catch and the catch-rate given that the catch is non-zero.

The analyses presented in this paper consider two primary objectives: (a) to estimate vessel effects for Dover sole, sablefish, shortspine thornyhead, and longspine thornyhead, and hence evaluate whether it is appropriate to treat vessel as a random effect in further analyses, and (b) if vessel can be treated as a random effect, to determine annual biomass indices and their coefficients of variation.

2. Methods

2.1. Basic data

The data were obtained from bottom trawl hauls made by the NWFSC and AFSC slope surveys. The primary difference between these surveys is that NWFSC slope survey is conducted using multiple vessels in an annual open bid charter system while the AFSC slope survey is conducted by a single vessel, the *FRV Miller Freeman*. While several other differences also exist between these surveys, only tow distance and net width (which are used to derive bottom area swept) are discussed in detail here because catches need to be standardized by appropriate measures of effort. Interested readers should see Lauth et al. (1998) for details of the AFSC slope survey, and Turk et al. (2001) for details of the NWFSC slope survey.

In the NWFSC slope survey, operational protocols between the industry-based vessels have been standardized including a target towing speed of 2.2 knots over ground with a nominal tow duration of 15 min between net set time and lift-off. GPS navigation, a Simrad ITI net mensuration system and a bottom contact sensor are used to monitor trawl performance and calculate haul distance and net width. While the *FRV Miller Freeman* uses a similar towing speed (2.3 knots), tow duration was standardized to 30 min in recent years, but lasted as long as 1 h in deep water in early years. Net performance was monitored using SCANMAR and distances fished for hauls made by the *FRV Miller Freeman* were calculated using a bottom contact sensor with GPS.

2.2. Stratum definitions

The purpose of stratifying a survey area is to account for spatially heterogeneous density patterns of the target species (Gunderson, 1993). In essence, this reduces within-stratum variance while increasing between-stratum variance. Preliminary analysis of the survey data using a GLM showed a statistically significant interaction between depth and latitude. This indicated, for example, that catch rates for both Dover sole and shortspine thornyhead increased with depth and at lower latitudes. Studies of slope groundfish have shown ontogenetic movements from shallow water on the continental shelf to deep water on the slope (Jacobson and Hunter, 1993; Jacobson and Vetter, 1996). It is likely that the rate of movement from shallow to deep water also differs from south to north due to changes in the width of the continental shelf. Ten strata (or spatial cells) were therefore selected based on specifications for both latitude and depth to account for possible gradients in abundance and life history characteristics. The latitude boundaries were based on area defined by the International North Pacific Fishery Commission (INPFC), except that the Vancouver–Columbia boundary was moved to just south of the Columbia River (43°00′N latitude) (Fig. 1) to increase sample size in the Vancouver region. Only two depth strata (depth stratum 1, 183–566 m; depth stratum 2, 567–1280 m) were

Table 1
Summary of the basic data for each of the ten spatial cells

Year	Vancouver		Columbia		Eureka		Monterey		Conception	
	183–566 m	567–1280 m	183–566 m	567–1280 m	183–566 m	567–1280 m	183–566 m	567–1280 m	183–566 m	567–1280 m
1998	17 (17, 14, 17, 6)	21 (11, 19, 19, 21)	47 (47, 38, 47, 13)	31 (20, 30, 26, 31)	29 (29, 23, 28, 8)	29 (28, 28, 25, 27)	43 (43, 26, 29,7)	55 (50, 54, 53, 55)	11 (11, 6, 6, 4)	17 (16, 15, 17, 17)
1999	36 (36, 32, 35, 6)	56 (25, 53, 54, 55)	61 (59, 55, 88, 21)	69 (45, 68, 62, 69)	43 (43, 36, 41, 10)	57 (51, 55, 53, 57)	72 (72, 61, 64, 23)	75 (68, 73, 75, 75)	27 (27, 26, 23, 13)	29 (22, 28, 29, 29)
2000	36 (36, 31, 34, 9)	44 (22, 41, 41, 44)	60 (60, 57, 58, 18)	67 (43, 67, 63, 67)	41 (43, 39, 41, 9)	56 (51, 55, 49, 56)	76 (75, 66, 60, 23)	92 (86, 89, 89, 91)	30 (29, 28, 23, 11)	33 (27, 32, 33, 33)
2001	39 (37, 33, 37, 8)	47 (24, 45, 45, 23)	74 (74, 70, 71, 18)	55 (32, 52, 45, 38)	46 (45, 46, 44, 12)	50 (47, 46, 41, 37)	73 (72, 71, 61, 30)	90 (83,88, 90, 89)	33 (33, 31, 27, 14)	33 (27, 32, 33, 33)
2002	25 (23, 22, 23, 8)	23 (11, 20, 22, 23)	48 (48, 48, 45, 18)	38 (20, 36, 33, 38)	33 (33, 33, 32, 12)	37 (34, 35, 30, 37)	65 (64, 62, 53, 24)	54 (46, 47, 50, 53)	52 (51, 39, 37, 19)	52 (35, 45, 48, 50)

The entries for each combination of year and stratum are the total number of hauls, followed by the numbers of non-zero hauls for Dover sole, sablefish, shortspine thornyhead and longspine thornyhead, respectively.

