Diurnal variation in bottom trawl survey catches: does it pay to adjust?

Vidar Hjellvik, Olav Rune Godø, and Dag Tjøstheim

Abstract: The diurnal bias of bottom trawl survey catches is studied with the purpose of adjusting for it and thereby improving the accuracy of abundance estimates. The correction term is estimated with uncertainty and thus increases the variance of the resulting abundance estimate. To investigate this adequately, we use a stochastic model describing diurnal fluctuations and examine the annual variation of the diurnal amplitude as a function of species and length. The diurnal amplitude is fairly stable for large fish, and for these, the bias-corrected estimate leads to a moderate increase in variance. For small fish, the diurnal amplitude is unstable, however, and the correction of diurnal bias occurs at the expense of a large increase in variance. This unstable amplitude also leads to a large year-to-year variation in catchability for small fish. For haddock (*Melanogrammus aeglefinus*), the diurnal amplitude depends heavily on fish length, indicating a strong decrease in catchability with decreasing fish length.

Résumé : Nous avons étudié l'erreur systématique reliée à la période de la journée dans les inventaires faits à partir de récoltes au chalut de fond dans le but de trouver un ajustement et ainsi améliorer les estimations d'abondance. L'estimation du terme de correction comporte de l'incertitude, ce qui augmente la variance des estimations d'abondance. Un modèle stochastique qui décrit les fluctuations diurnes et l'examen de la variation annuelle de l'amplitude diurne en fonction des espèces et des longueurs nous ont permis d'étudier ce problème adéquatement. L'amplitude diurne est assez stable chez les gros poissons et, pour eux, une estimation corrigée entraîne une augmentation modeste de la variance. Pour les petits poissons, cependant, l'amplitude diurne est instable et toute correction de l'erreur amène une importante augmentation de la variance. Cette amplitude instable cause aussi une grande variation inter-annuelle de la capturabilité des petits poissons. Chez l'Aiglefin (*Melanogrammus aeglefinus*), l'amplitude diurne est fortement reliée à la longueur des poissons, avec une forte décroissance de la capturabilité en fonction de la diminution de la longueur du poisson.

[Traduit par la Rédaction]

Introduction

A main motivation for studying diurnal oscillations, besides understanding them biologically, is to obtain more accurate abundance estimates. Efficiency of trawl surveys is dependent on the ability of the gear to catch the available fish and further, the availability of target species to the trawl (Godø 1994). Diurnal fluctuations may affect both of these factors, and if not adjusted for, diurnal oscillations could cause a large bias. Implicitly, the bias is included in the catchability coefficient when converting survey catches to abundance estimates. This coefficient is usually considered constant from one year to another but is well known to be subject to considerable variation (e.g., Pennington and Godø 1995). If the impacts of the diurnal fluctuations vary from year to year, this may serve to explain temporal variation in the catchability coefficient. A main goal of our paper is to examine the effect of diurnal bias on catchability, and its implications for survey abundance estimation.

To avoid diurnal bias, surveys are often conducted during the day only, under the assumption that the highest catches (usually during daytime) give the most representative picture of the stock (Wakabayashi et al. 1985). This leads to at least three problems. First, for some species nighttime catches are higher than daytime catches (Walsh 1988; Casey and Myers 1998). Second, it is not quite clear how daytime catches should be defined. As shown in Hjellvik et al. (2001), simply using the times of sunrise and sunset is not optimal in general (cf. also the problem of polar night in the Barents Sea). Third, there is a feeling that only using daytime measurements is tantamount to throwing away valuable information and not exploiting the measurements in an adequate way. Survey vessel time is expensive and optimal use of available resources is crucial for the quality of the survey results (Pennington and Vølstad 1991).

One might expect that including nighttime measurements will, if properly adjusted for diurnal bias, lead to more accurate abundance estimates. To our knowledge, this hypothesis has not been carefully examined. An important reason for this is the lack of an appropriate way of handling the associated uncertainty. A major objective of this paper is to show that the modeling technique of Hjellvik et al. (2001) can be used to approach this problem. The answer will depend critically on the annual stability of the diurnal variations. The bias adjustment depends on an uncertain estimate of the

V. Hjellvik[1] **and O.R. Godø.** Institute of Marine Research, P.O. Box 1870, Nordnes, N-5817 Bergen, Norway.
D. Tjøstheim. Department of Mathematics, University of Bergen, Johs. Brunsgt 12, N-5008 Bergen, Norway.

[1]Corresponding author (e-mail: vidarh@imr.no).

34

Can. J. Fish. Aquat. Sci. Vol. 59, 2002

bias, and this added uncertainty is more serious if the diurnal fluctuations vary significantly from one year to another, because the bias then must be estimated separately each year and the added uncertainty may cancel the benefit of including nighttime measurements. Demer and Hewitt (1995) applied a temporal compensation function to adjust upward acoustic biomass estimates of Antarctic krill (*Euphausia superba*). In their modeling, they did not take into account the uncertainty involved in estimating the compensation function, but they found that the coefficient of variation in most cases increased after adjusting for diurnal variation, contrary to their expectation.

Demonstration of annual changes in diurnal variation in trawl catches is simultaneously a manifestation of varying catchability. Trawl survey catches are often one of the basic sources of information used in modeling ecosystem interactions and competition both between sizes of the same species and between different species. If catchability varies substantially in time and space, the traditional approach where catchability is considered constant (Godø 1994) might lead to misinterpretation of the dynamics of the ecosystem and hence reduce the quality of advanced modeling in stock assessment. In this paper, we demonstrate that it is possible to evaluate quantitatively changes in catchability based on the accumulated information collected during the surveys.

Material and methods

The data

Combined acoustic and bottom-trawl surveys for demersal fish have been conducted annually in the Barents Sea during winter (January–March) and in autumn (August–September) by The Institute of Marine Research, Bergen. We use the data from 1985 to 2000. For the winter survey, the time span does not vary much during the period (Fig. 1), but there have been some changes in the area covered. In 1993, the survey area was expanded northeastwards because it was clear by then that the small fish were not completely covered in the traditional survey area. In 1997 and 1998, the vessels were not allowed to enter the Russian zone, and thus the eastern part of the area was not covered. In 1999, the same area remained partially uncovered because of ice conditions. For the autumn survey, the time span varies considerably from year to year, and in 1995, the area covered was substantially expanded to the southeast in an attempt to cover the whole stock. Various adjustments in equipment have also taken place (Jakobsen et al. 1997). Two of the most important adjustments were the introduction of rockhopper gear in 1989 and a reduction in cod-end mesh size in 1994, and both strongly influenced the catch of small fish (<20 cm).

We started out with the length groups (length in centimetres) 0–10, 11–15, 16–22, 23–31, 32–44, 45–63, 64–90, and 90+, which are roughly the same as in Korsbrekke and Nakken (1999). However, we found it convenient to join the two first groups to a 0- to 15-cm group and the three last ones to a group of 45+ cm. The five resulting length groups roughly correspond to ages 1 through 4 and 5+ for cod.

We have considered six species: cod (*Gadus morhua*), haddock (*Melanogrammus aeglefinus*), deep-sea redfish (*Sebastes mentella*), Norway haddock (*Sebastes marinus*), Greenland

halibut (*Reinhardtius hippoglossoides*), and long rough dab (*Hippoglossoides platessoides*). The number of species is much smaller than in Casey and Myers (1998) and the geographical region is more restricted, but the purpose of our study is rather different with the emphasis on adjusting for diurnal effects and the uncertainty involved in the adjustment.

Models

Technical details on models and methods are given in Hjellvik et al. (2001) and in Appendix A. Here we only give a brief summary. The basic observations are the number of fish n_i caught in the ith haul. However, the model treats log-transformed data

$$y_i = \log(n_i \cdot l_i^{-1} + 1) \doteq \log(x_i)$$

where l_i is the towed distance. We assume that the total variation in fish density is caused by a day-to-day variation, a superimposed diurnal variation, and random noise, so that

(1) $y_i = \mu_{d(i)} + f(t_i) + \varepsilon_i$

Here, t_i is the apparent local time of the ith haul (cf. Appendix A), f is a function describing the diurnal variation, and $\mu_{d(i)}$ is the daytime catch level on day $d(i)$ when haul i is taken. Finally, ε_i represents the random noise component.

The function f can be estimated nonparametrically without making any assumptions on its shape, or we can assume that it is given by a known parametric function for which parameters must be estimated.

We have found two parametric functions to be especially useful: the sinusoid

(2) $f_S(t) = f_S(t; D) = (D/2)\sin((t - 6)\pi/12) - (D/2),$
$$0 \le t < 24$$

where $D/2$ is the amplitude; and the symmetric logistic function

(3) $f_L(t) =$

$$f_L(t; D, \alpha, \beta) = \begin{cases} \dfrac{De^{\alpha(t-\beta)}}{1 + e^{\alpha(t-\beta)}} - \dfrac{De^{\alpha(12-\beta)}}{1 + e^{\alpha(12-\beta)}}, & 0 \le t \le 12 \\[3mm] \dfrac{De^{\alpha(24-t-\beta)}}{1 + e^{\alpha(24-t-\beta)}} - \dfrac{De^{\alpha(12-\beta)}}{1 + e^{\alpha(12-\beta)}}, & 12 < t < 24 \end{cases}$$

Both functions have been normalized so that their maximum is at $f(12) = 0$, which means that $\mu_{d(i)}$ can be interpreted as the expected value of y_i at day $d(i)$ at noon. The sinusoid function f_S represents a situation with smooth diurnal variations, whereas the logistic function f_L describes a scenario with an approximately constant night level, another constant day level, and a transition phase between them. The length and location of the transition phase are determined by the parameters α and β, respectively.

Note also that both eqs. 2 and 3 are linear in D. Thus, if α and β are kept fixed in eq. 3, then eq. 1 with $f = f_L$ is a linear model (see Appendix A for details). The main advantage of the linear approximation is that it is much faster, especially when bootstrapping (see the section Uncertainty estimates). If fixed values of α and β are used, we take $\alpha = 2$ (corre-

