CM 1999/J:06 Application of Acoustic Techniques to Bottom Trawl Surveys.
Not to be sited without prior reference to the author.

# Modelling Diurnal Variation in Bottom Trawl Catches and Potential Application in Surveys

V. Hjellvik, O. R. Godø and D. Tjøstheim

## Abstract

Diurnal variability in bottom trawl survey catches may influence the quality of the survey results when not taken properly into account. This is especially the case if diurnal effects vary substantially from year to year. In this study, catches of cod from the winter and autumn surveys in the Barents Sea during the years 1985-1999 were explored using both parametric and nonparametric statistical methods. Diurnal variation in winter has a smooth threshold form with a high catch period at daytime lasting approximately 8 hours. Small cod seems to be triggered more by the clock than by the time of sunrise. In the autumn the diurnal variation is much less distinct. Both in winter and autumn the effect tends to decrease with fish size and increase with depth. Diurnal effects on small cod vary substantially from year to year, whereas larger fish show a higher degree of stability. We suggest a bootstrap based method for adjusting the catches for diurnal variation. This can be applied to the catches prior to the calculation of indices, and also includes measures of the uncertainty. Analyses are done by size groups and the effects of size dependent vertical migration dynamics and catching efficiency are discussed and considered in relation to the performance of the simultaneous acoustic survey.

Keywords: bottom trawl catches, cod, diurnal variation, modeling, year to year effects.

*Vidar Hjellvik and Olav Rune Godø: Institute of Marine Research, P.O. Box 1870, Nordnes, N-5817 Bergen, Norway [tel: +47 55 23 85 00, fax: +47 55 23 86 87, email: vidarh@imr.no, olavrune@imr.no]. Dag Tjøstheim: Department of Mathematics, University of Bergen, Johs. Brunsgt 12, N-5008 Bergen, Norway [tel: +47 55 58 28 85, fax: +47 55 58 96 72, email: dagt@mi.uib.no]*

# 1 Introduction

Diurnal variability in survey results is known to affect both bottom trawl and acoustic surveys. Approaches to avoid negative effects on survey results do vary and are in many cases determined more by practical circumstances related to operation than a real understanding of diurnal variability. Survey vessel time is expensive and optimal use of available resources is crucial for the quality of the survey results (Pennington and Vølstad 1991). Often, trawl catches are higher during day, and many surveys are conducted at day only to avoid diurnal effects and under the assumption that the highest catches give the most representative picture of the stock (Wakabayashi et al. 1985). For some flatfish species, however, night catches are higher than day catches and affect survey results accordingly (Walsh 1988, Casey and Myers 1998). When arrangements on survey vessels permit operation around the clock, surveys might be carried out on a continuous basis (Doubleday 1981, Jakobsen et al. 1997) under the assumption that diurnal effects influence the surveys equally from year to year and more observations reflect the stock better.

Efficiency of trawl surveys is dependent on the effectiveness of the trawl to catch available fish and further, the availability of target fish to the trawl (Godø 1994). Both features might be size dependent (Aglen et al. 1999), and hence, disentangling diurnal effects in trawl catches is a complex matter where behaviour and vertical migration patterns of the fish in relation to variation in the environmental stimuli need to be understood. A full comprehension is probably not obtainable without methodology that integrates information about the vertical density profile from acoustics including behavioural characteristics induced by the survey vessel (see e.g. Ona and Godø 1990). As both efficiency of trawls and vertical behaviour patterns are potentially seriously diurnally affected, a combination of results demands an appropriate statistical handling and modelling of the data.

Both the above survey strategies ('around the clock' and 'day' design) might seriously affect results if assumptions related to diurnal effects fail to be true, or if design is not properly balanced among years (distribution of day and night station by stratum, depth etc., see e.g. Engås and Soldal 1992, to avoid diurnal related effects over the time series). In this paper historic survey catch data of North-East Arctic cod (Gadus morhua L.) from the Norwegian standardised surveys for cod and haddock (Jakobsen et al. 1997) are analysed with respect to diurnal variation. The goal is to design a general model for adjusting of trawl catches by fish size. Detailed analyses are accomplished to uncover annual, depth, and geographic effects on the model. Results are evaluated according to effects on measures of precision of the abundance estimates from catches.

Our first task is to determine the nature of the diurnal variation. For example, does it change continuously as time passes from midnight to noon and from noon to midnight? Or is it a threshold effect with an abrupt change as the sun rises and a symmetric reversed effect as it sets? Or perhaps the threshold occurs at a given time of the day independent of the time of the sunrise which varies a lot during the winter survey. This would be the case if fish behaviour is triggered by some biological clock rather than directly by light. A third possibility is fairly stable catches during day and night time but with a transition period, the extent of which should be determined. For all of these three scenarios the goal is to establish a simple parametric function describing the changes. In Section 2 we suggest some models, which are further explored in Section 3, utilising all of the data described in Section 2, but distinguishing between the winter survey and the autumn survey because of the different light conditions.

The parametric description obtained in Section 3 is important in that it enables a fairly complete investigation of possible changes in the diurnal pattern from year to year and its dependence, if any, on the length distribution of the fish, the depth at which the catches are taken, the geographical position, and the light conditions. Using the parametric representation, such a study can be done in terms of a few parameters instead of a visual comparison of patterns. It also facilitates use of statistical tests for examining the significance of variations in diurnal patterns. We look at these problems in Section 4, again separately for the winter and autumn data.

In Section 5 we discuss how the results from Section 3 and 4 can be used to adjust resource estimates for diurnal variation. This include quantitative estimates of the diurnal effect as a function of the explanatory variables of the preceding paragraph. Both the bias and the estimation errors will be considered.

In Sections 3-5 we have only been concerned with bottom trawl catches of cod, but the techniques are general and can be adapted to cover other species, and to analyse diurnal fluctuations of acoustic surveys. In Section 6 we discuss the findings of this paper and point out further possibilities.

# 2 Materials and methods

## 2.1 The data

Combined acoustic and bottom-trawl surveys for demersal fish in the Barents sea are conducted annually in Winter (January - March) and in Autumn (August - September) by The

Institute of Marine Research, Bergen (IMR). There are several target species, but in this work we concentrate on cod. We use data from 1985 - 1999 for the winter survey and from 1985 to 1998 for the Autumn survey. The geographical coverage and the time span of the surveys are given in Figure 1. For the winter survey the time span does not vary much during the period, but there have been some changes in covered area. In 1993 the area was expanded north-eastwards since by then it was clear that parts of the small fish was situated in this 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 partially remained uncovered due to ice conditions. In 1994 there was also a change in equipment, as the mesh size was reduced from 40 to 22 mm to enable a larger catch of small fish. Further description is given by Jacobsen et al. 1997. For the autumn surveys the time span varies substantially, and in 1995 the area covered was substantially expanded in direction southeast in an attempt to cover the whole stock.

Usually abundance indices are computed for each 5 cm interval in length, but for our purpose of exploring the dependence of diurnal variation on fish size, we have used the rather coarse division into the 3 length groups (length in cm) 7-15; 16-31 and 32-90. We started out by the finer division 7-10; 11-15; 16-22; 23-31; 32-44; 45-63; 64-90, which is roughly the same as in Korsbrekke and Nakken (1999), and for which the relative range of length groups $(L_{\max} - L_{\min})/L_{\text{mean}}$ is approximately constant, but found that most of the dynamics is taken care of by subsuming these into the 3 groups mentioned above.