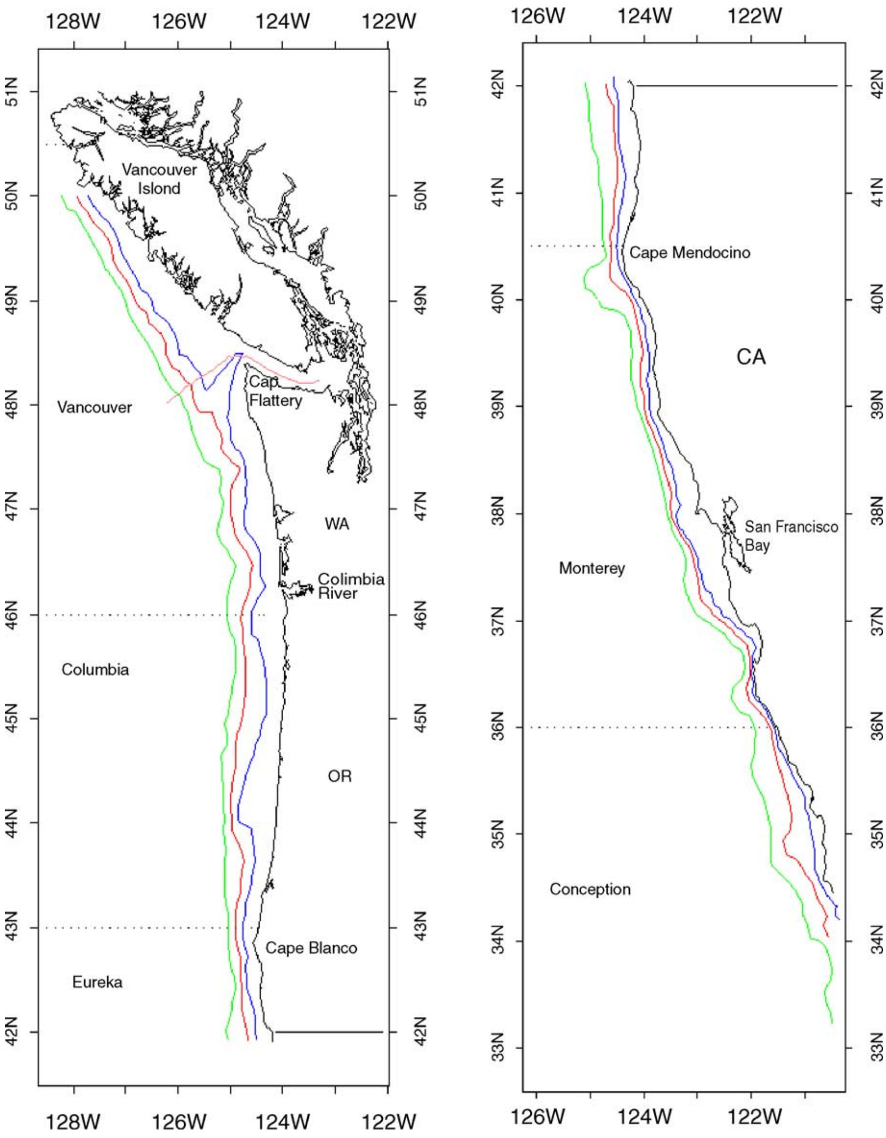

Fig. 1. Area by depth stratification scheme used to develop biomass indices. The strata consist of the Vancouver, Columbia, Eureka, Monterey and Conception INPFC areas each subdivided into two depth strata: 183–566 m, and 567–1280 m.

considered. Table 1 provides the total number of hauls within each stratum as well as the number of hauls with non-zero catches for each of the four species. The number of hauls between 1998 and 2000 is substantially larger than for previous years because both the ASFC and NWFSC surveys occurred during these years. Note that data prior to 1998 could not be used in this analysis because the NWFSC slope survey only started in this year.

2.3. Standardization of survey catch and effort data

Survey effort was calculated in terms of bottom area swept by taking the product of the distance fished and the net wingspread. Distance fished, net wingspread and swept area differed substantially between the *FRV Miller Freeman* and the industry vessels (Fig. 2). The bottom area swept during the NWFSC surveys

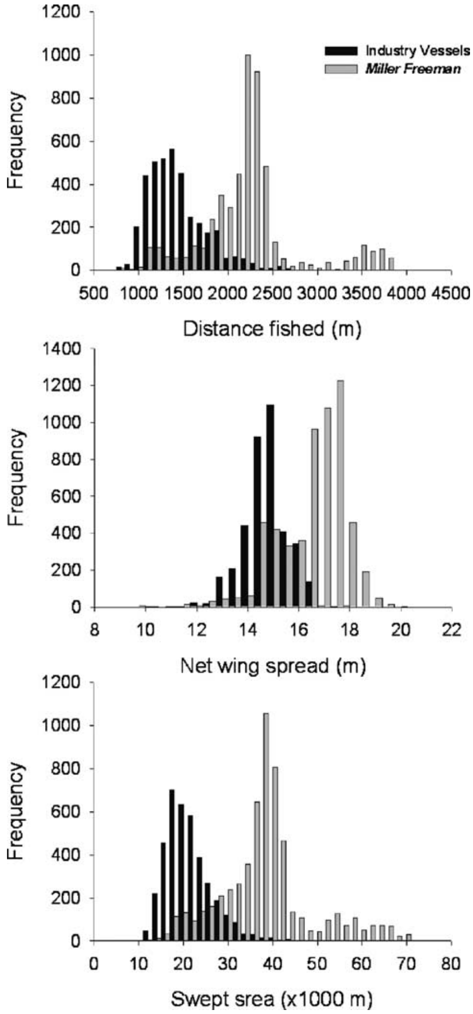

Fig. 2. Histograms of the distance fished (m), net wingspread (m) and bottom area swept (000's m; distance fished × wingspread) for years in the AFSC and NWFSC surveys.

(10 000–40 000 m^2; mean approximately 20 000 m^2) is considerably less than during the *FRV Miller Freeman* surveys (15 000–70 000 m^2; mean approximately 40 000 m^2) because the industry vessels used during the NWFSC slope surveys use trawl nets with a narrower wingspread and a shorter towing distance. In the GLMMs, survey effort, standardized on a 2 ha basis (effort/20 000 m^2), was used as an offset. We recognize the uncertainty in the accurate measurement of net wingspread, particularly the influence of the effective wingspread and its potential effect on herding. However, because of the limited 15 min tows, we suspect that this is not a problem with the species analyzed, but such effects have been estimated for groundfish in the Australian northern trawl fishery (Ramm and Xiao, 1995).

The response variables used in the GLMMs included a zero catch coded as 0 or a positive catch coded as 1. If the catch was positive then the response variable was simply taken as the catch in weight (kg) per haul. The catch distributions for all four species are highly skewed (Fig. 3); the modes of the catch rate distributions occurred at low values for all species and up to 10% of the catch rates was zero. Even when the zero tows were eliminated and the data examined on the logarithmic scale, there was still evidence for heavy skewness.

2.4. Estimation of biomass indices

The objective of the analysis is to determine biomass indices and their coefficients of variation[1], e.g. B_i and $\mathrm{cv}(B_i)$, where i denotes year. The values of both quantities are given by

$$B_i = \sum_s a^s D_i^s \qquad \mathrm{cv}(B_i) = \frac{1}{B_i}\sqrt{\sum_s (a^s)^2 \mathrm{sd}(D_i^s)^2} \tag{1}$$

where a^s is the area of stratum s, D_i^s the estimated density in stratum s during year i, and $\mathrm{sd}(X)$ denotes the standard error of X.