Fig. 1. Geographical extent and time span of the winter and autumn surveys. All stations were within the area or time indicated in the figure, and 80% were within the intervals indicated by solid lines. The medians are also given. Horizontal dotted lines indicate a change of month.
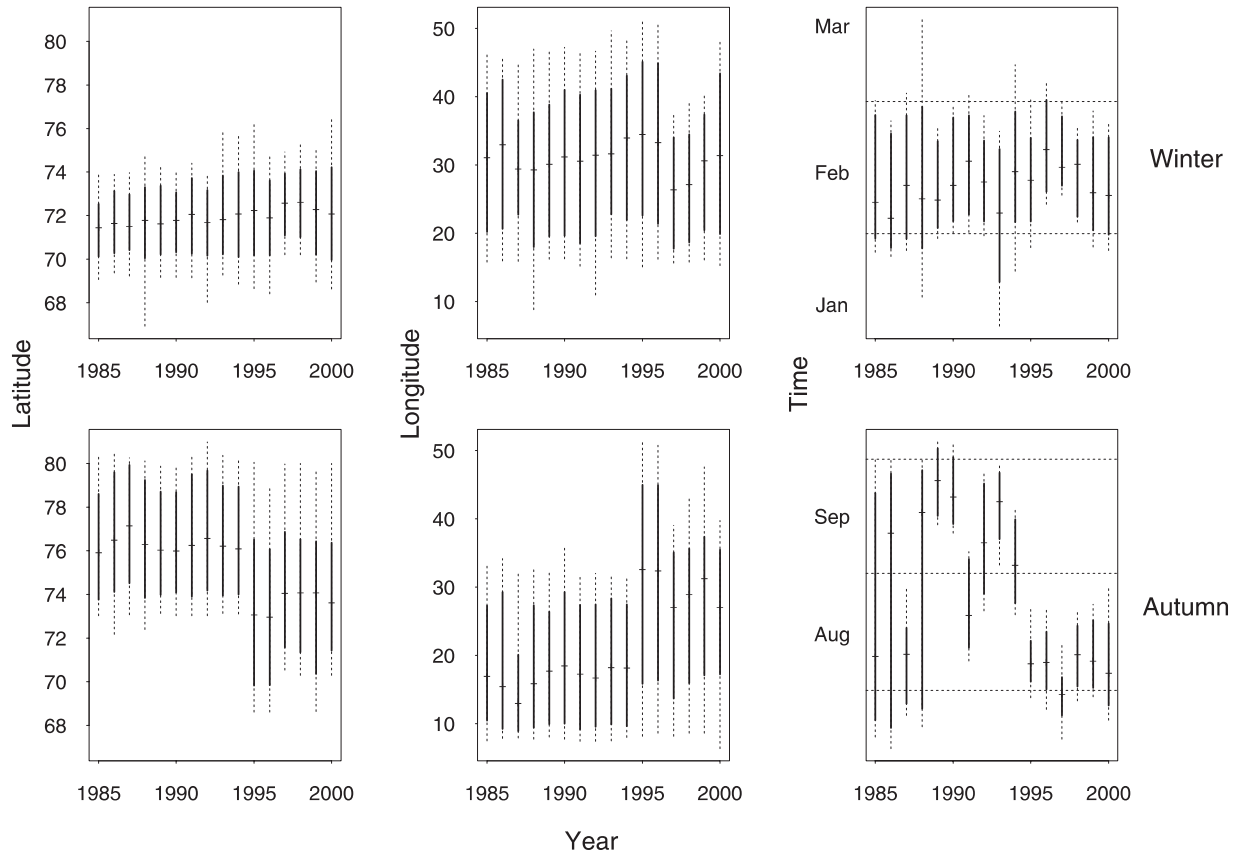
sponding to a transition phase of approximately 3 h) and β between 7 and 8. These values are based on our experience in Hjellvik et al. (2001) and on the computer runs in this paper.

The S-PLUS functions "nls" and "lm" (cf. Venables and Ripley 1997, sections 6 and 9) were used for estimating the parameters in the nonlinear and linear cases, respectively. The standard errors of the parameter estimates from these functions have been used to calculate confidence intervals.

Zero catches

In any study of diurnal variation, it is important to look carefully at zero catches. In Hjellvik et al. (2001), where the data were restricted to cod, it was found that the zero catches were roughly evenly distributed throughout the 24-h cycle and were therefore omitted. But clearly an uneven distribution of zero catches could be an added indication of diurnal effects, and omitting them would lead to an underestimation of the diurnal oscillations. Including all zero catches, as seems to be the case in some earlier studies, would also lead to an underestimation, because getting zero catches may simply mean that there are no fish of various length groups and species in certain locations. If included, the catches from such areas would contribute to the estimate of D just as much as do non-zero catches. Thus, if only zero catches are taken on a particular day, that day would draw the estimate of the diurnal amplitude D downwards. The same argument is valid for the simpler night–day level mod-

els used earlier. The zero catches taken outside the distribution area of a given length group and species should be excluded, and only those. In practice, this is not easy to achieve, but one possible strategy is to exclude all catches, zero and non-zero, taken on days with more than a certain percentage of zero catches, and include all other catches. Varying the threshold percentage yields a sensitivity study as to what degree the zero catches influence the results.

Adjusting for diurnal variation

When a functional relationship such as eq. 2 or eq. 3 is established, the catches can be adjusted accordingly to correct for diurnal bias. In a situation where f (f_L or f_S) gives a good fit and f has its maximum $f(t) = 0$ at $t = 12$, the natural adjustment of a measurement y_i taken at local time t_i is

(4) $y_{i,\mathrm{adj}} = y_i - f(t_i)$

where in practice f has to be replaced by its estimate $\hat{f}(t; D, \alpha, \beta) = f(t; \hat{D}, \hat{\alpha}, \hat{\beta})$.

We are ultimately interested in the influence of the adjustment on the abundance indices, but here we simplify by looking at the bias and uncertainty of the means $\bar{y}$ and $\bar{x}$. Alternatively, the medians $m(y)$ and $m(x)$ could be used. For the nonlogarithmic data, $m(x)$ is much more robust than $\bar{x}$ in terms of coefficient of variation, but smaller (if $y \sim N(\mu, \sigma^2)$ and $x = \exp(y)$, then $\mathrm{mean}(x)/\mathrm{median}(x) = \exp(\sigma^2/2)$). We have followed the traditional approach by using the mean.

Uncertainty estimates

Calculating the uncertainty of abundance indices is not straightforward because of, among other things, the uneven geographical distribution and the movement of the fish. Ideally, the survey should be repeated under identical conditions. If such hypothetical surveys could be conducted, then for the kth survey we would have

$$(5) \qquad y_i^{(k)} = \mu_{d(i)} + f(t_i) + \varepsilon_i^{(k)}$$

with $\mu_{d(i)}$ and $f(t_i)$ being fixed because of the identical circumstances of the surveys. The mean and its uncertainty could then be computed by looking at $y_i^{(k)}$, $k = 1,2,\dots$.

This is impossible in practice, but we can mimic the above situation by using the bootstrap technique. First we estimate $\mu_{d(i)}$ and $f(t_i)$ from the data from a particular survey to obtain $\hat{\mu}_{d(i)}$ and $\hat{f}(t_i)$. Then we calculate the estimated residuals

$$(6) \qquad \hat{\varepsilon}_i = y_i - \hat{\mu}_{d(i)} - \hat{f}(t_i)$$

The estimated random variation is now contained in $\{\hat{\varepsilon}_i\}$, and we can create new realisations $\hat{\varepsilon}_i^{(k)}$, or $\varepsilon_i^{*(k)}$ in bootstrap notation, by drawing random samples, with replacement, from $\{\hat{\varepsilon}_i\}$, creating bootstrap realisations

$$(7) \qquad y_i^{*(k)} = \hat{\mu}_{d(i)} + \hat{f}(t_i) + \varepsilon_i^{*(k)}$$

where $\hat{\mu}_{d(i)}$ and $\hat{f}(t_i)$ are kept fixed in accordance with eq. 5. For each realisation $\{y_i^{*(k)}\}$, we can estimate the mean $\bar{y}^{*(k)}$, and bootstrap confidence intervals can be formed from the empirical distribution of $\{\bar{y}^{*(k)}\}$.

To examine the effect of bias correction, we first go back to the idealized situation (eq. 5). Now f and $\mu_{d(i)}$ are unknown quantities that need to be estimated in order to correct for diurnal bias. We can use the nls/lm algorithm to obtain $\hat{f}^{(k)}$ (and $\mu_{d(i)}^{(k)}$) for each realisation and then compute the adjustment

$$y_{i,\text{adj}}^{(k)} = y_i^{(k)} - \hat{f}^{(k)}(t_i)$$

and finally the mean and its uncertainty is found from $\{y_{i,\text{adj}}^{(k)}\}$. When employing the bootstrap, we use re-estimates $f^{*(k)}(t_i)$ (and $\hat{\mu}_{d(i)}^{*(k)}$) for each bootstrap realisation $\{y_i^{*(k)}\}$, with $y_i^{*(k)}$ given by eq. 7, and calculate

$$(8) \qquad y_{i,\text{adj}}^{*(k)} = y_i^{*(k)} - f^{*(k)}(t_i)$$

The fact that we have to estimate f means that although we correct for the bias, the added variability of the estimated f leads to larger variability in $y_{i,\text{adj}}$ than in y_i in eq. 4, which again leads to larger uncertainty in the adjusted abundance estimate. If f varies from one year to another, it has to be estimated separately for each year, which implies that there are fewer data points available for this estimation. This situation is represented by the bootstrap estimate in eq. 8, with $\{y_i\}$ being equal to the observed data points in each particular year. The effect on uncertainty is illustrated by a simple example given in Appendix A.

If, however, f is constant (i.e., the parameters α, β, and D are the same) from year to year, a more accurate estimate of f is obtained by pooling the data for the various years. As the number of available years increases, f can be considered

to be known compared to the other unknown parameters $\mu_{d(i)}$, $i = 1,\dots,n_d$. We can then use the estimate $\hat{f}$ (instead of $f^{*(k)}$) for all of the bootstrap realisations and the adjustment is given by

$$(9) \qquad y_{i,\text{adj}}^{*(k)} = y_i^{*(k)} - \hat{f}(t_i)$$

The resulting bootstrap confidence intervals correspond to a situation where f is known or approximately to a situation where f is stable from one year to another so that it can be estimated with high precision.

If diurnal fluctuations are ignored, or if only day catches are used, eq. 1 may be replaced by the simpler model

$$(10) \qquad y_i = \mu_{d(i)} + \varepsilon_i$$

An alternative, which is often used in practice, is to stratify geographically instead of by day.