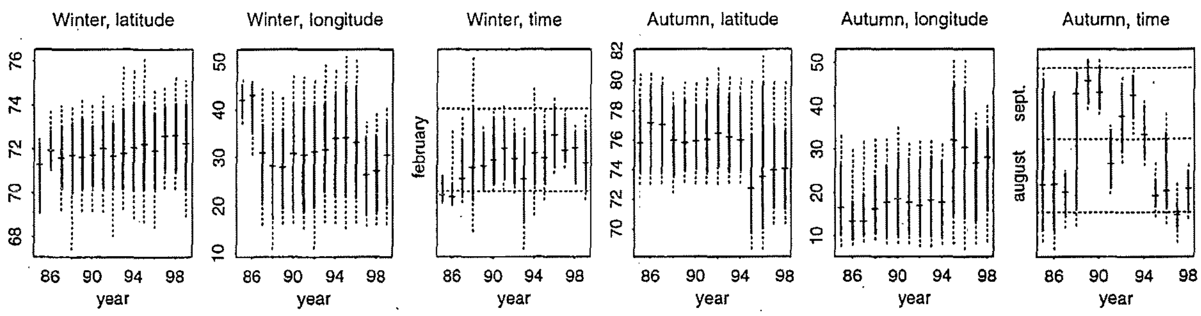


**Figure 1** *Geographical span and time span of the winter and autumn surveys. All non-zero catches of 7-90 cm cod were taken within the area/time indicated in the figure, and 80 % where taken within the intervals indicated by solid lines. The medians are indicated by horizontal bars.*

## 2.2 The model

Log transformed data were used to reduce heterogeneity of the variance. Time of the day $t$ ($0 \leq t < 24$), i.e. local time, is given by $t = t_{\mathrm{UTC}} - \mathrm{longitude}/15$ where $t_{\mathrm{UTC}}$ is the time in Universal Time Coordinates. Let the index $i$ refer to the $i$-th haul, and further, let $D_i$ be the towed distance and $N_i$ the number of cods caught for the $i$-th haul. We make the simplifying assumption that the total variation in the fish density is made up by diurnal variation, a day-to-day variation and random noise. Thus, $Y_i = \log(N_i/D_i)$ is given by the model

$$Y_i = \mu_{d(i)} + f(t_i) + \epsilon_i \tag{2.1}$$

where $t_i$ is the local time of the $i$-th haul, and $f$ is a function describing the diurnal variation. Moreover, $\mu_{d(i)}$ is the day-time level on day $d(i)$ when haul $i$ is done. Here $d(i)$ varies between 1 and $n_d$, $n_d$ being the number of days with hauls. In a multi-vessel operation each vessel is treated separately when modelling $\mu_{d(i)}$. For example, if two vessels are operating over four days, we have $n_d = 8$. In an alternative model local time $t_i$ in the argument of $f$ is replaced by the altitude $s_i$ of the sun at the $i$-th haul. Both $t_i$ and $s_i$ refer to the start of the haul, which typically has a duration of 30 minutes. Note that the function $f$ is normalised so that $f(t) = 0$ for $t = 12$, which means that $\mu_{d(i)}$ can be thought of as the expected value of $Y$ of day $d(i)$ at noon.

To examine the nature of diurnal variation the shape of the function $f$ is needed. Initially, this can be done nonparametrically with a pure curve fitting method (cf. Section 3), which in turn can be used to suggest parametric functional forms. Potential candidates for such functions are sine functions, threshold functions with thresholds near sunrise and sunset, or smooth threshold functions allowing for a transition period between night-time and day-time behaviour. Exploratory studies (see also Section 3) show that the diurnal variation is essentially symmetric around noon, and hence we have used symmetric models. The sine model is then given by

$$\text{ST: } f(t) = f(t; D) = \frac{D}{2} \sin\left\{\frac{(t-6)\pi}{12}\right\} - \frac{D}{2}, \quad 0 \leq t < 24, \tag{2.2}$$

with amplitude $D/2$. Threshold type behaviour is modelled by a symmetric logistic function, i.e.

$$\text{LT: } f(t) = f(t; B, C, D) = \begin{cases} g(t; B, C, D) - g(12; B, C, D), & 0 \leq t \leq 12 \\ g(24 - t; B, C, D) - g(12; B, C, D), & 12 > t > 24 \end{cases} \tag{2.3}$$

where $g$ is the logistic function given by

$$g(x; B, C, D) = D \, \frac{e^{(x-B)/C}}{1 + e^{(x-B)/C}}. \tag{2.4}$$

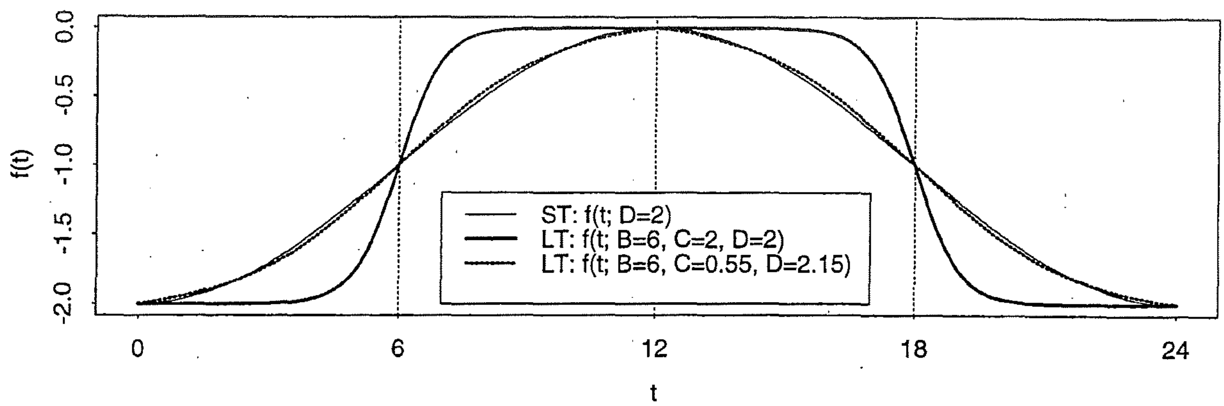


**Figure 2** *The functions defined by (2.2) and (2.3).*

| $t$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Winter, $F$: | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.05 | 0.28 | 0.66 | 0.91 | 0.95 |
| Autumn, $F$: | 0.32 | 0.35 | 0.44 | 0.55 | 0.60 | 0.75 | 0.90 | 1 | 1 | 1 | 1 | 1 | 1 |

**Table 1** *Cumulative relative frequencies $F$ of hauls with 7-90 cm fish taken at days and latitudes where the sun rises before time $t$. The table is based on all surveys from 1985-1999. For the autumn survey $t < 0$ indicate midnight sun.*

The parameter $B$ is an indicator of the time when the transition from night-level to day-level takes place in the sense that $g(x = B; B, C, D) = D/2$. The parameter $C$ determines the duration of the transition. Large values of $C$ give short transition phases, and as $C \to \infty$ a pure threshold function is obtained. For reasonable values of $B$ and $C$ the difference between day- and night-level is equal to $D$, i.e. the ratio in (2.4) is approximately equal to 1 and 0 for $x = 12$ and $x = 0$, respectively, but for small values of $C$ this is not the case. This is illustrated in Figure 2, where $f(t; B, C, D)$ of (2.3) is drawn for two different parameter sets, and compared to the sine function $f(t; D)$ of (2.2). For the dotted line we have $C = 0.55$ and $D = 2.15$, i.e. $D > f(12; B, C, D) - f(0; B, C, D) \approx 2$. Thus, for small values of $C$, the parameter $D$ does not properly describe the difference between day- and night-level, but for the data we have considered this is not a problem. The dotted line also demonstrate that $f(t; B, C, D)$ for certain parameter values has a shape very similar to the sinusoid.

In equations (2.2) and (2.3) $f$ is independent of the number of hours with sun above the horizon on day $d(i)$, Consequently, days in the beginning of the winter survey (when the north-most stations have polar night) and at its end (when the sun is above the horizon for more than 8 hours) are modeled in the same way. As can be seen from Table 1 approximately 5% of the hauls are taken under each of these extreme conditions. A refinement of model (2.3) taking the time of sunrise into account is presented in equation (2.9). Alternatively, the altitude $s$ of the sun can be used as an argument in $f$ as in the model

$$\text{LS: } f(s; B, C, D) = g(s; B, C, D). \tag{2.5}$$

There is also an underlying symmetry assumption in the LS-model in that we do not distinguish between the cases where $s$ is observed before and after noon.

The parameters to be estimated in the models (2.1)-(2.3) and (2.5) are $\mu_1, \ldots, \mu_{n_d}$, $B$, $C$ and $D$. The estimation is done by nonlinear least squares, using the S+ routine `nls` (cf. e.g. Venables and Ripley 1997, Sec. 9). This routine yields standard errors as well, which have been used as a basis for forming confidence intervals for $B$ and $D$, assuming a Gaussian approximation for the distribution of the parameter estimates.

To study effects of covariates such as depth, latitude, longitude and light, one may stratify the data in two or more groups according to the value of the covariate in question. The groups are then analysed separately and the resulting estimates of $B$, $C$ and $D$ of (2.2) and (2.3) are compared.