The delta distribution (Aitchison and Brown, 1957) was used to model the survey data because there are many zero catches. This model assumes that it is possible to treat separately the question of whether or not a catch is zero, and the size of the catch given that it is non-zero (Vignaux, 1996). Pennington (1983) and Stefánsson (1996) have shown that this approach leads to efficient estimators of the abundance for surveys in which there are many zero catches. The expected density in stratum s during year i is calculated by multiplying the probability of a non-zero haul in stratum s during year i by the expected catch-rate (kg/m^2) given that the catch is non-zero, i.e.

$$D_i^s = P_i^s V_i^s \tag{2}$$

[1] Also referred to as "relative standard errors".

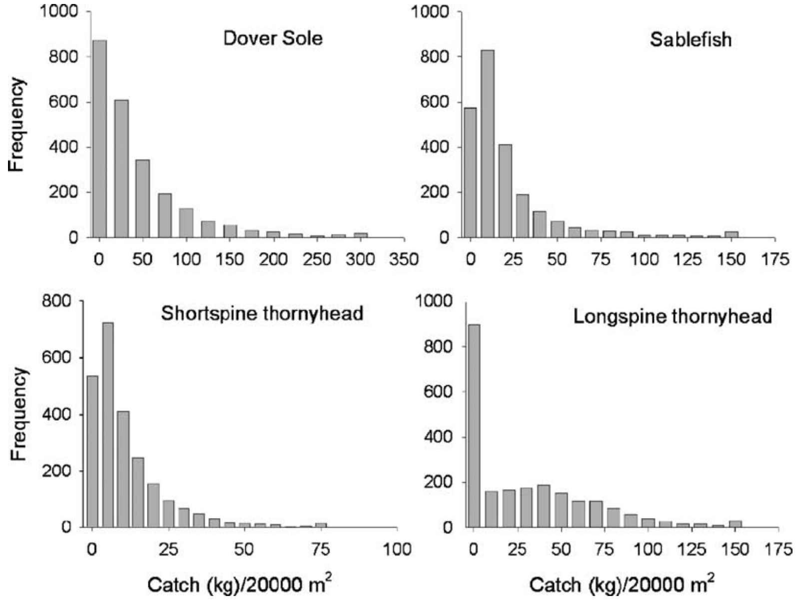

Fig. 3. Distribution of the catch rates for the four species (all years and surveys combined).

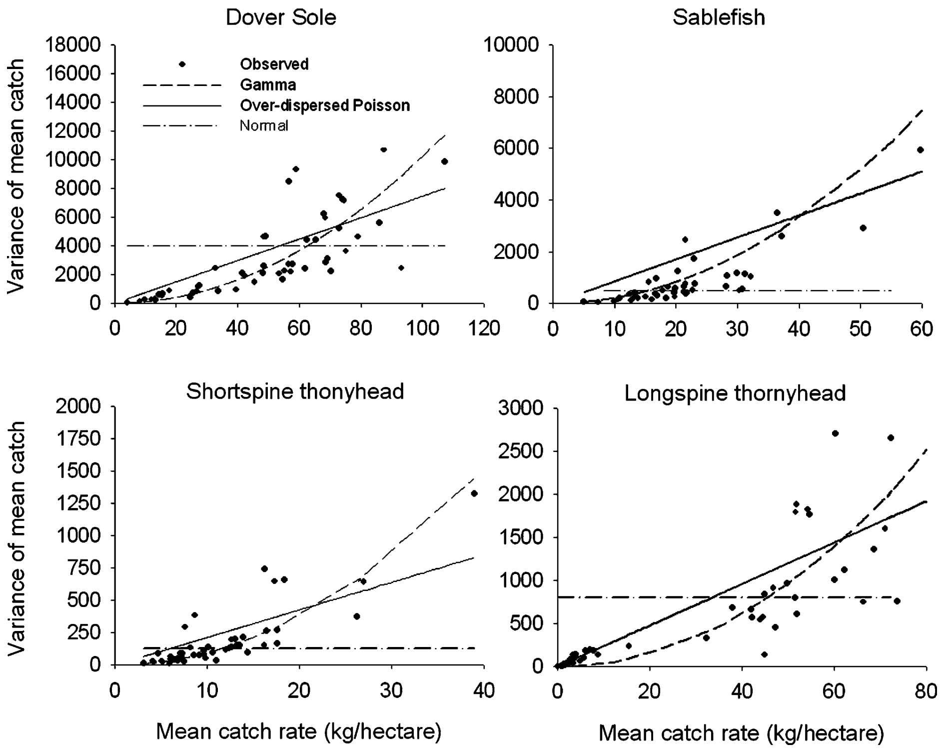

Fig. 4. Variance in catch rate against mean catch rate by species. Mean and variances were calculated for each combination of stratum and year (data aggregated over the AFSC and NWFSC surveys).

where P_i^s is the expected probability that a haul in stratum s during year i leads to a non-zero catch, and V_i^s is the expected catch-rate in stratum s during year i given that the catch-rate is non-zero.

The information needed to apply Eq. (2) is obtained from two separate GLMMs both fitted using restricted maximum likelihood (REML). Whether or not a haul is zero is assumed to arise from a Bernoulli trial and the data are consequently modeled using a binomial error model. However, further investigation is necessary to determine the appropriate error model for the non-zero catch rates. A simple test of candidate error models is to plot the variance of the catch rate against the average catch rate (Dong and Restrepo, 1996). Mean catch rates and their variances were calculated on the arithmetic scale for all strata and years (Fig. 4). The variance would be independent of the mean if the data were normally distributed, which was not evident for any of the four species. Similarly, an over-dispersed Poisson error model, in which the variance in catch rate would fall along a straight line, did not seem to fit the data adequately (Fig. 4). Instead, the data were seemingly well approximated by the assumption that the variance in catch rate is proportional to the square of the mean catch rate. Therefore, all subsequent analyses of positive catch rates were based on the gamma error model. Stefánsson (1996), who applied the delta–gamma distribution to haddock, provides a theoretical treatment of the delta–gamma model, which will not be reiterated in detail here. The probability of a positive haul, P_i^s, and the catch-rate given that the catch is non-zero, V_i^s, for a given stratum and year, is obtained from the inverse link (inverse logit and exponential in this case) of the linear predictors of the 'best model'.

Three alternative models with the same fixed effects structures but different assumptions about the random effects structure are fitted to the data: (1) no random effects, (2) random vessel effect, and (3) random vessel × year interaction. The Akaike Information Criterion (AIC) (Akaike, 1974) can be used to compare these three models (Pinheiro and Bates, 2000).