Results
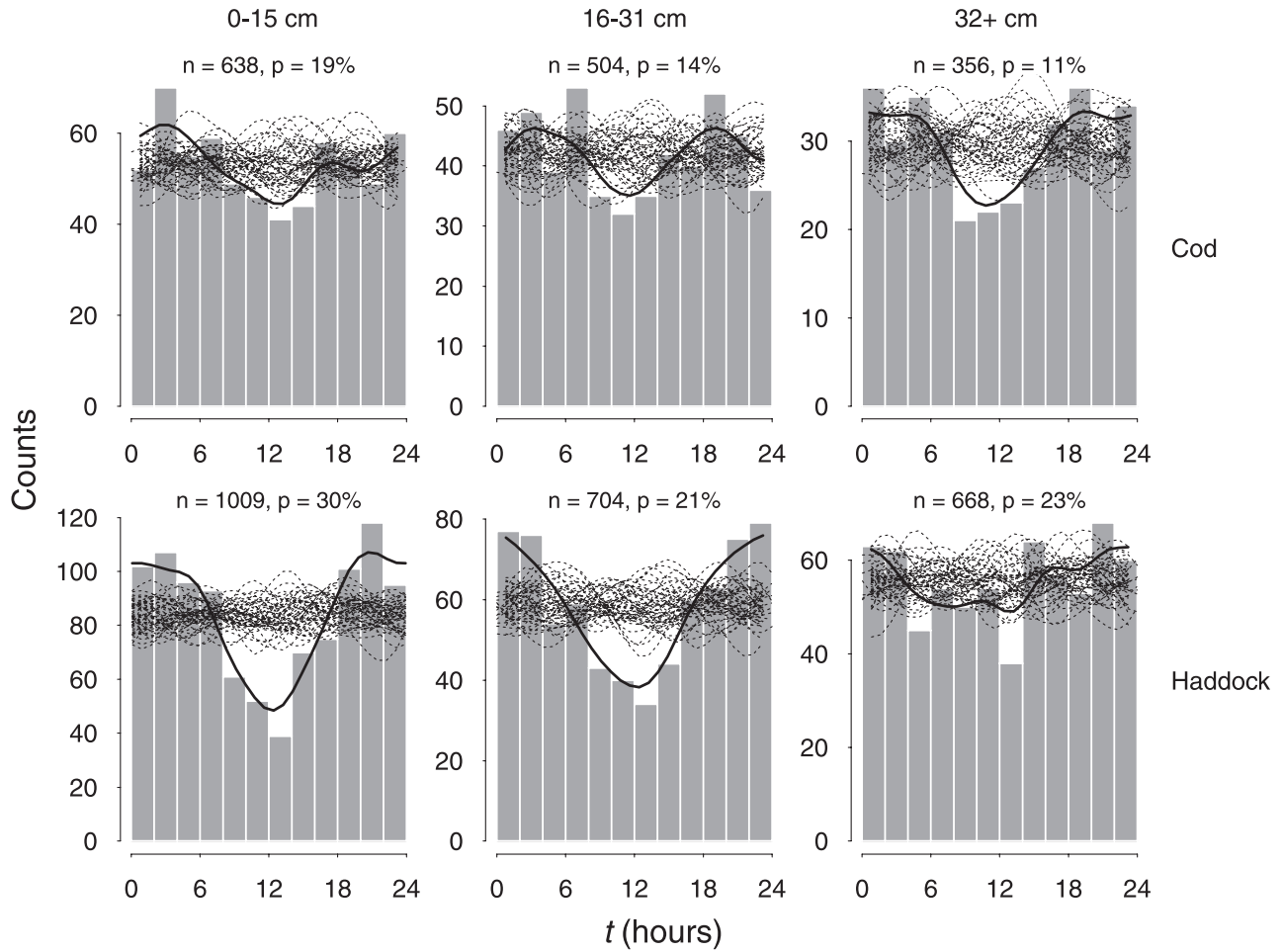
Zero catches

The effect of zero catches was examined by analysing the winter data for cod and haddock. For small- and medium-sized haddock, there was a clearly significant overweight of zero catches at night time (Fig. 2). In most other cases, the tendency was the same, but not so clear. For comparison, density estimates for simulated data sets of the same size but with the zero catch observations randomized over the 24-h cycle are included in the graphs. As will be seen, this is consistent with larger diurnal variation for small- and medium-sized haddock. Also note that zero catches constitute a much higher percentage of the total data material for haddock than for cod, even after the catches taken at days with only zero catches were eliminated. As expected, the results for haddock more than those for cod depend on the strategy for handling zero catches (Table 1). The average diurnal amplitude $\hat{D}$ varied when different strategies were applied, and including all zero catches yielded the smallest $\hat{D}$ in all cases as expected from the argument in the Zero catches section of Material and methods. On the other hand, there was no single strategy yielding the highest $\hat{D}$ in all cases, but setting the threshold at 50% non-zero catches seemed to yield overall reasonable results, and this strategy has been used in all subsequent calculations. Note that in Hjellvik et al. (2001) strategy a "exclude all" was employed, but it is seen that for cod this gives rather similar results to those obtained using the 50% strategy.

The shape of the diurnal variation

The nature and shape of the diurnal oscillations for winter cod data were analysed quite extensively in section 2 of Hjellvik et al. (2001). Nonparametric estimates for the species considered in this paper are depicted in Fig. 3. The S-PLUS function smooth.spline (see Venables and Ripley 1997) was used here and elsewhere for nonparametric estimates. We carried through a successive sequence of tests, as in Hjellvik et al. (2001), of the hypotheses $f = C$ (constant), $f = f_S$, and $f = f_L$ (see Appendix A for details) and found that the winter data are best described by f_L and the autumn data by f_S (Fig. 3). These are adopted as our standard in the fol-

Fig. 2. Histograms and scaled density estimates (thick lines) for the time points at which zero catches were taken during days with at least one nonzero catch. The number of zero catches and their percentage of the total number of hauls are given by *n* and *p*, respectively. The dotted lines show density estimates for 50 data sets of size *n* simulated from a uniform (0,24) distribution. The S-PLUS function density was used for density estimates.


Table 1. Estimates of *D* in eq. 3 with $\alpha = 2$ and $\beta = 8$ when various strategies for handling zero catches are applied.

Species	Method	0–15 cm	16–31 cm	32+ cm
Cod	(*a*) exclude all	0.74 (0.049)	0.62 (0.042)	0.39 (0.042)
	(*b*) include all	0.57 (0.040)	0.58 (0.042)	0.38 (0.043)
	(*c*) include 1	0.64 (0.045)	0.59 (0.043)	0.42 (0.048)
	(*d*) include 25%	0.70 (0.049)	0.60 (0.043)	0.44 (0.049)
	(*e*) include 50%	0.74 (0.052)	0.63 (0.045)	0.45 (0.049)
	(*f*) include 75%	0.83 (0.056)	0.73 (0.046)	0.42 (0.047)
	(*g*) include 100%	0.89 (0.058)	0.70 (0.048)	0.39 (0.043)
Haddock	(*a*) exclude all	1.36 (0.060)	1.09 (0.060)	0.30 (0.062)
	(*b*) include all	1.21 (0.048)	1.06 (0.053)	0.26 (0.047)
	(*c*) include 1	1.33 (0.052)	1.18 (0.058)	0.33 (0.060)
	(*d*) include 25%	1.43 (0.056)	1.24 (0.061)	0.37 (0.064)
	(*e*) include 50%	1.59 (0.064)	1.31 (0.065)	0.45 (0.070)
	(*f*) include 75%	1.66 (0.074)	1.29 (0.070)	0.48 (0.074)
	(*g*) include 100%	1.65 (0.082)	1.25 (0.072)	0.38 (0.073)

Note: Standard errors are given in parentheses. (*a*) and (*b*), all zero catches are excluded or included; (*c*), all catches taken on days with at least one non-zero catch are included, all other catches are excluded; (*d–g*), all catches taken on days with at least 25, 50, 75, and 100% non-zero catches are included, all other catches are excluded.

38

Can. J. Fish. Aquat. Sci. Vol. 59, 2002

Fig. 3. Dots represent mean-adjusted catches $\{y_i - \bar{y}_{d(i)}\}$ from the winter surveys 1985–2000 for cod, haddock, deep-sea redfish, Norway haddock, Greenland halibut, and long rough dab. Solid curves represent the corresponding nonparametric estimates of $f(t)$. Broken curves represent the nonparametric estimates of $f(t)$ for the autumn data, adjusted downwards by 4 units. Vertical broken lines are drawn at $t = 6$, 9, 15, and 18. Results from tests of the hypotheses $f = C$, $f = f_S$, and $f = f_L$ are given in the upper (winter) and lower (autumn) right corners. C indicates that $f = C$ was not rejected at the 5% level. S indicates that $f = C$ was rejected, whereas $f = f_S$ was not. L indicates that $f = C$ and $f = f_S$ were rejected, whereas $f = f_L$ was not. A minus indicates that all three hypotheses were rejected.
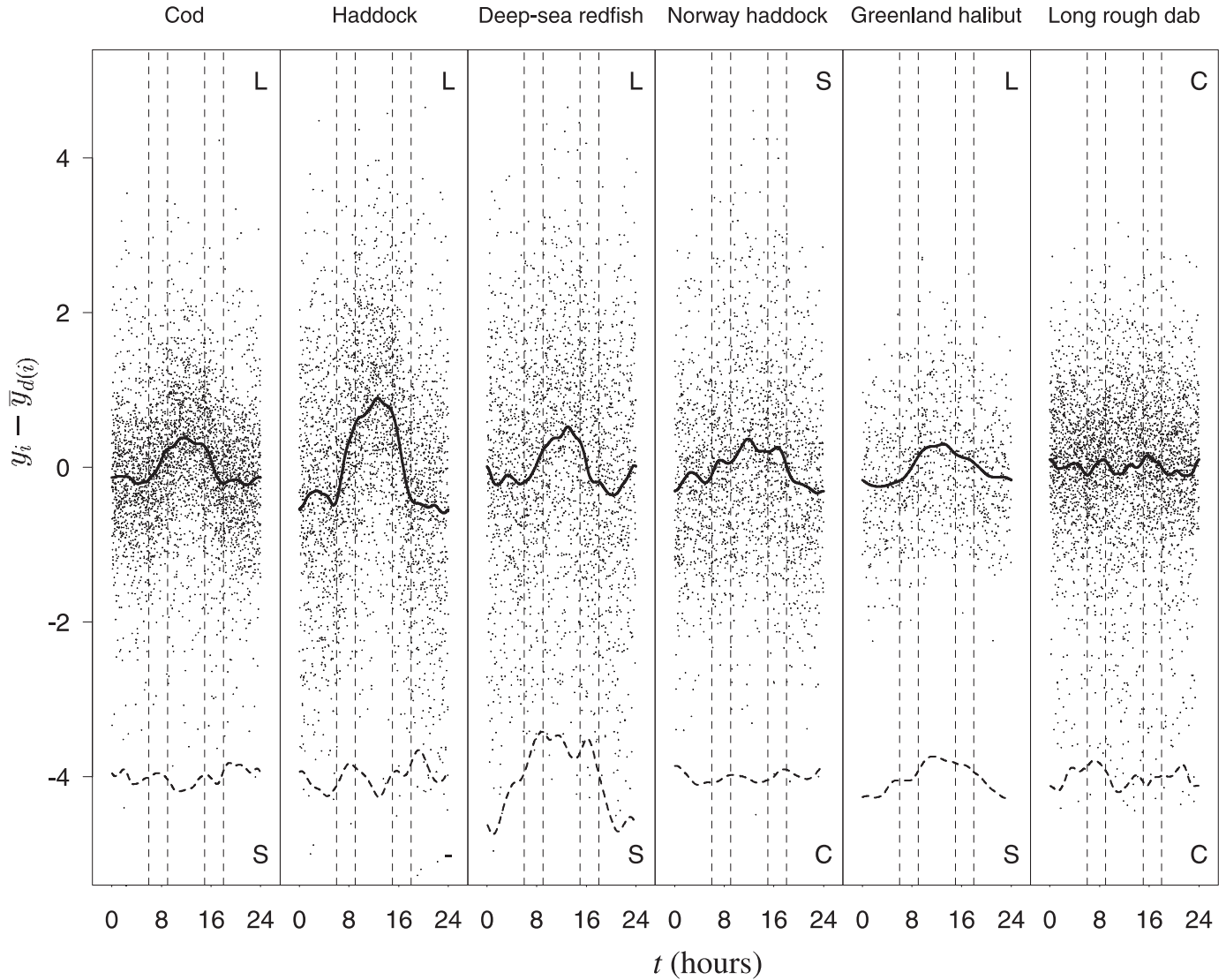
lowing with f_L fitted to the winter data and f_S to the autumn data. Comparing f_L with nonparametric estimates for the length stratified winter data, there is generally a quite good agreement (Fig. 4). In particular, the choice of $\alpha = 2$ in eq. 3 seems appropriate.

**Yearly estimates of diurnal variation**

Motivated by the problem of obtaining adequate measures of uncertainty outlined in the Uncertainty estimates section, we were interested in detecting possible year-to-year differences in diurnal oscillations. We did this by fitting model 2 for the autumn data and model 3 (or, more precisely, model A3 in Appendix A) for the winter data for each year.

For small cod, some of the last years differ markedly from other years (Fig. 5), and the same is true to some extent for small haddock. For large cod and haddock and for other spe-

cies, annual differences are smaller and can be ascribed to random fluctuations in many cases. We applied Simes-modified Bonferroni test (see Appendix A for details) to test the null hypothesis of no year-to-year differences in D for each combination of species and length group. The p values are given on the figure. At a 5% level, the null hypothesis was rejected in 8 of 10 cases for the two smallest length groups but only in 1 of 15 cases for the three largest length groups.