An alternative way of describing dependencies on covariates is to undertake a regression analysis. We assume that $f$ depends on $d(i)$ (cf. the above discussion), but restrict the dependence to the parameter $D$, which measures the strength of the diurnal variations and hence is of main interest. That is, the functional form of $f$ is the same, but the scaling as measured by $D$ may differ from one day to another so that

$$f_{d(i)}(t) = f(t; B, C, D_{d(i)}). \tag{2.6}$$

For example, potentially a day, $d(i)$, with a large average depth could have more extensive diurnal variations than a day with a small average depth. To examine the dependence in more detail, $D_{d(i)}$ is regressed on the explanatory variables $x_1, \ldots, x_p$, with an error term $e_{d(i)}$, so that

$$D_{d(i)} = b_0 + b_1 \overline{x}_{1,d(i)} + \ldots + b_p \overline{x}_{p,d(i)} + e_{d(i)} \tag{2.7}$$

where

$$\overline{x}_{k,d(i)} = \frac{1}{N_{d(i)}} \sum_{j=1}^{N_{d(i)}} x^j_{k,d(i)} \tag{2.8}$$

is the average of $x_k$ over the number of hauls $N_{d(i)}$ of day $d(i)$ with $x^j_{k,d(i)}$, $j = 1, \ldots N_{d(i)}$, being the value of $x_k$ for haul $j$ on day $d(i)$. Doing a regression of $Y_i$ directly on $\mu_{d(i)}, D_{d(i)}, d(i) = 1, \ldots, n_d$, $B$ and $C$ will lead to a regression problem with $2n_d + 2$ parameters to estimate. In principle the estimated $D_{d(i)}$-s could then be inserted in (2.7) to do a regression on $x_1, \ldots x_p$. However, since $n_d$ is of the order of 500 day units for both the winter survey and the autumn survey when the whole data material is considered, we will have too many parameters. Instead, we have chosen to insert the expression (2.7) for $D_{d(i)}$ into (2.6) to obtain the nonlinear regression equation

$$Y_i = \mu_{d(i)} + f(t_i; B, C, b_0 + b_1\overline{x}_{1,d(i)} + \ldots + b_p\overline{x}_{p,d(i)} + e_{d(i)}) + \epsilon_i$$

which only has $n_d + p + 3$ parameters, namely $\mu_1, \ldots, \mu_{n_d}, B, C, b_0, b_1, \ldots, b_p$. Since the function $f$ in (2.2) and (2.3) is linear in $D$, we can rewrite this as

$$Y_i = \mu_{d(i)} + f(t_i; B, C, b_0 + b_1\overline{x}_{1,d(i)} + \ldots + b_p\overline{x}_{p,d(i)}) + \epsilon'_i$$

where now $\epsilon'_i$ will depend on $f$ and $t_i$ but not on the regression variables $x_1, \ldots, x_p$.

The regression principle we have applied above to $D$, will also be applied to $B$ to examine to which extent fish behaviour is triggered by the light level or the clock. Let $\overline{x}_{d(i)}$ be the time of the sunrise on day $d(i)$ averaged over the visited stations. We then insert the simple linear regression model

$$B_{d(i)} = b_0 + b_1\overline{x}_{d(i)} + e_{d(i)} \tag{2.9}$$

into $f_{d(i)}(t) = f(t; B_{d(i)}, C, D)$ (cf. (2.6)) if model LT is used, or into $f_{d(i)}(s) = f(s; B_{d(i)}, C, D)$ if model LS is used.

## 2.3 Adjusting catches for diurnal variation

Diurnal variation in the catches reflects differences in availability of the fish to the trawl, and clearly not in the fish density itself. Thus, if the the catches taken at night on average are significantly lower than those taken at day, this should be compensated for by appropriately adjusting the night catches. This can easily be done if $f$ is known. In the simplest case, where $f$ is assumed to be day-independent, and using the models (2.2) or (2.3), the function $f(t)$ normally has its maximum for $t = 12$, the exception being some cases in the autumn surveys where night catches are significantly higher than day catches. In the latter cases the function $f$ of equations (2.2) and (2.3) should have its maximum for $t = 0$. This can be achieved by letting $D$ be negative (or by redefining $f$ as $f' = -f$). The adjusted catches in the two situations are defined by

$$Y_{i,\text{adj}} = \begin{cases} Y_i + f(12) - f(t_i), & 0 \le t_i < 24, \quad f(12) \ge f(0) \\ Y_i + f(0) - f(t_i), & 0 \le t_i < 24, \quad f(12) < f(0). \end{cases} \tag{2.10}$$

In practice, of course, $f$ has to be replaced by its estimate $\hat{f}$. Ultimately we are interested in determining how the yearly abundance indices are influenced by adjusting the catches for diurnal variation, but in this paper we will restrict ourselves to a comparison of the catches; more precisely to the medians of the distributions of $X = \exp(Y)$. We prefer the median

$M(X)$ to the mean $E(X)$ both because the median is more robust and because it can be found directly by taking antilogarithms of $M(Y)$, whereas $\exp(\overline{Y})$ is not an unbiased estimate of $E\{\exp(Y)\}$. To compare $M(X)$ with $M(X_{\text{adj}})$, we need an estimate of the uncertainty involved. This is obtained by bootstrapping. We start by estimating $\mu_{d(i)}$ and $f$ of (2.1) and computing the residuals as $\hat{\epsilon}_i = Y_i - \hat{\mu}_{d(i)} - \hat{f}(t_i)$. Then $\{\hat{\epsilon}_i\}$ is centralised by subtracting the mean of $\{\hat{\epsilon}_i\}$ and bootstrap replicas $\{\epsilon_i^\star\}$ and $Y_i^\star = \hat{\mu}_{d(i)} + \hat{f}(t_i) + \epsilon_i^\star$ are formed. Further $\mu_{d(i)}^\star$ and $f^\star(t_i)$ are fitted to $\{Y_i^\star\}$, and $Y_{i,\text{adj}}^\star$ is calculated by replacing $f$ by $f^\star$ in (2.10). Taking sufficiently many bootstrap replicas we get the bootstrap distribution of $Y_{i,\text{adj}}$, and hence of $X_{i,\text{adj}} = \exp(Y_{i,\text{adj}})$.

If diurnal variation is modelled on a daily basis by $f_{d(i)}(t)$, we simply replace $f$ by $f_{d(i)}$ in (2.10) and in the bootstrap algorithm.

## 2.4 Additional remarks

Some years, there are many hauls with zero catches within one or more of the length groups. These hauls are distributed more or less evenly throughout the 24 hour cycle, and since they represent an anomaly in the distribution of the catches, they are eliminated from the analysis. Another reason for eliminating them is that the reason for getting zero catches may be that in fact there are no fishes of a certain length group in certain areas, and hence no diurnal variation. If included, they would contribute to the estimate of $f$ just as much as the non-zero catches, and the diurnal variation as measured by $D$ would be underestimated.

Since the level $\mu_{d(i)}$ is fitted individually for each day, we obtain an exact fit with $\epsilon_i = 0$ in (2.1) on days with only one haul, and these days do not contribute meaningfully to the estimation of $f$. Consequently they are excluded from the analysis.

A few observations (4 for the winter data and 4 for the autumn data) are taken at depths larger than 600 meters. These we consider as atypical and we exclude them from the material when $f_{d(i)}$ is applied and depth is included as a covariate.

# 3 The general nature of the diurnal variation

## 3.1 The form of the function $f$ of (2.1)

In this section we investigate the general nature of the diurnal variation parametrically by means of the models described in Section 2 as well as by nonparametric estimates of the function $f$ of (2.1). Assuming that $f$ takes roughly the same form each year, it can best be investigated by merging data from all years into one large data set. For the nonparametric approximation, we eliminate the effect of $\mu_{d(i)}$ by subtracting first the daily averages from the catches to remove the day-to-day variation. That is, we use $\{Y_i - \overline{Y}_{d(i)}\}$ instead of $\{Y_i\}$, where $\overline{Y}_{d(i)}$ is defined in the same way as $\overline{x}_{d(i)}$ in (2.8). As regards the parametric analysis, we perform this on $\{Y_i\}$.