The process used to compute the point estimates of the catch-rates for each year × spatial cell combination implies that the variances for the densities cannot be computed analytically. Instead, a Monte Carlo approach (based on 10 000 replicates) was used in which replicate draws were taken from a multivariate normal distribution for the estimates of the fixed effects pa-

Table 2
Diagnostic statistics for GLMMs with different random effects structures

Model effect	Binomial error model				Gamma error model			
	AIC	ΔAIC[a]	σ_R[b]	χ[c]	AIC	ΔAIC	σ_R	χ
Sablefish								
Fixed, no random	13588.4	221.2	–	0.9236	7487.1	296.2	–	2.0214
Vessel random	13492.0	124.8	0.1446	0.9321	7307.1	116.2	0.0311	1.7685
Vessel × year random	13367.2		0.1944	0.8964	7190.9		0.0165	1.6622
Dover sole								
Fixed, no random	15391.2	71.0	–	1.407	6038.5	27.4	–	1.1382
Vessel random	15355.4	35.2	0.0896	1.4205	6021.7	10.6	0.0121	1.1249
Vessel × year random	15320.2		0.0834	1.3863	6011.1		0.0205	1.1157
Shortspine								
Fixed, no random	12831.1	–	–	1.0377	6705.1	86.8	–	1.3835
Vessel random	1283.1	–	–	1.0377	6641.3	23.0	0.0081	1.3388
Vessel × year random	12831.1	–	–	1.0377	6618.3		0.0128	1.3360
Longspine								
Fixed, no random	14558.8	156.4	–	0.9552	3835.5	57.2	–	0.6655
Vessel random	14453.7	51.3	0.733	0.9107	3812.8	34.5	0.0108	0.6508
Vessel × year random	14402.4		0.0834	0.8863	3778.3		0.0129	0.6489

[a] $\Delta AIC = AIC_i - AIC_{min}$.
[b] σ_R is the standard deviation of the random effects.
[c] χ is the extra dispersion parameter.

rameters. This idea is consistentwith the fact that the parameters of a generalized linear mixed model are approximately distributed according to a multivariate normal distribution (Littell et al., 1996).

3. Results

3.1. Model diagnostics

Convergence of the GLMMs was achieved for all species. The only peculiarity occurred with the logistic regression model (binomial error model) that predicted the proportion of hauls with non-zero catches of shortspine thornyhead, where the random effects components were estimated at zero (Table 2). This can occur when the within-group variance is several orders of magnitude greater than between-group variance. This may, in fact, be the case here because a very large fraction of tows caught some shortspine thornyhead (Fig. 5), so that there is limited scope for between-vessel variation in the probability of a tow making a catch. Therefore, further analysis of the probability of catching some shortspine thornyhead was based on a fixed effects model.

The GLMM incorporating the random vessel × year effect had the smallest ΔAIC (Burnham and Anderson, 2002) and was thus selected as the best model. This result was consistent across all GLMMs, with little support for other models (ΔAIC > 100).

The extra dispersion parameter, χ (Table 2), provides a measure of the degree to which the size of the deviance corresponds to that expected given the error model. The value of χ decreases as model complexity increases from fixed effects only to random vessel and random vessel × year effects. It ranges between 0.66 and 1.77 for models with a random vessel × year interaction, which indicates little evidence of over- or under-dispersion. Furthermore, the goodness-of-fit diagnostics indicate that the model fits the data reasonably well for all four species (Figs. 5 and 6). Most of the observations occur at probabilities between 0.7 and 1.0 for shortspine thornyhead and sablefish, and the model replicates this well (Fig. 5). How-

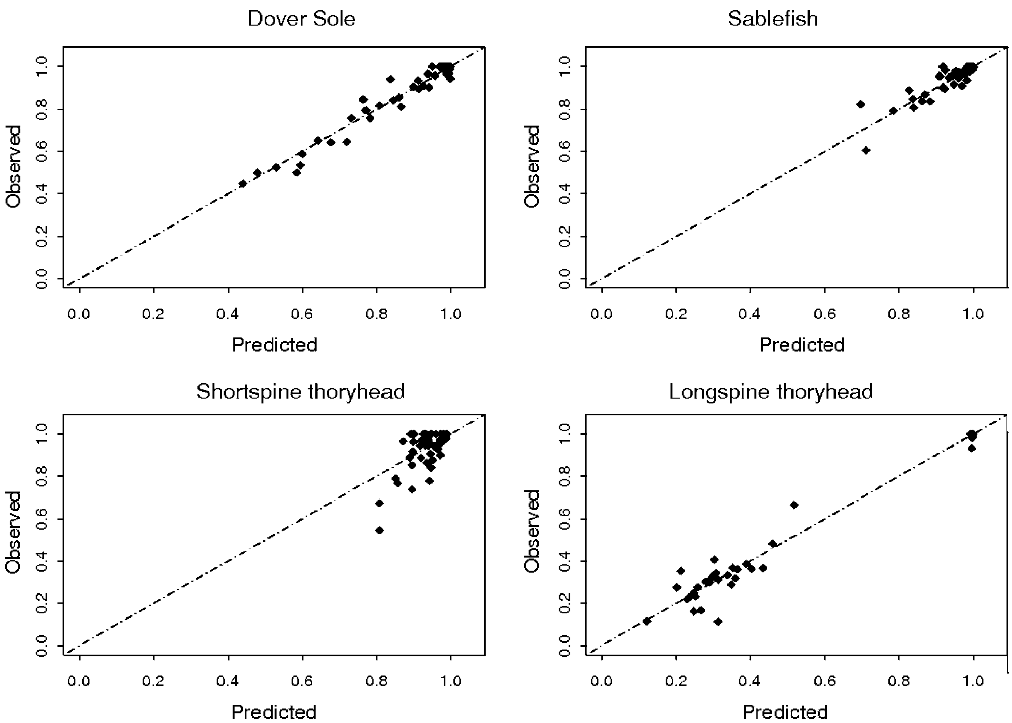

Fig. 5. Observed and predicted probabilities of non-zero tows for the four species. The predictions are based on models with a random vessel × year effect.