For the autumn data, there were relatively few cases with significant diurnal variation for other species than deep-sea redfish (Fig. 6). One exception is 1994 when D was significantly negative for all length groups for cod and long rough dab. As regards year-to-year differences in D, these were significant for all length groups of deep-sea redfish and long rough dab and for two length groups of cod. However, re-

Fig. 4. Solid curves represent estimates of $f_L(t)$ ($\alpha = 2$ kept fixed) for length-stratified data from the winter surveys 1985–2000 for cod, haddock, deep-sea redfish, Norway haddock, Greenland halibut, and long rough dab. Broken curves represent the corresponding nonparametric estimates of $f(t)$. Some length groups were excluded for some species because of few observations. Confidence intervals (95%) for D and β are indicated on the figure. Vertical broken lines are drawn at $t = 6$, 9, 15, and 18.
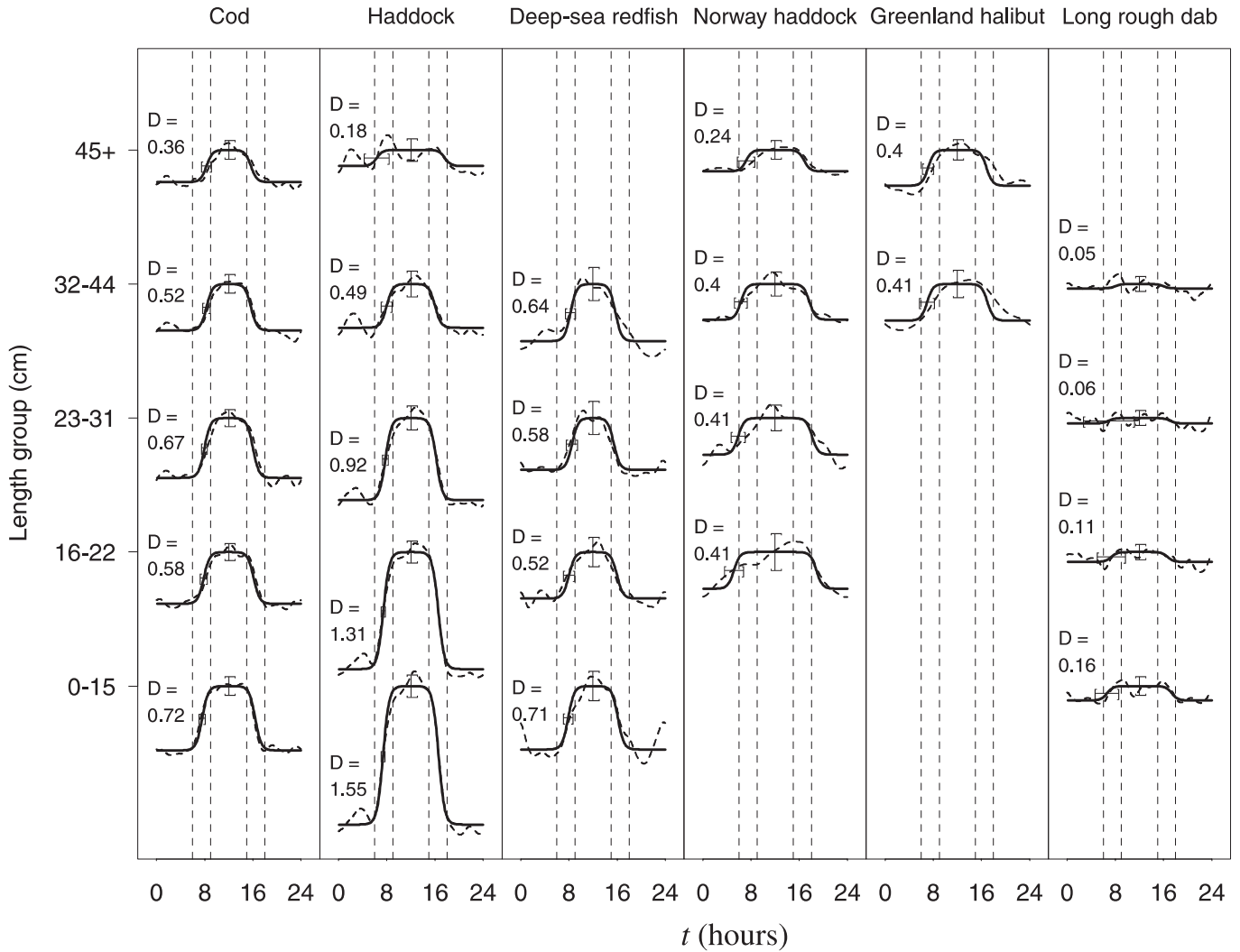
moving the estimates for 1994 for cod and long rough dab, year-to-year differences remained significant only for two length groups of long rough dab (Fig. 6).

Adjusting for diurnal effects

The variation in the diurnal amplitude D from one year to another contributes to the uncertainty in the abundance estimates as measured by the mean $\bar{y}$ of $\{y_i\}$. In this section, we show the results of applying the bias correction described in eq. 4 with accompanying bootstrapped confidence intervals based on eqs. 7–9, including the possibility of a varying D. Only the results for small cod and haddock (0–15 cm) will be displayed. For this length group, the diurnal variations are substantial and undergo significant yearly changes. All results in this section are based on 1000 bootstrap replicas.

The average catches $\bar{y}$ are represented by the point estimates marked as 1 in Fig. 7, which displays 6 years in which D values range from small to large. The corresponding 90% confidence intervals are obtained by bootstrapping as in formula 7. The bias-adjusted estimates are marked as 2 in

Fig. 7, and as can be seen, the effect of the adjustment is quite dramatic when D is large. The adjustment formula (eq. 4) has been used, and bootstrap confidence intervals have been computed from $\{y_{i,\mathrm{adj}}^{*(k)}\}$ in eq. 8 based on yearly estimates of D. The adjusted intervals are presumably correctly located, but their width has increased because of the added uncertainty in estimating D. Actually, comparing with the day-only estimates in 4 and 5 in Fig. 7, the effect of the larger number of observations is more or less cancelled by the added uncertainty of the bias correction. In view of the section Uncertainty estimates (cf. also the example in Appendix A), this is exactly what one would expect. For species and length groups with a stable D (as judged, for example, by using tests depicted in Figs. 5 and 6), however, a more precise estimate of D can be obtained. As more and more data are accumulated, the uncertainty in D can be ignored compared with the uncertainty in $\bar{y}$, and the bootstrap confidence intervals based on a fixed $\hat{f}$ as in eq. 9 (corresponding to f known) can be used. This is not justified for small cod, but to give an impression of the size of the uncertainty re-

40

Can. J. Fish. Aquat. Sci. Vol. 59, 2002

Fig. 5. Yearly estimates of D with 95% confidence intervals for the winter data 1985–2000. Years with $n - n_d - 1 < 20$ are excluded, and where $n - n_d - 1 < 50$, dotted lines are used. Here, n is the number of hauls and n_d the number of days. Model A3 has been used. The numbers at the bottom of each panel are p values for the observed $\hat{D}$s under the null hypothesis of no year-to-year differences in D for each combination of species and length group.
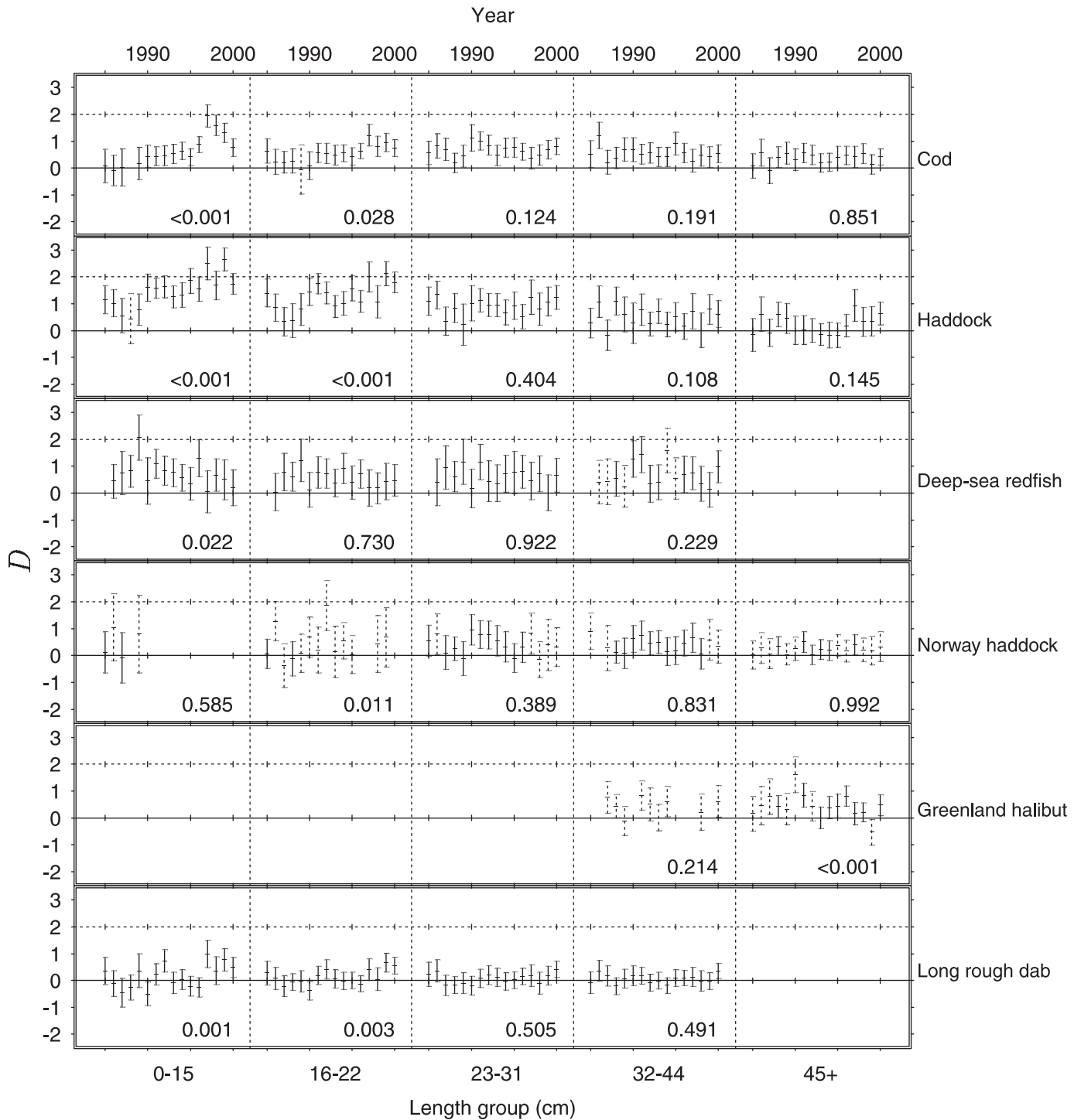

duction involved, these intervals have been included in Fig. 7, marked as 3. As anticipated, they have the same location as those adjusted with f unknown, and their widths are essentially the same as for the unadjusted intervals. In practice, with a moderate number of years available, there will be some uncertainty connected with the estimate of f even though it is based on data from all years. One should then

apply the bootstrap algorithm in the section Uncertainty estimates on the data set consisting of observations from all years, and the resulting confidence intervals would be somewhat broader. We can check the location by computing the intervals using daytime measurements only. This was done by fitting and bootstrapping model 10, but in the bootstrap, we used the residuals (eq. 6) from the full model. A catch is