**Winter:** We start by looking at the winter data. In Figure 3 the catches $\{y_i - \overline{y}_{d(i)}\}$ of 7-90 cm fish are shown as a function of local time, Figure 3 a), and as a function of the altitude of the sun $s$, Figure 3 b), with nonparametric estimates of $f$, as produced by the S+ function `smooth.spline` (cf. Venables and Ripley 1997, p. 326). The large dots

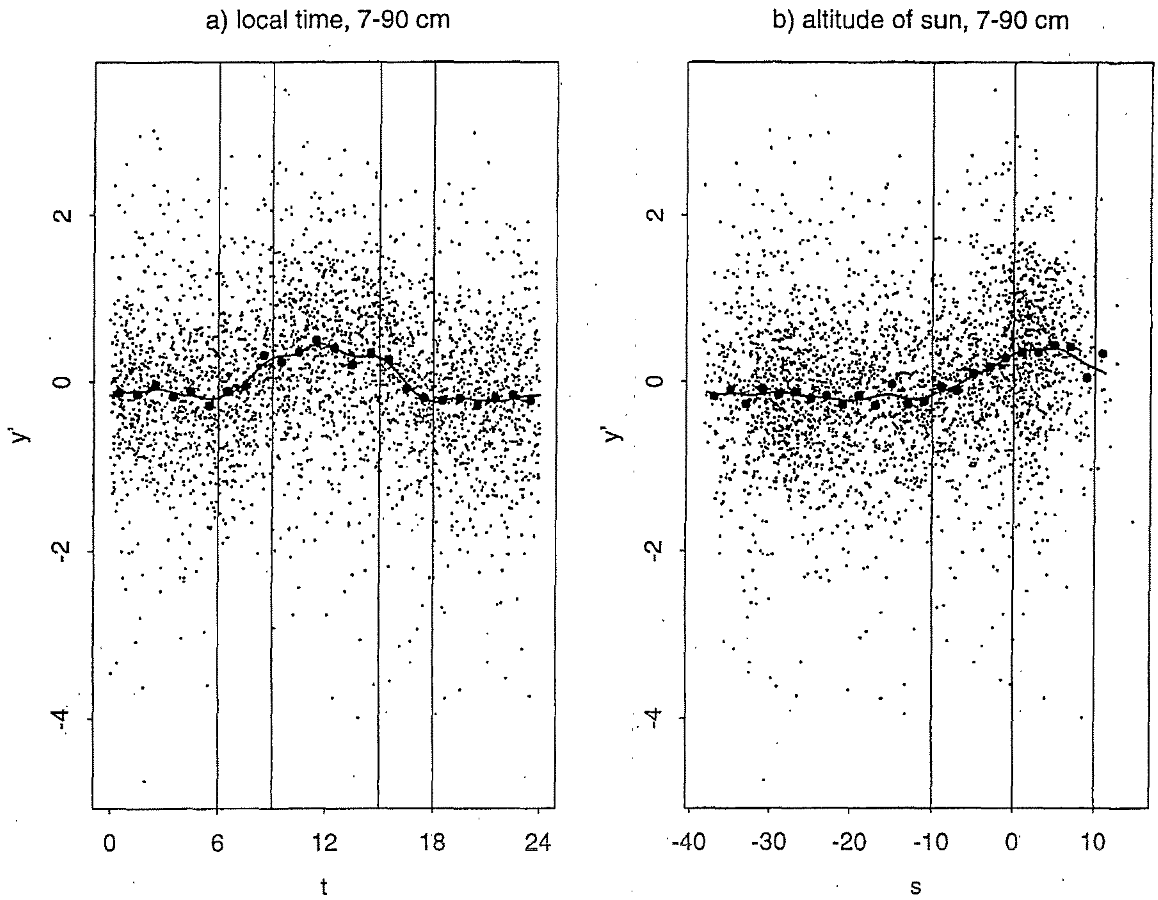


**Figure 3** *Mean-adjusted catches $y_i - \overline{y}_{d(i)}$ for the length group 7-90 cm from all winter surveys from 1985 to 1999 plotted against time (a) and altitude of the sun (b). The large dots in a) and b) represent the averages $\{\overline{y}_{j-1<t<j}, j = 1, \ldots, 24\}$, and $\{\overline{y}_{j-2<s<j}, j = -36, -34, \ldots, 12\}$, respectively. The solid lines are nonparametric estimates of the underlying relationships.*

signify hourly averages. Thus the leftmost big dot of Figure 3 a) is the average of all catches $y_i - \overline{y}_{d(i)}$ taken between midnight and one o'clock. If we denote this by $\overline{y}_{0<t<1}$, then in general $\overline{y}_{j-1<t<j}$ denotes the average of the catches $y_i - \overline{y}_{d(i)}$ taken at time $t_i$, where $j - 1 \leq t_i < j$, $j = 1, \ldots, 24$. In the right hand plot we use intervals of 2 degrees for $s$, so the averages shown in the plot are $\{\overline{y}_{j-2<s<j}, \quad j = -36, -34, \ldots, 12\}$. Since there are very few observations with $s_i < -38$ or $s_i > 12$, they are ignored in fitting the nonparametric curve to reduce end effects. Nevertheless the curve makes a dip at its rightmost end.

As a first impression, the nonparametric curve in Figure 3 a) indicates symmetry around noon. Further, the diurnal variation is relatively small compared to the total variation of $\{y_i - \overline{y}_{d(i)}\}$.

As diurnal behaviour is size related (Aglen et al., 1999), Figure 3 should be interpreted with care. Moreover, for each fixed local time point $t$ of Figure 3 a), the light intensity will depend on latitude and date. Figure 3 b) is more in accordance with an approach supposing that the fish is triggered by the light level rather than by the clock. The line is flat at night time and move through a transition zone to a day time plateau. (The behaviour at the extreme right is based on few observations and could be attributed to boundary effects). The transitional zone appears to start when the sun is about 10° under the horizon and seems to be completed by sunrise.

The same quantitative behaviour can be read from Figure 3 a). It is seen from Table 1 that for the winter survey, sunrise always occurs after 7, and for two thirds of the observations it occurs before 10. Correspondingly, on Figure 3 a) we have a flat night time section between $0 \leq t \leq 6$ and $18 \leq t < 24$; there are transitional phases for $t$ roughly between 6 and 9 and 15 and 18, and an indication of a plateau phase between 9 and 15. Figure 3 a) suggests that the logistic function in (2.3) might yield a better fit than the sinusoid of (2.3), and the curve in Figure 3 b) also has a logistic-like shape. This will now be examined in more detail for length stratified data.

In the three leftmost plots of Figure 4 we have stratified the winter data on length and fitted the parametric models LT, ST and LS to the data, as well as nonparametric curves. The last ones are fitted to $\{y'_i = y_i - \hat{\mu}_{d(i)}\}$, where $\{\hat{\mu}_{d(i)}\}$ are the estimates from the parametric fit, and to capture the dynamics of the transition phase better, the curves are somewhat less smoothed than those of Figure 3. The estimates of the parameters $B$, $C$ and $D$ of equations (2.2), (2.3) and (2.5) are given in the plots, as well as the R-squared values $R^2$, $R^{2\prime}$ and $R^{2\prime\prime}$. The first of these quantifies how much the full model (2.1) explains of the variation of $\{y_i\}$, and is defined by

$$R^2 = \frac{\left\{\sum_{i=1}^{n}(y_i - \overline{y}_i)(\hat{y}_i - \overline{y}_i)\right\}^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2 \sum_{i=1}^{n}(\hat{y}_i - \overline{y}_i)^2} \tag{3.1}$$