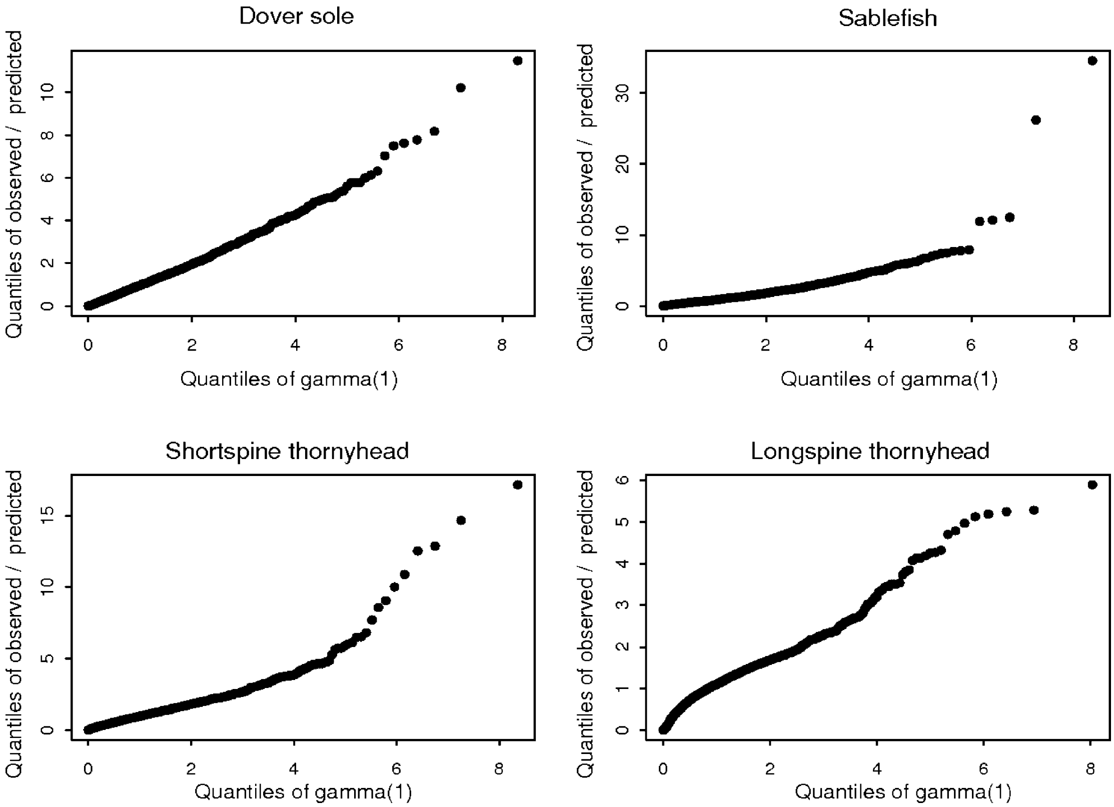

Fig. 6. Quantile–quantile plots of the ‘residuals’ for the four species. The predictions are based on models with a random vessel × year effect.

ever, the dispersion of the observed-predicted pairs off the diagonal indicates some lack of fit for shortspine thornyhead. For Dover sole and longspine thornyhead, numerous strata have probabilities of non-zero catches of <0.6, and the model mimics this adequately (Fig. 5).

The goodness-of-fit of the GLMM to the catch rates for the non-zero hauls is summarized in Fig. 6 in the form of quantile–quantile (*q*–*q*) plots. The ‘residuals’ (the *y*-axis in the *q*–*q* plots) are the ratios of the observed to the expected catch-rates and the values on the *x*-axis arise from a gamma distribution with a mean of unity. The fit of the model to the non-zero catch rates can be evaluated by assessing the degree to which the points in Fig. 6 fall along a straight line. For all four species, there are outlying catches, as reflected by a few large residuals (e.g. observed/predicted ≫ 1). However, there are usually only 5–10 outliers in Fig. 6 for each species – a relatively small fraction of the overall sample size of $n > 2000$. The deviance residuals plotted against the gamma error model’s linear predictors also indicate reasonably good goodness-of-fit of the models to the non-zero catch rate data for the four species (Fig. 7). The deviance residuals for shortspine thornyhead are separated into two distinct groups. This is attributed to the depth factor in the model and suggests that catch rates in weight are substantially greater in the deeper waters (>600 m) of the continental slope, where larger shortspine thornyheads are found. Despite this pattern, deviances are well distributed both above and below the zero reference line.

3.2. Comparison of vessel effects

Plots (showing the REML expected value and 95% confidence intervals) of the random vessel × year effect coefficients (Fig. 8) provide insight into how vessel power changes with vessel and time. The random effect coefficients for the *FRV Miller Freeman* (MF in Fig. 8) did not differ substantially from those of the industry-

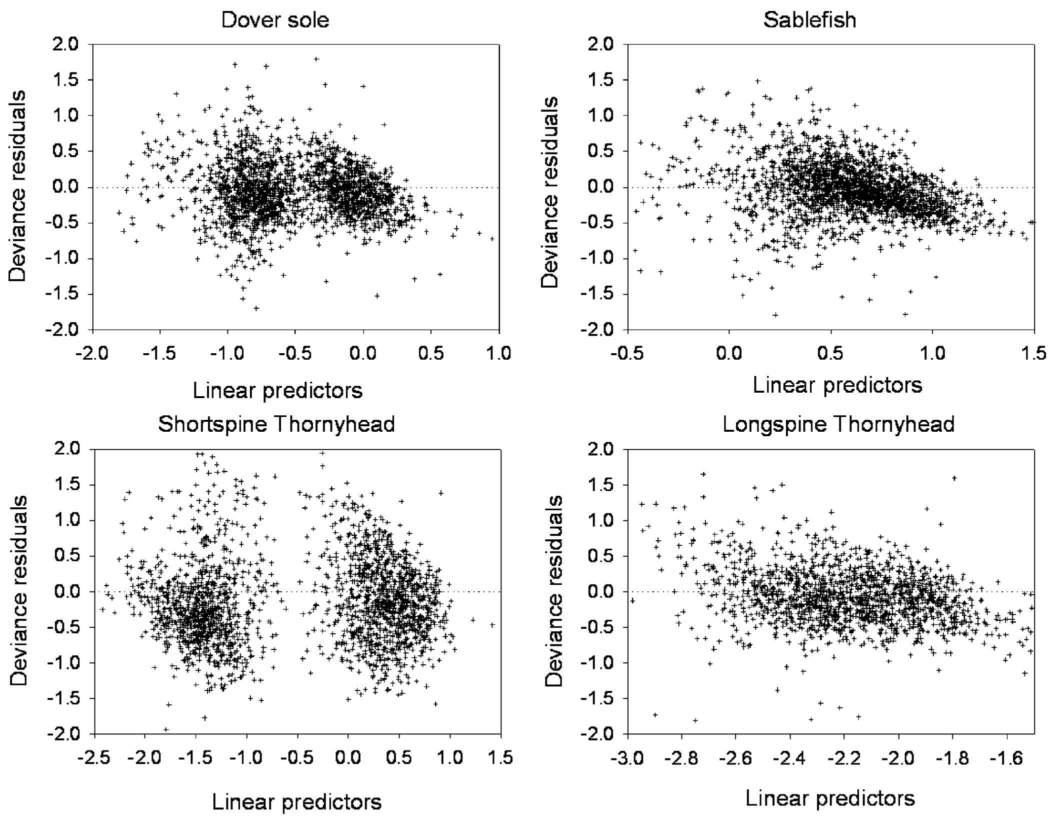

Fig. 7. Deviance residuals for the non-zero catches vs. the corresponding predictions from the generalized linear model for the four species. The predictions are based on models with a random vessel × year effect.