Fig. 6. Yearly estimates of D with 95% confidence intervals for the autumn data 1985–2000. Years with $n - n_d - 1 < 20$ are excluded, and where $n - n_d - 1 < 50$, dotted lines are used. Here, n is the number of hauls and n_d the number of days. Model 2 has been used. The numbers at the bottom of each panel are p values for the observed $\hat{D}$s under the null hypothesis of no year-to-year differences in D for each combination of species and length group. The numbers in parentheses for cod and long rough dab are the p values when 1994 is left out.
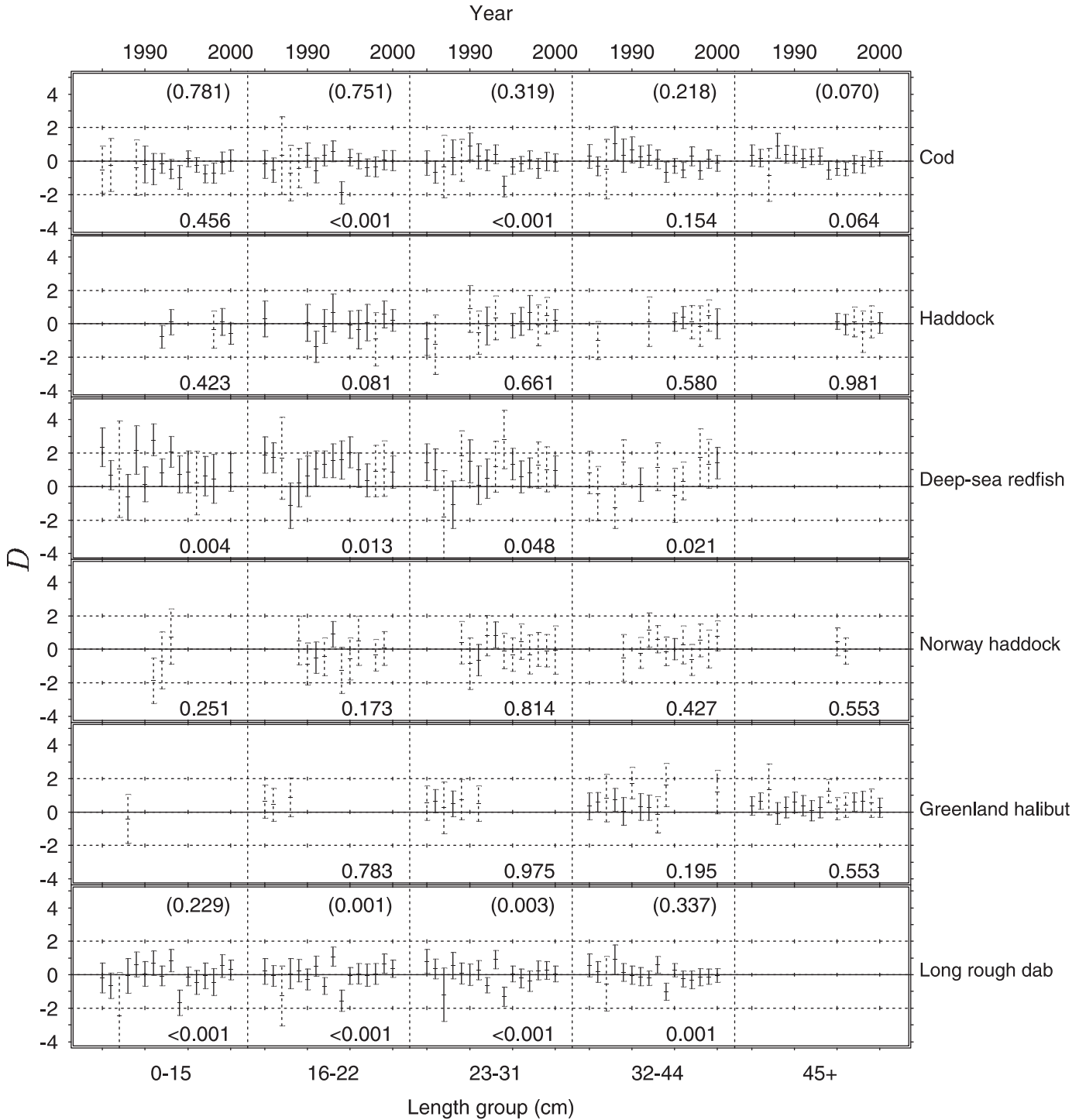
considered a daytime measurement if it is taken at time $\beta' \leq t_i \leq 24 - \beta'$, where $\beta' = 7.60$ is the estimate of β from the fit of eq. A3 in Appendix A. These intervals are marked as 4 in Fig. 7. Both location and width are about the same as for the bias-adjusted intervals obtained by using both night and day measurements, but with a tendency of the day intervals to be

biased downwards. This is because day time is defined so that half of the transition phase between day and night is included. Defining day and night in terms of the altitude of the sun also yields very similar results, as can be seen from the intervals marked as 5 in Fig. 7. Here, based on Hjellvik et al. (2001), the catches taken when the sun is less than 5° un-

Fig. 7. Bootstrapped means and 90% confidence intervals for mean catches $\bar{y}$ for 0- to 15-cm cod, winter. From left to right: 1) un-adjusted catches; 2) adjusted catches; 3) adjusted catches with f assumed known; 4) day catches with day defined as $\beta' \leq t_i \leq 24 - \beta$, where $\beta = 7.60$; and 5) day catches with day defined as the altitude of the sun being $\geq -5°$. The model $f_L(t)$ with $\alpha = 2$ and $\beta = 7.60$ fixed has been used in eqs. 8 and 9, and 1000 bootstrap replicas were used everywhere.
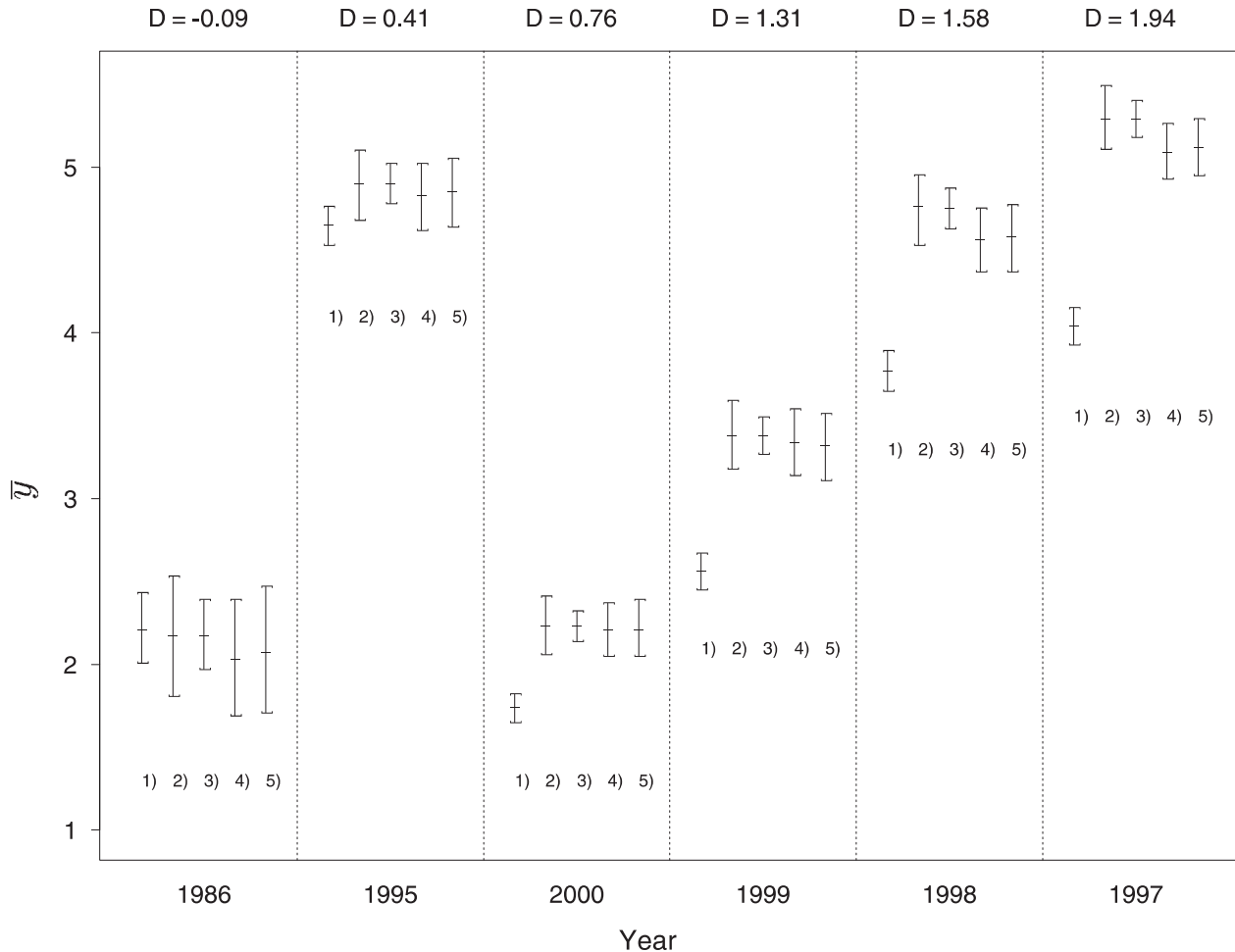
der the horizon are included (cf. also Korsbrekke and Nakken (1999)).

For the raw data, i.e., $x_i = e^{y_i}$, the mean for the adjusted catches of 0- to 15-cm cod is more than twice as high as the mean of the unadjusted catches in 1997–1999 when the diurnal variation is highest (Fig. 8). But that is also true for the standard deviation in both 1997 and 1998. The bootstrapped confidence intervals on the figure are centred at the observed means. The mean square error $\text{MSE}(\bar{x}) = \text{var}(\bar{x}) + \text{bias}^2(\bar{x})$ was calculated by setting $\text{var}(\bar{x})$ equal to the bootstrap variance of $\bar{x}$, and $\text{bias}(\bar{x}) = \bar{x}_{adj} - \bar{x}$. Similarly, $\text{MSE}(\bar{x}_{adj})$ was calculated by setting $\text{bias}(\bar{x}_{adj}) = 0$. The MSE was smaller for the adjusted means than for the unadjusted means in all cases for both cod and haddock (Fig. 8). For haddock, the unadjusted curve reveals a zigzag structure when plotted against year (Fig. 8b). This is amplified in the adjusted curve because the highest values of $\hat{D}$ occur in the years with the highest catch.

The significant annual variation in D for small cod and haddock also leads to huge year-to-year differences in the adjustment factor $\Delta_Q = \bar{x}_{adj}/\bar{x}$ (Fig. 9). The year-to-year variation in Δ_Q tends to decrease with fish length, especially for haddock. Taking the uncertainty of the estimates into ac-

count, it is seen that as a rough approximation, Δ_Q for large fish can be taken to be constant. If the years 1997–1999 are excluded, Δ_Q also is much more stable for small cod. For haddock, the average Δ_Q is clearly decreasing with increasing fish length, whereas for cod, it is relatively independent of fish length, but somewhat larger for small fish.

Discussion

The main objective of this paper has been to study diurnal bias, to correct for it, and to examine its influence on the uncertainty of abundance estimates. The bias has been studied as a function of species, length group, and season. Particular emphasis was put on investigating its stability, or lack of such, from one year to another. The uncertainty of a bias-adjusted abundance estimate depends critically on that stability. In the following, we will discuss these aspects, starting with a short discussion of the model.

Some comments on the model

Our modeling approach is completely general and is applicable to any bias adjustment and uncertainty analysis of this kind. In particular, we believe that it may be useful in a