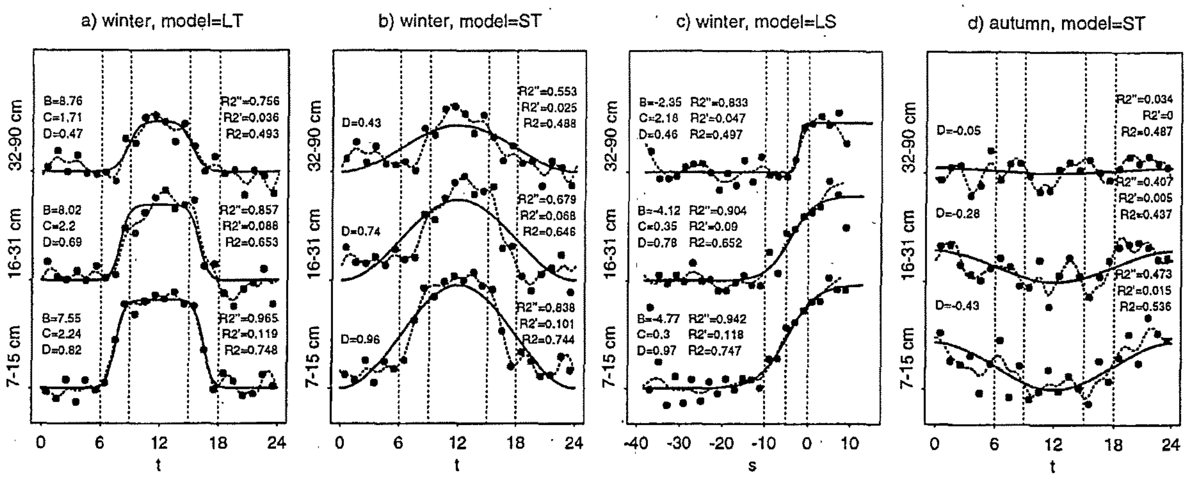


**Figure 4** *Solid lines: parametric estimates $\hat{f}$ of $f$ in (2.1) based on all catches from 1985-1999, stratified on length groups. The parameter estimates are given on the left hand side of each plot. Dotted lines: nonparametric smoothing spline estimates. Dots: $\{\overline{y}'_{j-1\le t<j}, j = 0,\ldots,23\}$ for models LT and ST, and $\{\overline{y}'_{j-2\le s<j}, j = -36,-34,\ldots,10\}$ for model LS. For all length groups and models the curves are plotted on the same scale, but they are vertically adjusted. For each parametric curve, $\hat{f}(t) = 0$ for $t = 12$ and $\hat{f}(s) = 0$ for $s = 10$. The R-squared values $R^2$, $R^{2\prime}$ and $R^{2\prime\prime}$ are given at the right hand side of the plots.*

where $\hat{y}_i = \hat{f}(x_i) + \hat{\mu}_{d(i)}$ as in (2.1) with $x_i = t_i$ if LT or ST is used and $x_i = s_i$ if LS is used. Since we primarily are interested in how much $f$ explains of the variation of $\{y'_i\}$, we also calculate $R^{2\prime}$ by replacing all entries of $y_i$, $\overline{y}_i$ and $\hat{y}_i$ in (3.1) by $y'_i$, $\overline{y}'_i$ and $\hat{y}'_i$, respectively, where $y'_i$ is defined above and $\hat{y}'_i = \hat{f}(x_i)$.

For the LT and ST models we compute R-squared values based on the hourly averages as well. These are denoted by $R^{2\prime\prime}$ and are found as in (3.1) by replacing each observation $y_i$ by the corresponding hourly average of $y'_i$. Thus, if the time $t_i$ of haul $i$ satisfies $j-1 \le t_i < j$, we define $y''_i = \overline{y}'_{j-1<t<j}$, where $\overline{y}'_{j-1<t<j}$ is defined in analogy with $\overline{y}_{j-1<t<j}$ above. The corresponding fitted value $\hat{y}'_i = \hat{f}(t_i)$ is replaced by the mid-hour value $\hat{y}''_i = \hat{f}\{(2j-1)/2\}$, and $\overline{y}_i$ is replaced by $\overline{y}''_i$. For the LS model $R^{2\prime\prime}$ is constructed by averaging over two degrees intervals of $s$. The main purpose of $R^{2\prime\prime}$ is that it gives a comparison between the parametric and a nonparametric estimate, since the hourly averages (large dots) can be interpreted as a coarse nonparametric smoothing. Alternatively we could have used $R^{2\prime\prime}$-values based on the nonparametric curve itself. Since $\{\overline{y}'_{j-1<t<j}\}$ differs from $\{\overline{y}'_{j-2<s<j}\}$, $R^{2\prime\prime}$ for the time based models LT and ST are not directly comparable to $R^{2\prime\prime}$ for the sun based model LS. Thus $R^2$ and $R^{2\prime}$ should be used for comparison between LT and LS.

As can be seen from Figure 4 a), LT fits very well to the 7-15 cm length group, both as judged by $R^{2\prime\prime}$ and by visual comparison to the nonparametric curve. It also fits quite well to the 32-90 cm group. For the 16-32 cm group, which has a slightly asymmetric nonparametric estimate, we did not find starting values so that `nls` managed to do the minimisation for all of the 3 parameters. However, with a fixed value of $C$ there were no problem, and $C = 2.2$ which seems reasonable both as compared to the two other length groups and to the nonparametric estimate, yields the results shown in the figure. Both as judged visually and from the R-squared values, LT is clearly a better choice than ST for all 3 length groups. This confirms the tentative conclusion based on Figure 3 that the diurnal variation basically consists of a flat night-phase and day-phase and a transition phase between them. As regards LT, $D$ decreases with length, $C$ is more or less length independent whereas $B$ increases with half an hour for each length group. The estimated 95% confidence intervals for $D$ are [0.69, 0.95], [0.59, 0.80] and [0.31, 0.63] for length groups 7-15, 16-31 and 32-90, respectively. The results thus show that the diurnal effects significantly increase with reduced fish size.

For $B$ the corresponding confidence intervals for model LT are [7.2, 7.9], [7.6, 8.4] and [8.0, 9.6]. For $C$, the estimates have an asymmetric distribution and hence the confidence intervals based on the standard errors from `nls` are not presented.

<u>Autumn:</u> The light conditions during the autumn surveys are very different from those experienced during winter (cf. Table 1), and not unexpectedly, the character of the diurnal variation also changes, as shown in Figure 4 d). Data from 1985 to 1998 are used, and the fitted model is ST. Compared to the corresponding winter data in the leftmost plot, there are several differences. First, we see that for small fish, the highest catches in the autumn are taken at night. The 95 % confidence interval for $D$ for 7-15 cm fish is [-0.63, -0.24], so $D$ is significantly negative (by convention we allow the amplitude of the sine curve (2.2) to be negative instead of redefining its phase). It should be noted, though, that $R^{2''}$ only indicates a moderately good fit of the sine function and the value of $R^{2'}$ shows that the diurnal variation only explains 1.5% of the total day-level adjusted variation. Next, the magnitude of the diurnal variation is much smaller in the autumn than in the winter, with almost no variation for large fish. Thirdly, the nonparametric estimates do not have a clear threshold form as for the winter data. In fact we did not manage to fit LT for the two largest length groups, and for the smallest length group the fitted logistic curve looked very similar to the sine curve shown in the figure. LS is also problematic to fit, and no distinct threshold appears for this model as it does for the winter data.

## 3.2 Light versus clock as triggering factors of fish behaviour in winter

As mentioned above (cf. Table 1), the light conditions changes considerably during the time span of the winter survey. It has often been an implicit assumption that sunrise and sunset are the factors that trigger the behaviour of the fish; see e.g. Korsbrekke and Nakken (1999) where hauls taken when the sun was 5° or more below the horizon were defined as night hauls. If this is the case, the period with high catches will have a duration varying from 0 hours at the start of the winter survey to 8-9 hours at its end, and the model LT is obviously not optimal since in its simplest form it assumes a fixed duration of the high catch period.

Looking at the plot of the LS model in Figure 4 c), where the fish behaviour is measured as a function of the altitude of the sun, the most striking difference between the different length groups is, in addition to the decrease of $D$ with length, the steep gradient of the curve for 32-90 cm fish as compared to 16-31 and 7-15 cm fish. For 32-90 cm fish the transition phase spans an interval of approximately 3° whereas for smaller fish it spans an interval of about 15°. This suggests that the transition phase lasts longer for small fish. However, an alternative explanation is that the gradient is in fact sharper than indicated for small fish in the plots, and that it occurs at an approximately fixed local time $t_0$. In this case the estimated LS-curve of Figure 4 represents an average of a collection of LS-curves each having a relatively sharp transition, but at different values of $s$ corresponding to the altitudes $s_{d(i)} = s_{d(i)}(t_0)$ of the sun at day $d(i)$ at local time $t_0$.

This hypothesis can be investigated by means of model (2.9). Assume first that the fish is triggered by the clock and that the transition takes place at a fixed time $t = t_0$ each day. If model LT is used with the interpretation of the parameter $B$ given just after (2.4), this means that $B_{d(i)} \equiv t_0$ (cf. also Figure 2) independent of $\overline{x}_{d(i)}$, and consequently $b_1$ in (2.9) must be equal to zero. However, the altitude of the sun $s_0$ at time $t_0$ varies with $d(i)$ in the course of the survey. In the first part of the survey when the sun rises late and $\overline{x}_{d(i)}$ is large, $s_0$, and hence $B_{d(i)}$ if model LS is used, is smaller than in the end of the survey. Thus, $B_{d(i)}$ decreases as $\overline{x}_{d(i)}$ increases, and $b_1$ should be negative for the LS model. On the other hand, if the transition takes place at sunrise, i.e. at $s_0 = 0$ (or at a *fixed* value $s_0 \neq 0$), then $b_1 = 0$ in (2.9) if model LS is used and $b_1 > 0$ if model LT is used.