based vessels (Fig. 8). For sablefish, for example, the *FRV Miller Freeman* exhibited a slightly greater probability of catching sablefish, as the random effect coefficients for this vessel generally lie above the reference line of zero in 2001 and 2002 (Fig. 8a, upper panel). This may not be completely unexpected because of the greater average effort expended by this vessel and the use of effort as an offset in the GLMM. However, given that the catch rate is positive, the *FRV Miller Freeman* appears to have catch rates that are largely consistent with the industry-based vessels (Fig. 8a, lower panel). This pattern of generally consistent random vessel coefficients for the positive catch rate models associated with the *FRV Miller Freeman* and industry-based vessels is consistent across the other species (Fig. 8b–c).

Interestingly, the random effect coefficients associated with either the *FRV Miller Freeman* or industry-based vessels are not very consistent over time (Fig. 8). Vessel #1, which has been involved in the survey for all of the 5 years, illustrates this point – the random effect coefficients lie both above and below the mean zero reference line; e.g. greatest fishing power in 1999 and 2000 for the positive catch rates of sablefish but less than or equal to zero in 1998, 2001, and 2002. These results suggest that the random vessel coefficients for both the *FRV Miller Freeman* and industry-vessels come from a common distribution of random effects and, as such, that it is valid to treat the catch data from these different surveys together when developing abundance indices.

3.3. Biomass indices

Given the results in Fig. 8, the biomass indices (Fig. 9) are based on the data for both the *FRV Miller*

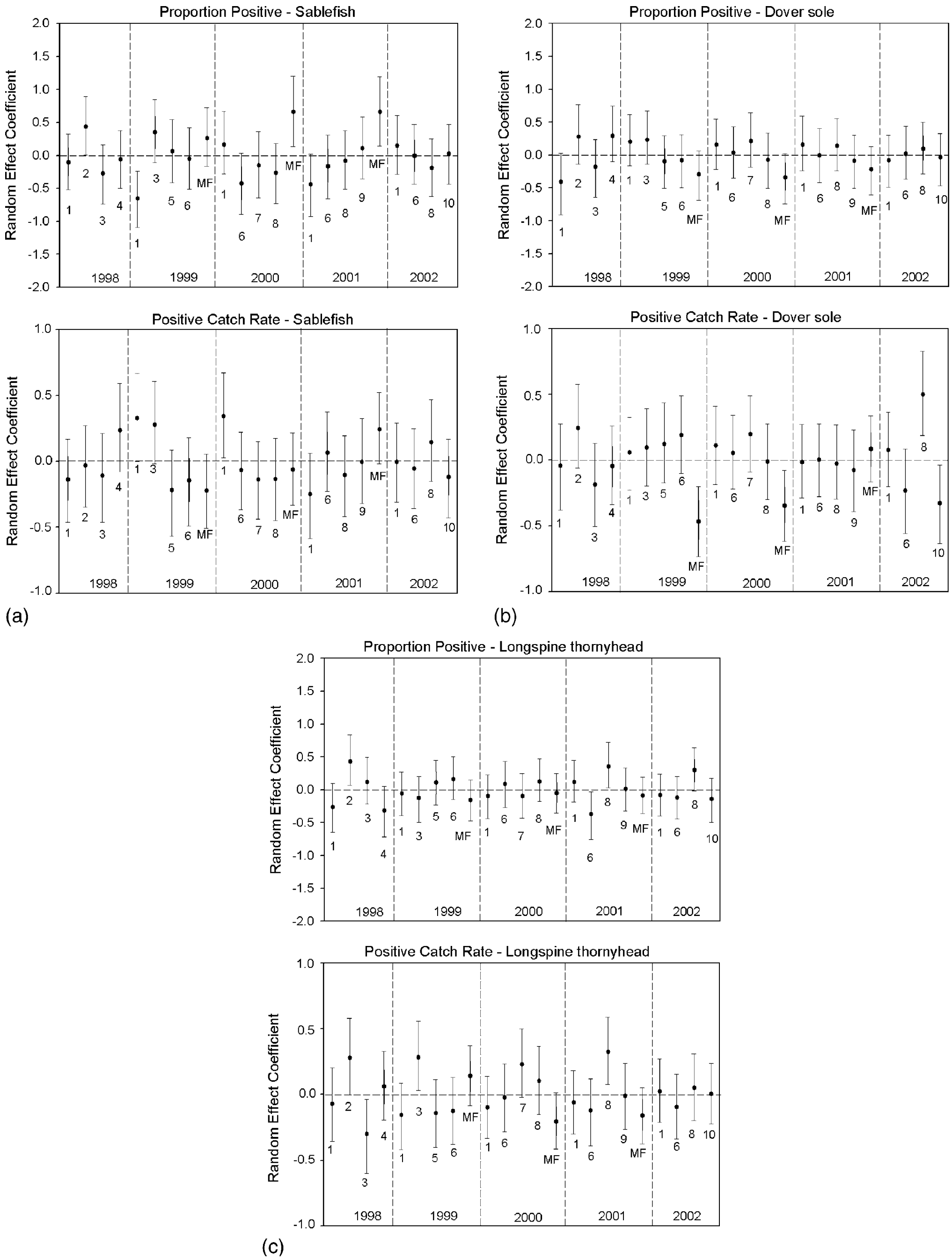

Fig. 8. Maximum likelihood point estimates of the random effects coefficients (year and vessel) from generalized linear mixed models applied to survey catch data for sablefish, Dover sole and longspine thornyhead. Error bars show two standard errors of the mean. Results are shown in the upper panels for the fit to whether a catch is zero or not and in the lower panels to the catch given that the catch is non-zero. 'MF' denotes *FRV Miller Freeman* and all others are industry-based vessels.