Fig. 8. Observed values of $\bar{x}$ (solid lines) and $\bar{x}_{adj}$ (dotted lines) for (*a*) 0- to 15-cm cod, winter, and (*b*) 0- to 15-cm haddock, with bootstrapped 90% confidence intervals. Unadjusted and adjusted estimates correspond to 1) and 2) in Fig. 7. For each year, $\hat{D}$ and the ratio MSE($\bar{x}$)/MSE($\bar{x}_{adj}$) are given at top and bottom, respectively, where MSE is mean square error.
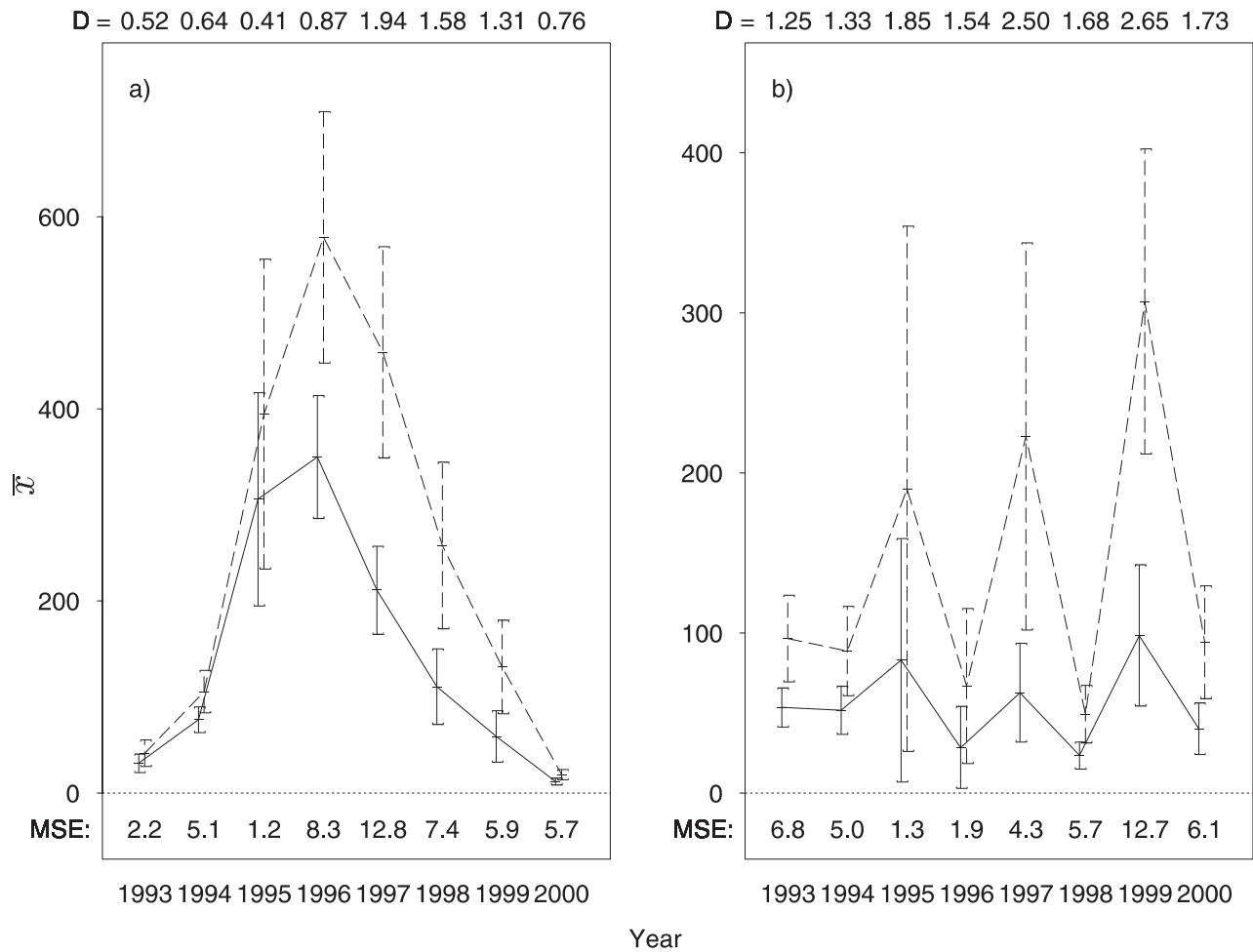
combined study of acoustic surveys and trawl catches. The logistic model f_L with its smooth transition between night and day level is a refinement compared with the models of only two levels used by Casey and Myers (1998) and Korsbrekke and Nakken (1999). Casey and Myers (1998) base their classification on local times of sunrise and sunset, but to remove transitory effects, they omit hauls within 1 h of sunset or sunrise. Korsbrekke and Nakken (1999) use the altitude of the sun, and catches taken when the sun is less than 5° below the horizon are classified as daytime catches, which in view of the results of Hjellvik et al. (2001) seems sensible.

Still, our model is quite crude for several reasons. First, the function f is taken to be independent of the time of sunrise and sunset of day $d(i)$. Second, explanatory variables such as location and depth have been omitted. Third, species are interacting in an ecosystem, and the true dynamics and variation in abundance and composition cannot be fully explained merely from a simple single-species approach (Bogstad and Mehl 1997). For instance, can the larger diurnal variation for small fish be better understood in a multispecies context? The first point is discussed in Hjellvik et al. (2001), and a model containing explanatory variables and a

preliminary investigation of the multispecies aspect are parts of the effort in Hjellvik et al. (unpublished data).

The assumption of independent, identically distributed residuals is not quite fulfilled, as there is a tendency of getting large residuals ε_i at stations with low predicted values. When taking antilogarithms, biased bootstrap distributions of $\bar{x}$ and $\bar{x}_{adj}$ result because large residuals ε_i^* in combination with large predicted values $\hat{y}_i^*$ may yield very large values of $x_i^* = \exp(y_i^*) = \exp(\varepsilon_i^* + \hat{y}_i^*) = \exp(\varepsilon_i^*)\exp(\hat{y}_i^*)$. Thus, the means of the bootstrap distributions are typically higher than the corresponding observed means. We have chosen to centre the confidence intervals so that they are symmetric around the observed means. In most cases, the largest residuals occur at nighttime (Fig. 10), but bootstrapping night and day data separately did not have much effect on the skewness of the confidence intervals. However, this approach typically produced slightly narrower confidence intervals. For the unadjusted means for haddock in 1995 and 1996, where the tendency of getting largest residuals at night is strongest, the variance was reduced to about 50%, but the ratio MSE($\bar{x}$)/MSE($\bar{x}_{adj}$) remained practically unchanged.

As a technical detail, it should be noted that for small values of α (slow transition), the parameter D is no longer

44

Can. J. Fish. Aquat. Sci. Vol. 59, 2002

Fig. 9. The catchability ratio $\Delta_Q = \bar{x}_{adj}/\bar{x}$ for (*a*) cod, winter, and (*b*) haddock plotted against length for the years 1985–2000 (dotted lines). The thick lines are the averages of the dotted lines, weighted with the yearly number of hauls used.
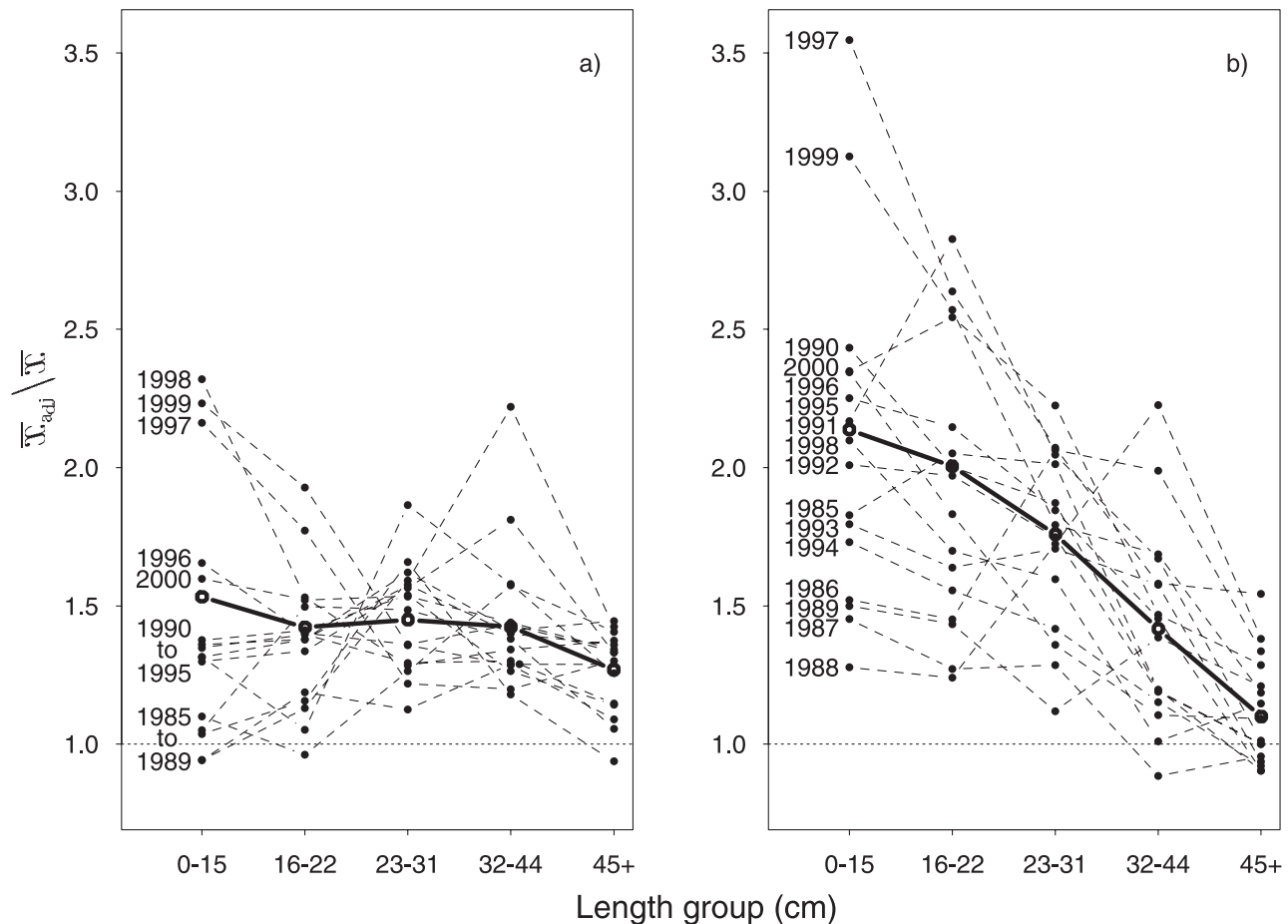
equivalent to the difference between day and night levels. This is because the maximum and (or) minimum of $e^{\alpha(t-\beta)}/(1 + e^{\alpha(t-\beta)})$ in eq. 3 is not obtained for $0 \leq t \leq 12$. In such cases, the sinusoidal model f_S is a better choice, and in the case of small oscillations, when the estimation of α is encumbered with large uncertainty, we recommend that it be kept fixed, as has been done in most of this paper.

**Instability of diurnal variation in time and dependence on depth and length**

Relatively large seasonal differences were found in both amplitude and shape of diurnal oscillations. This is not surprising in view of the different light conditions, but it does not really create much trouble in the adjustment procedure as the most appropriate of the sinusoid and the logistic function can be applied in eq. 4. More serious from an assessment point of view is the instability of diurnal amplitude from one year to the next. The year-to-year differences are largest for small fish, in particular for small cod and haddock. There are no commercial data available for recruiting age groups, and thus survey-based assessment of these age groups is particularly important. To understand these fluctuations in D, it should be remembered that juvenile groundfish often have a pelagic distribution and therefore would be partly unavailable to the bottom trawl during their 1st year of life. Age-1 fish (corresponding more or less to the 0- to 15-cm length group) have not always completed bottom set-

tlement at the time of the winter survey (Godø and Sunnanå 1992), and hence there will be a year effect on the availability of these groups to the trawl. Also, as shown by Aglen et al. (1999), small fish may have substantial vertical migration dynamics.

We know also that larger-sized cod migrate vertically, but the results of the present paper and new information from data storage tags indicate that the diurnal pattern is not as pronounced as for small fish (Godø and Michalsen 2000). Further, the substantial vertical herding (fish swimming from the safe pelagic zone into the catching volume of the trawl) observed for large fish (e.g., Ona and Godø 1990) will probably reduce the effect of diurnal migrations on catchability for these size groups.