The results of applying (2.9), with $\overline{x}_{d(i)}$ set equal to 12.1 for days with polar night, are shown in Figure 5 both for models LS and LT. Instead of drawing one curve for each day, only the fitted values $\hat{f}_{d(i)}(s)$ and $\hat{f}_{d(i)}(t)$ are shown, but the range of $\hat{B}_{d(i)}$ is clearly seen. For model LS, for 16-31 cm fish, we see that $\hat{B}_{d(i)}$ ranges approximately from -10 for the darkest days to 0 for the lightest days. With $x = 7.5$ and $x = 12$ we get $\hat{b}_0 + \hat{b}_1 x = 1.4$ and -9.8, respectively, where $\hat{b}_0$ and $\hat{b}_1$ are given in the figure. For 7-15 cm fish the span is approximately the same, whereas for 32-90 cm fish there is practically no dependence of $\hat{B}_{d(i)}$ on the sunrise time. If this is correct, i.e. if the transition as measured by $B_{d(i)}$ takes

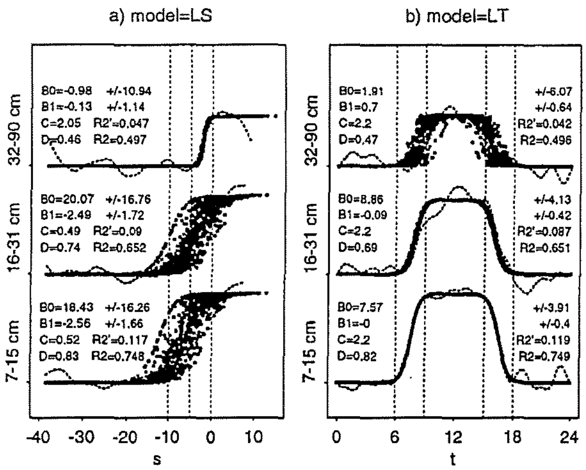


**Figure 5** *Models LS and LT with B regressed on sunrise according to equation (2.9). Winter data from 1985-1999 is used. For all hauls* $\{\hat{f}(t_i)\}$ *for model LS and* $\{\hat{f}(t_i)\}$ *for model LT are plotted as dots. The dotted lines are nonparametric smoothing spline estimates. Parameter estimates for* $b_0$ *and* $b_1$ *of equation (2.9) and for C and D of equations (2.3) and (2.2), with 95% confidence intervals for* $b_0$ *and* $b_1$*, are given, as well as* $R^2$ *and* $R^{2\prime}$.

place just before sunrise, independent of the clock, one would expect $\hat{b}_1 > 0$ for the LT model (cf. the discussion above). It is seen from Figure 5 b) that this is indeed the case, and the entire 95% confidence interval for $b_1$ is above zero. We also see that for the two smallest length groups, $B_{d(i)}$ is practically independent of the sunrise time for model LT, confirming the above presumption that $b_1 \approx 0$ for this case. Here the transition takes place at 7.30 and 8 o'clock for the 7-15 and 16-31 cm groups, respectively, throughout the whole survey.

For the LT model the `nls` algorithm required a fixed value of $C$ for it to be stable for the two largest length groups, so we have used a fixed $C = 2.2$ for all three groups. This value is motivated by the results in Figure 4 a). For the LS model we see that $\hat{C}$ is larger than on Figure 4 c) for the two smallest length groups, meaning that each individual transition curve is steeper, which is consistent with the hypothesis that for these two length groups the curves on Figure 4 c) represent average behaviour.

Comparing the estimates of $D$ in Figure 4 c) to those in Figure 4 a), we see that LS yields estimates approximately 15% higher than LT for the two smallest length groups. In Figure 5 this difference has vanished completely for 7-15 cm fish, whereas for 16-31 cm fish it is halved. This may imply that when we do the seemingly wrong assumption that the fish behaviour is triggered by light, and consequently use model LS, $D$ tends to be overestimated for these length groups. For the largest length group this effect is smaller.

To further explore the results described above, we stratified the winter data in two groups; those with sunrise before and after the median sunrise time, respectively, and fitted the

| | model=LS | | model=LT | |
|---|---|---|---|---|
| length gr. | light | dark | light | dark |
| 32-90 cm | $-2.61 \pm 2.21$ | $-1.77 \pm 1.82$ | $7.96 \pm 0.91$ | $9.46 \pm 0.92$ |
| 16-31 cm | $-3.42 \pm 2.11$ | $-6.69 \pm 3.36$ | $8.00 \pm 0.47$ | $8.03 \pm 0.79$ |
| 7-15 cm | $-4.68 \pm 1.94$ | $-8.83 \pm 3.17$ | $7.58 \pm 0.44$ | $7.47 \pm 0.72$ |

**Table 2** *Estimates of $B$ with 95% confidence intervals when models LT and LS are applied on the winter data stratified on the time of the sunrise. The "light" group contains hauls with $u_i < M(u_i)$ where $u_i$ denotes the time of the sunrise on the actual station at day $d(i)$, and $M$ denotes the median. For the LS model $C \equiv 0.5$ and 1.5 for the two smallest and the largest length group, respectively. For the LT model $C \equiv 2.2$ for all length groups.*

models LS and LT to the two groups separately without doing the regression of $B$ on the time of sunrise. In Table 2 the results are shown with $C$ equal to 2.2 for model LT and equal to 0.5 and 2 for the two smallest and the largest length group, respectively, for model LS. Thus, the results from the regression method are supported.

## 3.3 Dependency of diurnal variation on various covariates

In this subsection we investigate the dependency of the diurnal variation as measured by $D$ in equation (2.3) on depth, latitude, longitude and light. From the results of Section 3.2 model LT seems to be the best choice for length groups 7-15 and 16-31, whereas LS is to be preferred for the largest length group. However, for simplicity, we use LT for all three length groups, and in the rest of this section we keep $C$ fixed and equal to 2.2 to assure stability and thus get a better basis to investigate the dependence of $D$ on the various explanatory variables.

To give a rough impression of the significance of depth, the data from both seasons were split in two groups using the median depth for each season as stratification criterion. Diurnal variation for small fish appears to be strongly dependent on depth, whereas for large fish it is more or less depth independent (Figure 6).

We continue with a more detailed examination of the influence of depth, longitude, latitude and light on diurnal variation, applying the regression method described by equations (2.6) and (2.7) and the accompanying text in Section 2. We start with linear and polynomial regressions, treating each covariate separately. As a measure of light, the maximum altitude of the sun on day $d_{(i)}$ averaged over the visited stations is used for the winter data. For the autumn data the minimum altitude of the sun is used. Thus, negative values of the covariate indicate polar night in winter, and positive values indicate midnight sun in autumn. The fitted regression lines $D_{d(i)} = \hat{b}_0 + \hat{b}_1 x$ and regression curves $D_{d(i)} = \hat{b}_0 + \hat{b}_1 x + \hat{b}_2 x^2$ are shown

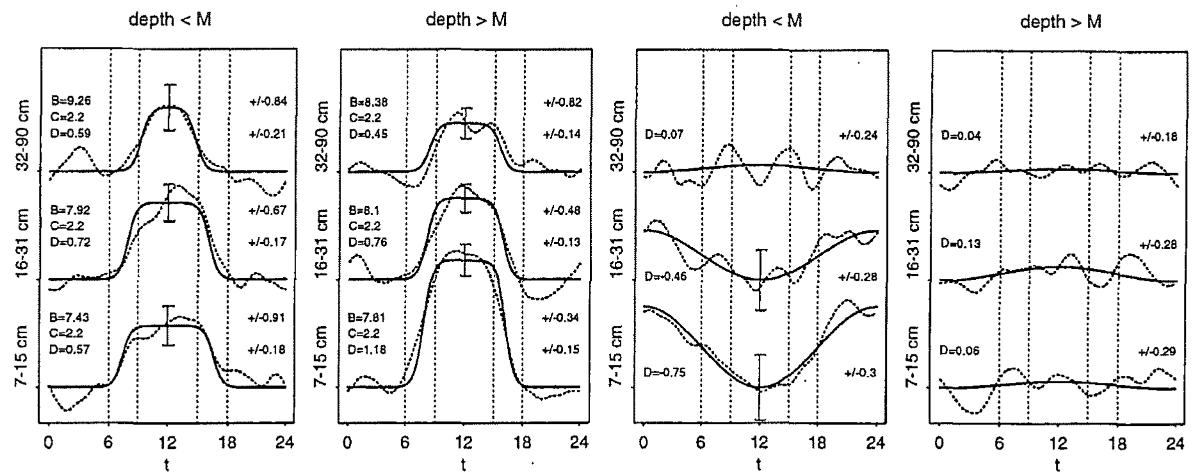


**Figure 6** *Parametric (winter: model LT with C fixed, autumn: model ST) and nonparametric estimates of diurnal variation with data stratified on depth, M being the median depth. In the significant cases, the estimated 95 % confidence intervals of D is drawn in the centre of the plots. The width of the intervals for both D and B (winter) are given on the right hand side of the plots.*