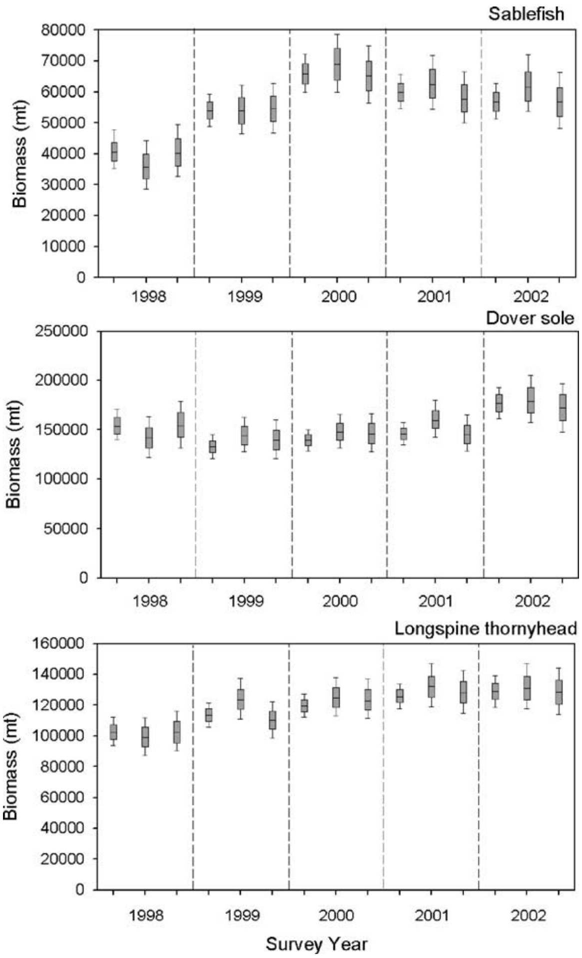

Fig. 9. Box and whisker plots for the biomass indices (1998–2002) for sablefish, Dover sole, and longspine thornyhead, based on data for the industry-based vessels only. The results for each year are (from left to right), estimates based on fixed effects only, a random vessel effect, and a random vessel × year effect.

Freeman and industry-based vessels. The degree of variability about the point estimates of biomass for all species and all years depends on the structure of the GLMMs. Specifically, ignoring differences in vessel power leads to the most precise estimates, while allowing for a random vessel effect and a random vessel × year effect leads to increasingly less precise estimates. These results suggest that ignoring the added variability due to membership of the hauls within vessels may lead to an underestimation of variance.

4. Discussion

Many researchers have analyzed data from fishery resource surveys or fishery CPUE statistics using fixed effects models to standardize for vessel or to derive fishing power corrections from vessel calibration experiments. Specifically, multiplicative models using fixed vessel effects are generally used to standardize commercial catch rates (Gavaris, 1980; Punt et al., 2000). For research trawl data, methods for analyzing fishing power differences between two fishery research vessels have included multiplicative linear models (Sissenwine and Bowman, 1978), as well as ratios of CPUE random variables with assumed identical distributions and different scales (Kappenman, 1992). In either approach, the assumption of vessel as a fixed effect seems justified due to the number of vessels involved (few levels of the factor) in the case of research surveys or if the intended purpose is to standardize catch and effort data from a relatively large and diverse fleet of vessels. In contrast, treating vessel as a random effect seems justified in this case from both theoretical and practical considerations.

From a theoretical perspective, a factor should be considered random if the levels of that factor represent only a random sample of a larger set of potential levels. The multi-vessel survey, in which vessels are added to and omitted from year to year due to an open bid system, essentially mimics the concept that the vessels represent a random sample from a larger set of potential vessels (fleet). This assumption remains valid only if the pool of vessels remains sufficiently large. On the practical side, differences in fishing power appeared to differ only nominally among vessels (including differences between the *FRV Miller Freeman* and industry-based vessels) and over years for a given vessel. Furthermore, the performance of a vessel appears to differ among years, implying that, for all practical purposes, even if the same vessel participates in the survey in several years, this is equivalent to different vessels participating in each year. This observation, in particular, cannot be dealt with in a model in which vessel is treated as a fixed effect, because the year-factor would be totally confounded with the vessel-factor. As such, using a mixed model and AIC to explore different specifications of random effects within the model seems a reasonable approach. In this context, the different models of random effects were essentially nested and AIC or likelihood ratio may be equivalently used. It should be pointed out that AIC is typically used in the case of non-nested models (Hilborn and Mangel, 1997).

The open bid vessel charter protocol used to contract vessels in this survey was largely based on the results of McAllister (1995). He analyzed data from the Eastern Bering Sea multi-vessel fishery survey and found that substantial gains in precision can be obtained by contracting several vessels by charter on an annual open bid system. Specifically, the variability in catch rates among vessels can be reduced by increasing the number of vessels used to survey a resource. This, in turn, should reduce the variance of the annual biomass indices (McAllister, 1995; McAllister and Pikitch, 1997). The use of four or more vessels chartered on an open bid can reduce the coefficient of variation of the biomass indices even when catching power varies substantially among vessels. Large within-vessel variability in catch-rates, and virtually negligible vessel effects, and the use of four vessels per year on the open bid system are largely consistent with McAllister (1995) suggestions.

The delta-distribution was used to describe separate but related characteristics of the response variable. Separate GLMMs were used to model the probability of a positive haul (binomial) and the catch rate given a positive haul (gamma). Pennington (1983, 1991) originally proposed that the delta approach in which a log-normal distribution was used to analyze the positive catch rates, but others have applied the gamma model (Vignaux, 1996; Stefánsson, 1996). There has been some controversy about the form of the distribution to use when modeling the positive catch rates. Syrjala (2000) and Myers and Pepin (1990) concluded that assuming a log-normal error structure for the non-zero catch rates when applying a delta distribution approach was not as robust as had been suggested by Pennington (1991). While simulation studies may be needed to test the robustness of the procedure adopted in this paper to the assumed distribution for the positive tows, the gamma distribution seemed most appropriate for the data in this study. The normal distribution for the logarithms of the non-zero catch rates for the four species was rejected in favor of a heavy-tailed distribution using the Shapiro–Wilk test (Shapiro and Wilk, 1965); the skewness of the non-zero catch rates was more consistent with that expected under a gamma distribution. Furthermore, the variance in catch rate is proportional to the square of the mean catch rate, a result consistent with a gamma distributed random variable.