This difference in behaviour is enhanced with increasing depth (Hjellvik et al., unpublished data), a larger depth generally leading to larger diurnal oscillations for small fish. Annual differences in average depth may therefore result in different D values. In fact, for small cod, there is a strong positive relationship ($R^2 = 0.84$) between average depth and $\hat{D}$ (Fig. 11). In particular, the high D values for 1997 and 1998 can be explained from this regression, as the winter survey did not enter the Russian zone during those years and the excluded eastern part is relatively shallow leading to a substantially larger average depth. The variation in average depth for the remaining years is minor and is caused partly by small year-to-year differences in the area covered (e.g.,

Fig. 10. Residuals from model 1 with $f = f_L$ plotted against time of day for 0- to 15-cm cod and haddock, winter. The model was fitted to data from each year separately, with $\alpha = 2$ and $\beta = 7.60$ (cod) or 7.34 (haddock) used in eq. 3. Dotted lines are drawn at $t = \beta$ and $t = 24 - \beta$.
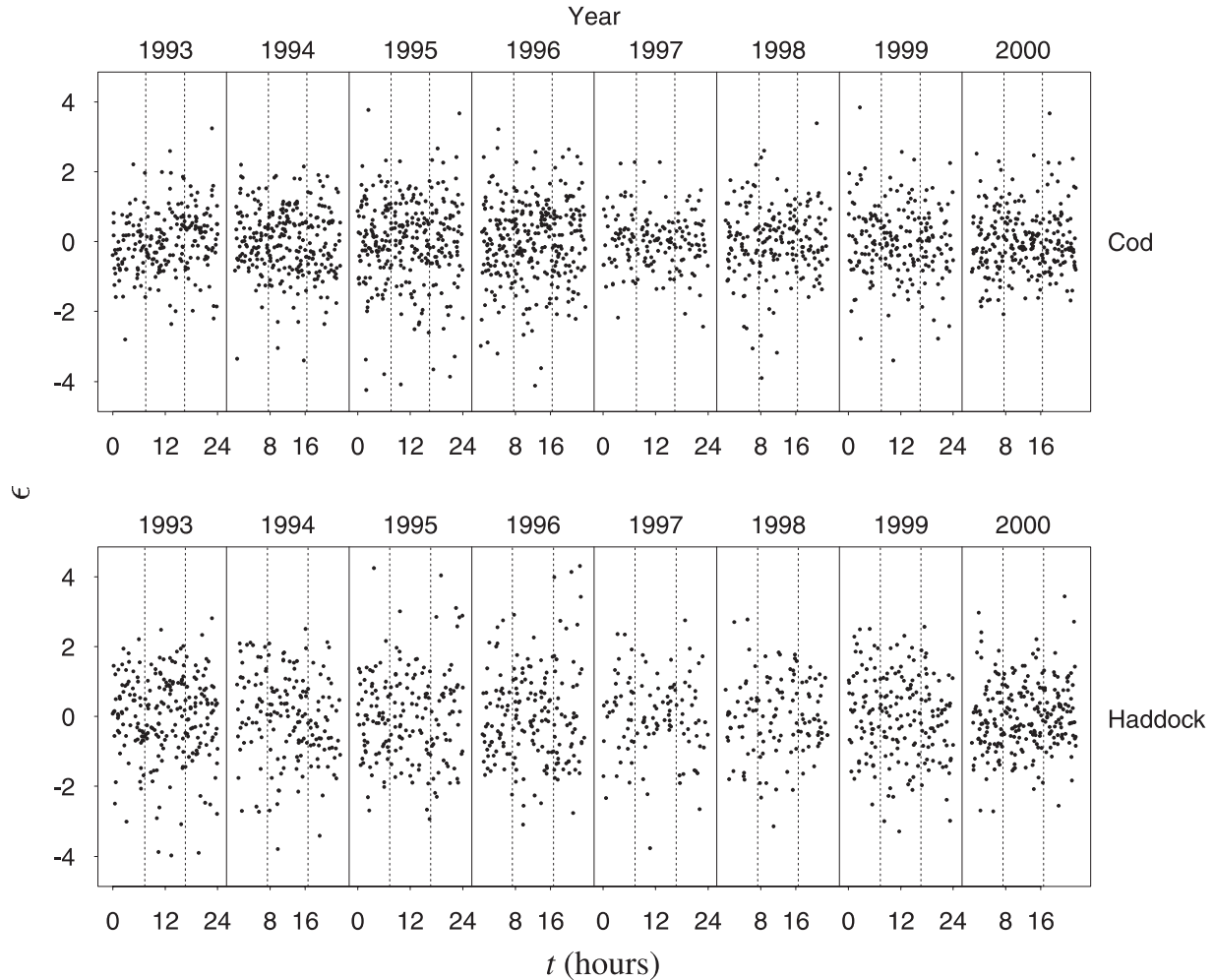
resulting from varying ice conditions and changes in survey strategy), partly by year-to-year variation in the spatial distribution of small cod combined with our strategy for discarding of zero-catch hauls.

For small haddock, the annual dependence of $\hat{D}$ on depth as measured by R^2 is much weaker ($R^2 = 0.15$). However, this does not mean that the depth dependency as such is weaker for haddock than for cod. Because of the more restricted distribution of haddock, particularly towards the east (see Bergstad et al. 1987), the average depth was not affected in the same way by restrictions in area coverage in the Russian zone. The resulting small year-to-year variation in average depth explains the low R^2 value. Changes in fish density as measured by $\bar{y}$ explain more of the variation in $\hat{D}$ ($R^2 = 0.45$).

The transition around 1990 from low to higher values of D for small cod and haddock could be due to the change of ground gear in 1989 (cf. Engås and Godø 1989; Jakobsen et al. 1997), but a similar transition is not observed for other species.

Clearly, the temporal instability of diurnal variation needs to be more fully investigated, for example, by integrating results from hydroacoustics and multispecies effects. Espe-

cially, it would be of interest to examine the effect of bottom settlement on diurnal variation for small fish.

Catchability

The results of this paper show that the diurnal correction, as measured by the adjustment factor $\Delta_Q = \bar{x}_{adj}/\bar{x}$, in many cases exceeds a factor of 2. This implies that a sizable proportion of the catchability coefficient is made up of diurnal bias. Thus, the temporal instability of the adjustment factor indicates that there is a corresponding temporal instability of the catchability coefficient. Similarly, variation in Δ_Q with fish length indicates a length-dependent catchability coefficient. For cod, the dependency of Δ_Q on fish length found in this paper is not in complete agreement with the results in fig. 3 of Korsbrekke and Nakken (1999). They give the average day–night ratios for abundance indices for the years 1989–1996, and the highest ratio is obtained for 23- to 31-cm fish. However, using data from 1989–1996 only, our results are more similar to theirs with the highest Δ_Q obtained for 23- to 31-cm and 32- to 44-cm fish.

If other factors that make up the catchability coefficient are length independent, our results mean that the catchability for haddock (averaged over night and day) is clearly de-

46

Can. J. Fish. Aquat. Sci. Vol. 59, 2002

Fig. 11. Yearly estimates of D for 0- to 15-cm cod, winter, plotted against average yearly depth for the stations used for the estimates. Confidence intervals (95%) for D are included.
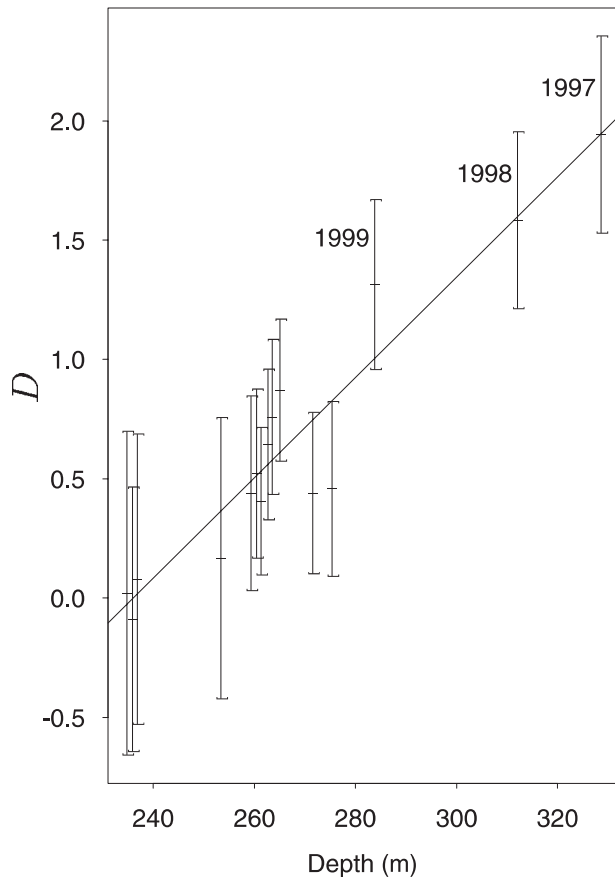
creasing with decreasing fish length in that night catches for small fish are much lower than day catches. Using a constant catchability, as is now done, would tend to underestimate the relative proportion of small fish by a factor that can be numerically derived from Δ_Q.

Does it pay to adjust?

We have argued for a model-based adjustment of survey trawl catches to minimise bias and variation caused by diurnal effects. In essence, the goal has been to examine integrity, within and between years, in the assessment results for individual species. Adjustment removes diurnal bias, but at the cost of increased uncertainty of the adjusted estimates. Thus, it is not obvious whether it pays to adjust. In fact, the answer to this question depends on our goal. Do we want to estimate absolute fish density? Are we only interested in tracking the changes in relative fish density for the different age or length groups over years? Or do we want to track each year-class over its life span?

Sometimes, in situations where the bias may be reduced at the cost of an increased variance, decision rules based on the MSE are used. Munro (1998) presents such a rule for a somewhat analogous situation. Our bootstrap method yields estimates of the variance of both the adjusted and unadjusted estimates and of the bias of the unadjusted estimates relative to the adjusted ones. But in terms of absolute abundance measures, the adjusted estimates are most likely biased

downwards as well. Considering the winter data from 1993 through 2000 for small cod and haddock and first assuming unbiased adjusted estimates, the MSE is much smaller for the adjusted means than for the unadjusted ones. If the downward bias of the adjusted estimates (relative to absolute densities) is taken into account, the difference in MSE is enhanced. In terms of the MSE, adjustment seems to pay even for relatively small diurnal fluctuations (in 1995, with $D = 0.41$ for small cod, the MSE was about 15% larger for the unadjusted data than for the adjusted ones). In a situation where diurnal fluctuations are stable from year to year, the variance, and hence the MSE, of the adjusted means will be even smaller compared with the unadjusted ones, and adjustment will pay for even smaller values of D.

If the only issue of interest is to track relative changes in density of a given age or length group over years and there are no significant year-to-year differences in diurnal amplitude, then the adjustment will just introduce an extra source of variation and is not recommended. On the other hand, if there are substantial year-to-year differences, adjustment is generally recommended to remove year-dependent biases. One could use the modified Bonferroni test in Appendix A or some other test to check whether such differences exist. If one is interested in tracking year-classes from birth to extinction, one should adjust if there are substantial differences in diurnal variation between length groups.

In the above discussion, it has been assumed that catchability at day time is more or less stable over years. However, year-to-year differences in bottom settlement for small fish (e.g., Godø and Sunnanå 1992) may affect both daytime catchability and diurnal variation, and ideally the bottom settlement process should be monitored to control its effect on the catchability of small fish. In fact, adjustment may potentially reduce the quality of the time series of abundance estimates of age-1 fish if bottom settlement varies substantially from one year to another.

Because the adjusted estimates, when the diurnal amplitude is based on the current year only, have more or less the same properties as those based on day catches only, it might be a proper strategy to restrict surveys to day if all species and length groups of interest exhibit strong year-to-year differences in diurnal variation. This is not the case for the Barents Sea surveys, and a 24-h survey strategy is adequate because more information is then obtained for medium- and large-sized fish.

Even though standard bottom trawl survey results have worked quite well for stock assessment in the past (Korsbrekke et al. 2001), new demands may in future call for more correct estimation of the absolute level. Particularly, correct density relations among size groups and species will be important in more advanced ecosystem approaches that demand correct data on trophic relationships. As diurnal behaviour dynamics are important for intra- and inter-specific relationships (e.g., Neilson and Perry (1990) and references therein), it will be important to establish models that can adjust for the associated bias. This may potentially improve our understanding of the ecosystem function and also enhance the efficiency of more advanced assessment models. Such adjustments could also improve the acoustic assessment because information from trawl catches normally is applied to convert acoustic back-scattering cross section to fish

densities by species and size. Finally, our results should be taken into account in the design of future survey strategies. This will be particularly important when combining density information from trawl catches with simultaneous density records from the acoustic method.

Aknowledgments

References

Aglen, A., Engås, A., Huse, I., Michalsen, K., and Stensholt, B. 1999. How vertical fish distribution may affect surveys results. ICES J. Mar. Sci. **56**: 345–360.