in Figure 7. The regression lines for depth in Figure 7 agree fairly well with the results of Figure 6. For the winter data there is no significant dependence on depth for fish larger than 15 cm, whereas there is a strong dependence for the smallest fish. The estimated regression parameters for this group is $\hat{b}_0 = -0.619$ and $\hat{b}_1 = 0.00528$ in the linear regression case. The $p$-value for $b_1$ is $3 \cdot 10^{-11}$. For $x = 223$, and 336, which corresponds to the 0.25 and 0.75 quantiles of depth for the whole data set, and thus approximately to the medians of the two subgroups of Figure 6 a) and b), we have $\hat{b}_0 + \hat{b}_1 x = 0.56$ and 1.16 in Figure 7, which should be compared to $\hat{D} = 0.57$ and 1.18 in Figure 6. For $x = 280$, which is the median depth, $\hat{b}_0 + \hat{b}_1 x = 0.86$, which should be compared to $\hat{D} = 0.82$ in the leftmost plot in Figure 4, so we see the agreement is good. The above described agreement between the results from the linear regression method and the stratification method of Figure 6, together with the fact that the polynomial regression yields a practically linear curve, indicate that the diurnal variation for small fish in the winter surveys in fact depends linearly on depth.

Since the difference between the 7-15 cm group and the 16-31 cm group are so big, we have undertaken a finer division in length groups to explore the dynamics in more detail. The estimates $\hat{b}_1$ resulting from the linear regression for the 7-10, 11-15 and 16-22 cm groups are then 0.0069, 0.0052, and 0.00053. It should however be noted that there were very few non-zero catches of 7-10 cm fish before 1994 when the reduction in mesh size took place.

Also for the autumn data there is a qualitative agreement between the results in Figure 6 and Figure 7. The regression lines and curves are mainly located below $D = 0$ and the strength

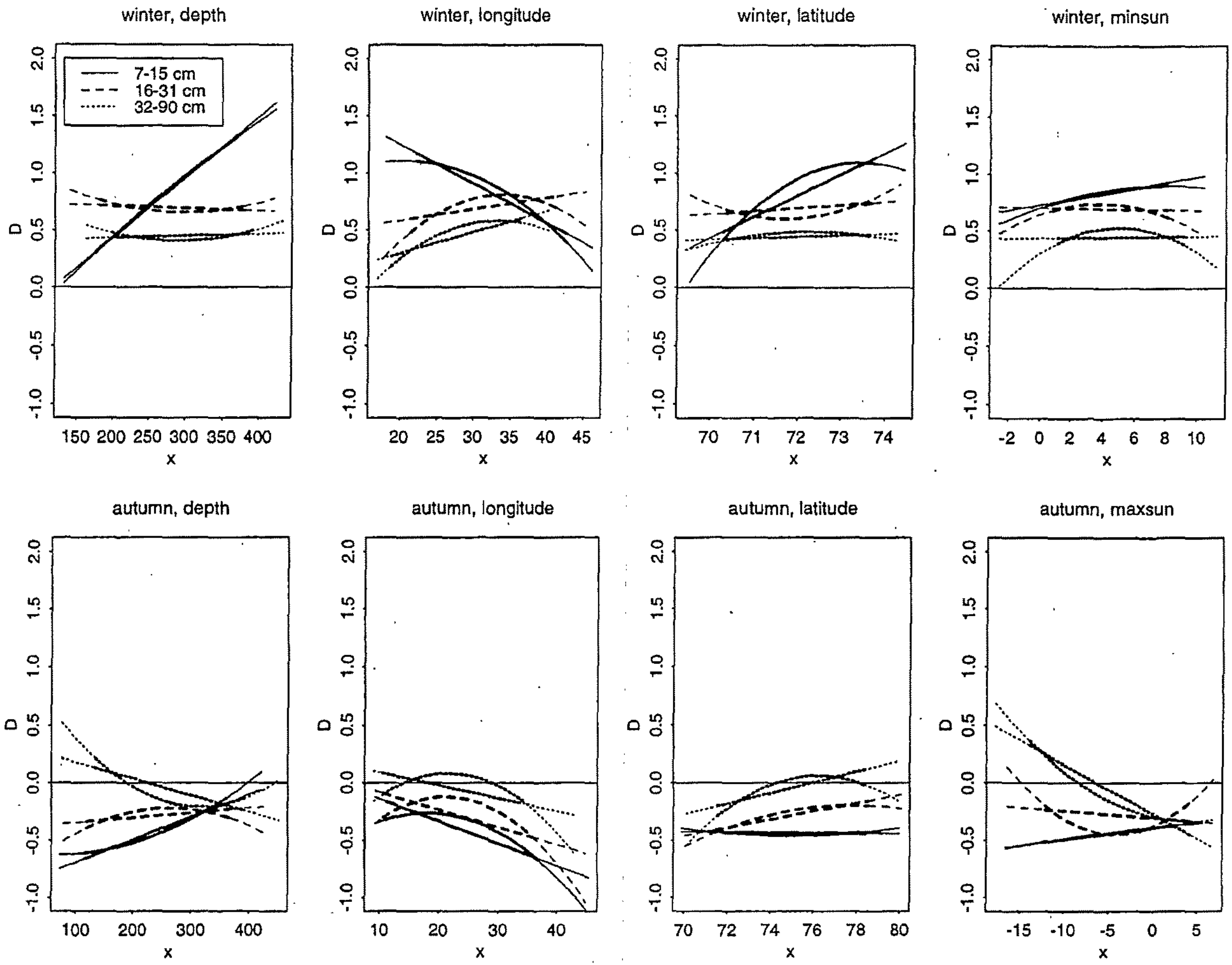


**Figure 7** *Estimated regression lines for linear and polynomial regressions of D on depth, longitude, latitude and light. The thick lines are limited by* $q_{0.10}$ *and* $q_{0.90}$ *and the thin lines by* $q_{0.025}$ *and* $q_{0.975}$*, where* $q_p$ *is the p-th quantile of the actual covariate. The model used for winter data is LT with* $C = 2.2$*, and for autumn data ST is used.*

of the diurnal variation is increasing with depth, except for the 32-90 cm length group, for which it is decreasing, also in accordance with Figure 6. The agreement between Figure 7 and Figure 6 are, however, not quite as good quantitatively as for the winter data. This may be due to the fact that the variation in depth per day is typically larger in the autumn than in the winter survey. It implies that hauls from the same day are more often split into the two groups of Figure 6 and that the regression method, which is based on daily averages, is less effective as compared to the winter survey. Optimally, for the depth regression problem all hauls taken at the same day should have been taken at the same depth.

Concerning latitude and longitude it should be noted that these covariates are relatively strongly correlated with depth in the winter survey, but not in the autumn survey. The estimated correlations for the winter survey are $\hat{r}(\text{depth}, \text{lat}) = 0.54$ and $\hat{r}(\text{depth}, \text{lon}) = -0.48$, whereas in the autumn survey $\hat{r}(\text{depth}, \text{lat}) = 0.02$ and $\hat{r}(\text{depth}, \text{lon}) = -0.15$. Thus

| | | Winter | | | | Autumn | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | depth | lon | lat | maxsun | depth | lon | lat | minsun |
| $b_1$ | 7-15 | 3.390 | -2.275 | 2.234 | 0.575 | 0.817 | -0.959 | -0.022 | 0.306 |
| | 16-31 | -0.148 | 0.657 | 0.328 | -0.059 | 0.188 | -0.777 | 0.487 | -0.209 |
| | 32-90 | 0.104 | 1.076 | 0.135 | 0.046 | -0.901 | -0.694 | 0.748 | -1.949 |
| $b_2$ | 7-15 | -0.205 | -0.818 | -1.309 | -0.311 | 0.239 | -0.639 | 0.105 | -0.023 |
| | 16-31 | 0.484 | -1.152 | 0.884 | -0.750 | -0.372 | -0.918 | -0.265 | 0.832 |
| | 32-90 | 0.447 | -0.764 | -0.351 | -1.161 | 0.835 | -1.002 | -0.839 | 0.628 |

**Table 3** *Parameter estimates for the regression lines in Figure 7, divided by half the length of the corresponding 95 % confidence intervals as estimated by* `nls` *for $b_1$ in the linear regression case and for $b_2$ in the polynomial regression case. Absolute values greater than one indicate estimates significantly different from 0.*

the apparent linear dependencies of diurnal variation on longitude and latitude for small fish in the winter may well be an artefact of its dependency on depth (cf. multiple regression analysis at the end). On the other hand, the dependency on longitude and latitude seems to be of a more nonlinear character than for depth.