The differences between the random effect coefficients associated with the *FRV Miller Freeman* and industry-based vessels suggest that the catch data from these two sources may be pooled. In essence, the vessel coefficients for the *FRV Miller Freeman* appear to arise from largely the same common distribution of random effects as the random vessel coefficients associated with the industry vessels. As such, catch rates from both surveys may be combined as a single survey for developing abundance indices for the species analyzed in this paper. It should be cautioned, however, that catch data be examined on a species-by-species basis before reaching the aforementioned conclusion. Finally, analysis of the catch rate data using the GLMM approach described in this paper seems to be a reasonable approach for a multi-vessel survey. In so doing, vessel or vessel $\times$ year should be treated as a random effect since, by ignoring this, the variances associated with biomass indices may be negatively biased.

Acknowledgments

The authors thank the survey teams of the FRAM Division, NWFSC, and the RACE Division, AFSC, for making available the survey data upon which the analyses were based. Thanks are also extended to Ray Hilborn and Dave Sampson for useful comments during the early phase of this work. We are grateful to Russel Millar, John Wallace (who also assisted in some data analysis) and Michael Pennington for thoughtful reviews of the manuscript. Financial support for AEP was provided through NMFS Grant No. NA07FE0473.

References

Aitchison, J., Brown, J.A.C., 1957. The Lognormal Distribution. Cambridge University Press, Cambridge.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Control AC-19, 716–723.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference, A Practical Information-Theoretic Approach, 2nd ed. Springer-Verlag, New York.

Byrne, C.J., Azarovitz, T.R., Sissenwine, M.P., 1981. Factors affecting variability of research vessel trawl surveys. In: Doubleday, W.G., Rivard, D. (Eds.), Bottom trawl surveys. Spec. Publ. Can. J. Fish. Aquat. Sci. 58, 258–273.

Cotter, A.J.R., 2001. Intercalibration of North Sea international bottom trawl surveys by fitting year–class curves. ICES J. Mar. Sci. 58, 622–632.

Dong, Q., Restrepo, V.R., 1996. Notes on the Poisson error assumption made to estimate relative abundance of west Atlantic bluefin tuna. ICCAT Col. Vol. Sci. Pap. 45, 125–131.

Fiorentini, L., Dremiere, P.Y., Leonori, I., Sala, A., Palumbo, V., 1999. Efficiency of the bottom trawl used for the Mediterranean international trawl survey (MEDITS). Aquat. Living Resour. 12, 187–205.

Gavaris, S., 1980. Use of a multiplicative model to estimate catch rate and effort from commercial data. Can. J. Fish. Aquat. Sci. 37, 2272–2275.

Gunderson, D.R., 1993. Surveys of Fisheries Resources. Wiley, New York.

Hilborn, R., Mangel, M., 1997. The Ecological Detective: Confronting Models with Data, Monogr. Pop. Biol. Princeton University Press, Princeton.

Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice Dynamics and Unceratinty. Chapman & Hall, London.

Jacobson, L.D., Hunter, J.R., 1993. Bathymetric demography and management of Dover sole. N. Am. J. Fish. Manage. 13, 405–420.

Jacobson, L.D., Vetter, R.D., 1996. Bathymetric demography and niche separation of thornyhead rockfish: *Sebasetolobus alascanus* and *S. altivelis*. Can. J. Fish. Aquat. Sci. 53, 600–609.

Kappenman, R.F., 1992. Robust estimation of the ratio of scale parameters for positive random variables. Commun. Stat., Theor. Meth. 21, 2983–2996.

Lauth, R.R., Syrjala, S.E., McEntire, S.W., 1998. Effects of gear modifications on the trawl performance and catching efficiency of the West Coast upper continental slope groundfish survey trawl. Mar. Fish. Rev. 60, 1–26.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., 1996. SAS System for Mixed Models. SAS Institute Inc, Cary, NC.

McAllister, M., 1995. Using decision analysis to choose a design of surveying fisheries resources. Ph.D. Dissertation. University of Washington, Seattle, WA.

McAllister, M.K., Pikitch, E.K., 1997. A Bayesian approach to choosing a design for surveying fisheries resources: application to the eastern Bering Sea trawl survey. Can. J. Fish. Aquat. Sci. 54, 301–311.

McCullaugh, P., Nelder, J.A., 1989. Generalized Linear Models, 2nd ed. Chapman & Hall, New York.

Methot, R.D., Wallace, J.R., West, C.W., 2000. Introducing a new trawl survey for the U.S. West Coast slope groundfish. ICES Document C.M. 2000/W:11, p. 15.

Myers, R.A., Pepin, P., 1990. The robustness of lognormally-based estimators of abundance. Biometrics 46, 1185–1192.

Pennington, M., 1983. Efficient estimators of abundance for fish and plankton surveys. Biometrics 39, 281–286.

Pennington, M., 1991. On testing the robustness of lognormal-based estimators. Biometrics 47, 1623–1624.

Pinheiro, J.C., Bates, D.M., 2000. Mixed-Effects Models in S and S-Plus. Springer, New York.

Punt, A.E., Walker, T.I., Taylor, B.L., Pribac, F., 2000. Standardization of catch and effort data in a spatially structured shark fishery. Fish. Res. 43, 129–145.

Ramm, D.C., Xiao, Y., 1995. Herding of groundfish and effective pathwidth of trawls. Fish. Res. 24, 243–259.

Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52, 591–611.

Sissenwine, M.P., Bowman, E.W., 1978. An analysis of some factors affecting the catchability of fish by bottom trawls. ICNAF Res. Bull. 13, 81–87.

Stefánsson, G., 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES J. Mar. Sci. 53, 577–588.

Syrjala, S.E., 2000. Critique on the use of the delta distribution for the analysis of trawl survey data. ICES J. Mar. Sci. 57, 831–842.

Turk, T.A., Builder, T., West, C.W., Kamikawa, D.J., Wallace, J.R., Methot, R.D., 2001. The 1998 Northwest Fisheries Science Center Pacific west coast upper continental slope trawl survey of groundfish resources of Washington, Oregon, and California: estimates of distribution, abundance, and length composition. NOAA Tech. Memo. NMFS-NWFSC-50, 122 pp.

Vignaux, M., 1996. Analysis of spatial structure in fish distribution using commercial catch and effort data from the New Zealand hoki fishery. Can. J. Fish. Aquat. Sci. 53, 963–973.

Wilkins, M.E., Zimmerman, M., Weinberg, K.L., 1998. Pacific West Coast bottom trawl survey of groundfish resources: estimation of distribution, abundance, and length and age composition. NOAA Tech. Memo. NMFS-AFSC-89, 138 pp.

Wolfinger, R., O'Connell, M., 1993. Generalized linear mixed models: a pseudo-likelihood approach. J. Statist. Comp. Sim. 48, 233–243.