Bergstad, O.A., Jørgensen, T., and Dragesund, O. 1987. Life history and ecology of the gadoid resources of the Barents Sea. Fish. Res. **5**: 119–161.

Bogstad, B., and Mehl, S. 1997. Interactions between cod (*Gadus morhua*) and its prey species in the Barents Sea. Forage fishes in marine ecosystems. *In* Proceedings of the International Symposium on the Role of Forage Fishes in Marine Ecosystems. Alaska Sea Grant College Program Report No. 97-01, University of Alaska Fairbanks. pp. 591–615.

Casey, J.M., and Myers, R. 1998. Diel variation in trawl catchability: is it as clear as day and night? Can J. Fish. Aquat. Sci. **55**: 2329–2340.

Demer, D.A., and Hewitt, R.P. 1995. Bias in acoustic biomass estimates of *Euphasia suberba* due to diel vertical migration. Deep-Sea Res. Part I, **42**: 455–475.

Engås, A., and Godø, O.R. 1989. Escape of fish under the fishing line of a Norwegian sampling trawl and its influence on survey results. J. Cons. Int. Explor. Mer, **45**: 269–276.

Godø, O.R. 1994. Factors affecting reliability of groundfish abundance estimates from bottom trawl surveys. *In* Marine fish behaviour in capture and abundance estimation. *Edited by* A. Fernø and S. Olsen. Fishing News Books, Oxford. pp. 166–199.

Godø, O.R., and Michalsen, K. 2000. Migratory behaviour of northeast Arctic cod, studied by use of data storage tags. Fish. Res. **48**: 127–140.

Godø, O.R., and Sunnanå, K. 1992. Size selection during trawl sampling of cod and haddock and its effect on abundance indices at age. Fish. Res. **13**: 293–310.

Hjellvik, V., and Tjøstheim, D. 1995. Nonparametric tests of linearity for time series. Biometrika, **82**: 351–368.

Hjellvik, V., Godø, O.R., and Tjøstheim, D. 2001. Modelling diurnal variation of marine populations. Biometrics, **57**: 189–196.

Jakobsen, T., Korsbrekke, K., Mehl, S., and Nakken, O. 1997. Norwegian combined acoustic and bottom trawl surveys for demersal fish in the Barents Sea during winter. ICES CM/Y:17.

Korsbrekke, K., and Nakken, O. 1999. Length and species dependent diurnal variation in catch rates in the Norwegian Barents Sea bottom trawl surveys. ICES J. Mar. Sci. **56**: 284–291.

Korsbrekke, K., Mehl, S., Nakken, O., and Pennington, M. 2001. A survey-based assessment of the Northeast Arctic cod stock. ICES J. Mar. Sci. **58**: 763–769.

Munro, T.M. 1998. A decision rule based on the mean square error for correcting relative fishing power differences in trawl survey data. Fish. Bull. **96**: 538–546.

Neilson, J.D., and Perry, R.I. 1990. Diel vertical migrations of marine fishes: an obligate or facultative process? Adv. Mar. Biol. **26**: 115–168.

Ona, E., and Godø, O.R. 1990. Fish reaction to trawling noise: the significance for trawl sampling. Rapp. P.-v. Réun. Cons. Int. Explor. Mer, **189**: 159–166.

Pennington, M., and Godø, O.R. 1995. Measuring the effect of changes in catchability on the variance of marine survey abundance. Fish. Res. **23**: 301–310.

Pennington, M., and Vølstad, J.H. 1991. Optimum size of sampling unit for estimating the density of marine populations. Biometrics, **47**: 717–723.

Simes, R.J. 1986. An improved Bonferroni procedure for multiple tests of significance. Biometrika, **73**: 751–754.

Smart, W.M. 1977. Textbook on spherical astronomy. Cambridge University Press, Cambridge, U.K.

Venables, W.N., and Ripley, B.D. 1997. Modern applied statistics with S-PLUS. 2nd ed. Springer, New York.

Wakabayashi, K., Bakkala, R.G., and Alton, M.A. 1985. Methods of the U.S.–Japan demersal trawl surveys. *In* Results of the cooperative U.S.–Japan groundfish investigations in the Bering Sea during May–August 1979. *Edited by* R.G. Bakkala and K. Wakabayashi. Int. North Pac. Fish. Comm. Bull. **44**: 7–26.

Walsh, S.J. 1988. Diel variability in trawl catches of juvenile and adult yellowtail flounder on the Grand Banks and the effect on resource assessment. N. Am. J. Fish. Manage. **8**: 373–381.

Appendix A

Details of the models in Models section

The apparent local time t in eq. 1 is calculated as $t = t_{UTC} + \text{longitude}/15 + \Delta$, where $\Delta = [-0.4083\sin(0.0172(\text{day} - 80)) - 1.7958\cos(0.0172(\text{day} - 80)) + 2.4875\sin(0.0344(\text{day} - 80))]/15$, with $1 \leq \text{day} \leq 365$ being the day of the year, is an approximation to the equation of time (Smart 1977, p. 150).

If the parameters α and β are kept fixed in eq. 3, then eq. 1 may be replaced by the linear model

$$(A1) \quad y_i = \mu_{d(i)} + Df_L^*(t_i) + \varepsilon_i$$

where

$$f_L^*(t_i) = f_L(t_i; D = 1, \alpha = \alpha_0, \beta = \beta_0)$$

with α_0 and β_0 being appropriately chosen constants. To get separate estimates of D for each year, we may write

$$(A2) \quad y_i = \mu_{d(i)} + D_{a(i)}f_L^*(t_i) + \varepsilon_i$$

where $a(i)$ is the year that haul i is taken. If the sinusoidal function is preferred, $f_L^*(t_i)$ in eqs. A1 and A2 can be replaced by $f_S^*(t_i) = f_S(t_i; D = 1)$.

Also, the nonlinear model may be refined to yield separate estimates of D for each year by simply replacing D in eq. 3 by $D_{a(i)}$. Then α and β can be estimated using data for all years, whereas $D_{a(i)}$ is estimated using data from year $a(i)$ only. In the section Yearly estimates of diurnal variation, we have used the nonlinear version with $\alpha = 2$ fixed, that is,

$$(A3) \quad y_i = \mu_{d(i)} + f_L(t_i; D = D_{a(i)}, \alpha = 2, \beta) + \varepsilon_i,$$

whereas for the bootstrap experiments in the section Adjusting for diurnal effects, we used eq. A2.

Stylized example illustrating the effect of uncertainty on bias correction

Let μ_N and μ_D be the nighttime and daytime levels of a pure threshold (night–day levels only) model, and let nighttime measurements $z_{i,N}$ and daytime measurements $z_{i,D}$ have expectations $E(z_{i,N}) = \mu_N$ and $E(z_{i,D}) = \mu_D$, respectively, with a common variance $\mathrm{var}(z_{i,N}) = \mathrm{var}(z_{i,D}) = \sigma^2$. If there are n_N nighttime measurements and n_D daytime measurements, then $\mu_N - \mu_D$ is estimated by

$$\bar{z}_N - \bar{z}_D = n_N^{-1} \sum_{i=1}^{n_N} z_{i,N} - n_D^{-1} \sum_{i=1}^{n_D} z_{i,D}$$

and the adjusted measurements are given by

$$z_{i,\mathrm{adj}} = \begin{cases} z_{i,D} & \text{for daytime measurements} \\ z_{i,N} - (\bar{z}_N - \bar{z}_D) & \text{for nighttime measurements} \end{cases}$$

It is then trivial to check that $\bar{z}_{\mathrm{adj}} = (n_N + n_D)^{-1} \sum_i z_{i,\mathrm{adj}} = \bar{z}_D$ so that the adjusted estimate and the daytime estimate are identical and therefore have the same error, irrespective of the values of n_N and n_D.

Testing for model fitness

Tests of the hypotheses $f = C$ (constant), $f = f_S$, and $f = f_L$ were executed by computing functionals of the data

$$\sum_i \{\hat{f}(t_i) - \hat{C}\}^2; \quad \sum_i \{\hat{f}(t_i) - \hat{f}_S(t_i)\}^2; \quad \sum_i \{\hat{f}(t_i) - \hat{f}_L(t_i)\}^2$$

measuring the distances between the nonparametric estimate $\hat{f}$ and the parametric estimates $\hat{C}$, $\hat{f}_S$, and $\hat{f}_L$ under the null hypotheses $f = C$, $f = f_S$, and $f = f_L$. The hypotheses are rejected for large values of the functionals, the null distribution of these functionals being established by bootstrapping. For more details, we refer to Hjellvik and Tjøstheim (1995) and Hjellvik et al. (2001). In general α, β, and D were all estimated under H_0: $f = f_L$, but in one case (deep-sea redfish), $\alpha = 2$ were kept fixed because of convergence problems with the nls algorithm.

It should be noted that our test of $f = C$, i.e., whether there are diurnal variations or not, does not suffer from the bias relative to a randomized procedure, observed for the test in Casey and Myers (1998). This is because we use the bootstrap principle so that the randomization is inherent in the test itself.

Testing for annual differences in diurnal variation

We have used Simes modified Bonferroni procedure (Simes 1986) to test for annual differences. This is an improved version of the Bonferroni correction for conducting multiple tests of significance. If $H = \{H_1, \ldots, H_n\}$ is a set of null hypotheses with corresponding test statistics $T_1, \ldots, T_n$, p values $P_1, \ldots, P_n$, and H_0 is the hypothesis that H_i, $i = 1, \ldots, n$ are true, then the suggested procedure rejects H_0 at level α if and only if

(A4)   there exists some value of $j (1 \le j \le n)$

such that $P_{(j)} \le j\alpha/n$

where $P_{(1)}, \ldots, P_{(n)}$ are the ordered values of $P_1, \ldots, P_n$.

This procedure can be applied to our situation. Let $\{D_1, \ldots, D_m\}$ denote the true values of D for the m years where estimates of D are available for the actual length group and species. The null hypothesis H_0: $D_1 = \ldots = D_m$ can then be restated as H_0: $\delta_{ij} \doteq D_i - D_j = 0$, $i = 1, \ldots, m - 1$, $j = i + 1, \ldots, m$. The global null hypothesis H_0 is then true if and only if all of the $n = m(m - 1)/2$ null hypotheses $\{H_{ij}: \delta_{ij} = 0\}$ are true. Under H_{ij}, we have that $E(\delta_{ij}) = 0$ and $\mathrm{var}(\delta_{ij}) = \sigma_i^2 + \sigma_j^2$, where $\sigma_i^2 = \mathrm{var}(D_i)$, $i = 1, \ldots, m$. We assume that the test statistic $T_{ij} = (\hat{D}_i - \hat{D}_j)/(s_i^2 + s_j^2)^{1/2}$, where s_i is the standard error for $\hat{D}_i$ yielded by the function lm in S-PLUS, follows a standard normal distribution. The p value for T_{ij} is then calculated as $P_{ij} = 2\min(S_{ij}, 1 - S_{ij})$, where $S_{ij} = \Phi(T_{ij})$ and Φ is the standard normal distribution function. The global p value is the lowest value of α for which eq. A4 is fulfilled ($P_{(1)}, \ldots, P_{(n)}$ being the ordered values of $\{P_{ij}\}$).

Simes modified Bonferroni test appears particularly advantageous compared with the classical Bonferroni procedure when several highly correlated test statistics are involved (Simes 1986). In our situation, many of the test statistics are positively correlated since $\mathrm{cov}(D_i - D_j, D_i - D_k) = \mathrm{var}(D_i)$ for $i \ne j \ne k$. Still, the actual level of the test is acceptable. Using a nominal level of 5% and 10 000 replicates, we simulated D_is from the $N(0, s_i^2)$ distribution for all species or length groups with $m > 3$ and got empirical levels ranging from 0.0344 to 0.0450, with an average of 0.0391.