In Table 3 the quantities $2\hat{b}_1/L_{\text{CI}}$ and $2\hat{b}_2/L_{\text{CI}}$ for all the linear and polynomial regressions, respectively, are given. $L_{\text{CI}}$ is here the length of a 95% confidence interval, so that an absolute value larger than 1 of the ratios in Table 3 indicate significance at a 5 % level. Converted to $p$-values, ratios of 2 and 3 corresponds to approximately $10^{-4}$ and $10^{-9}$, respectively, whereas the ratio 1 corresponds to 0.05. For depth, the only parameter that is significantly different from 0 is $b_1$ in the linear regression case for 7-15 cm fish.

Performing a multiple regression analysis on this group (7-15 cm), depth is the only covariate that remains significant. The estimated regression equation is $-1.3 + 0.0046x_1 - 0.0073x_2 + 0.015x_3 - 0.0045x_4$ where $x_1, \ldots, x_4$ are the daily averages of depth, longitude, latitude and maxsun, respectively. The $p$-values for the corresponding regression coefficients $\hat{b}_1, \ldots, \hat{b}_4$ are 0.00018, 0.50, 0.80 and 0.86. The values of $\hat{D}_{d(i)}$ ranges from -0.2 to 1.8, which is very close to the range of $\hat{D}_{d(i)}$ in Figure 7. Also $R^{2\prime} = 0.178$ both in the multiple linear regression and in the simple linear regression on depth. We also performed a multiple regression analysis where interaction between depth and longitude/latitude was allowed, as well as a second order polynomial multiple regression analysis. However, both of these analyses resulted in unstable parameter estimates in the sense that the width of the confidence intervals increased drastically. For example, the width of the confidence interval for the depth parameter $b_1$ increased from 0.0028 in the simple multiple model to 0.22 and 0.023 for the model with interaction and with second order polynomial terms, respectively. The results also depended to some degree on the initial values for `nls`. We therefore felt that the output could not be trusted, and the overall conclusion is that, when analysed simultaneously, depth for the 7-15 cm length group in the winter is the only clearly significant covariate.

# 4 Yearly differences in diurnal variation

Based on the results from Section 3 we now embark on the task of investigating possible year to year differences in diurnal variation, as measured by the parameter $D$ of (2.3) for the winter data and of (2.2) for the autumn data. Since the estimates of both $B$ and $C$ are rather unstable when we use data from one year only, especially in years with little diurnal variation, we keep both of these parameters fixed to get a better basis for comparing $D$ from year to year. We use $C = 2.2$ as before, and from Figure 4, $B = 8$ seems to be a reasonable choice for $B$.

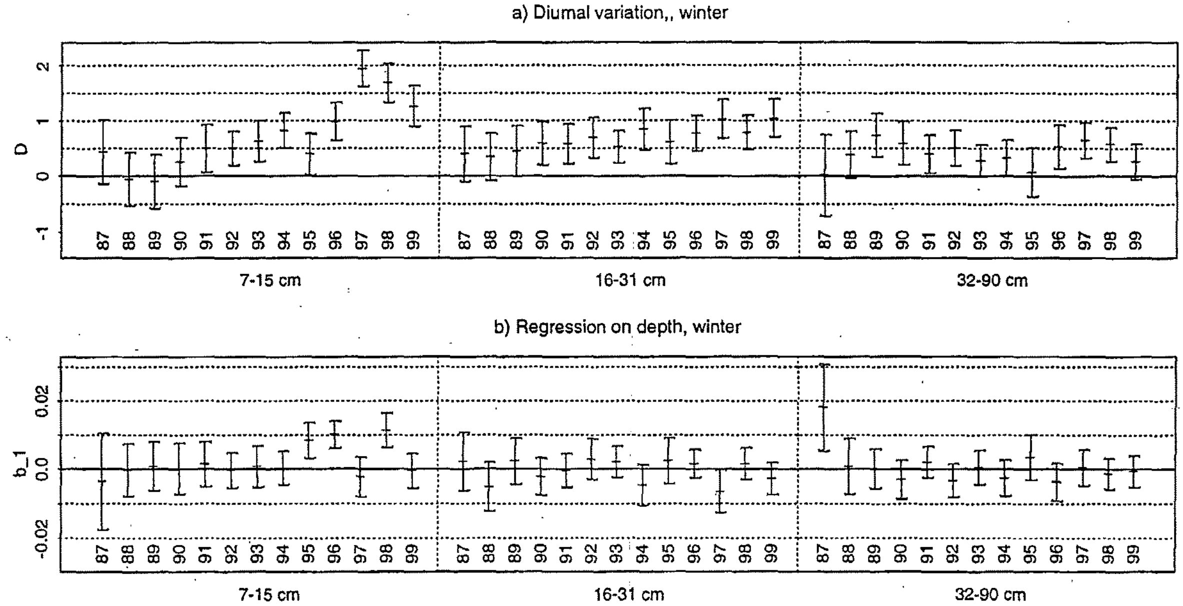


**Figure 8** *95% confidence intervals for $D$ in (2.3) (a) and for $b_1$ in (2.7) (b) for the winter survey. In both cases $B \equiv 8$ and $C \equiv 2.2$. The point estimates are indicated with horizontal bars.*

Winter: Figure 8 a) shows 95 % confidence intervals for $D$ in (2.3) for the years 1987 to 1999. Due to very few observations available in 1985 and 1986 (cf. the upper part of Table 4), these years are not included in the figure. As can be seen from Figure 8 a), the diurnal variation changes little from year to year for large and medium fish, whereas for small fish, there are clear differences between years, with the highest variation occurring the last three years. Looking at Figure 1 one may suspect that this might be affected by the east and southeast limitations of the area coverage these years. We therefore re-examined the years 1994-99 using data only from the area restricted by $70 <$ latitude $< 75$ and $14 <$ longitude $< 39$,

| | year: | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | tot |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | Winter | | | | | | | | |
| $N$ | | 23 | 38 | 191 | 247 | 177 | 223 | 300 | 223 | 245 | 288 | 292 | 315 | 177 | 197 | 224 | 3159 |
| $n$ | 7-15 | 14 | 29 | 59 | 22 | 60 | 143 | 213 | 178 | 201 | 261 | 273 | 303 | 155 | 181 | 193 | 2285 |
| | 16-31 | 22 | 19 | 164 | 180 | 123 | 144 | 236 | 182 | 206 | 237 | 241 | 279 | 156 | 178 | 205 | 2572 |
| | 32-90 | 4 | 2 | 129 | 185 | 149 | 148 | 222 | 204 | 224 | 210 | 177 | 231 | 164 | 188 | 191 | 2434 |
| $n_d$ | 7-15 | 5 | 7 | 21 | 10 | 18 | 27 | 46 | 31 | 45 | 59 | 59 | 61 | 28 | 38 | 41 | 496 |
| | 16-31 | 6 | 6 | 40 | 46 | 28 | 28 | 52 | 31 | 44 | 55 | 56 | 59 | 28 | 36 | 43 | 558 |
| | 32-90 | 2 | 1 | 32 | 49 | 31 | 25 | 49 | 34 | 47 | 47 | 41 | 49 | 30 | 40 | 40 | 516 |
| | | | | | | | | | Autumn | | | | | | | | |
| | year: | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | | total |
| $N$ | | 194 | 192 | 72 | 188 | 166 | 196 | 199 | 322 | 214 | 184 | 334 | 404 | 266 | 218 | | 3149 |
| $n$ | 7-15 | 60 | 31 | 0 | 8 | 40 | 64 | 80 | 196 | 174 | 120 | 271 | 351 | 221 | 164 | | 1780 |
| | 16-31 | 155 | 140 | 35 | 76 | 61 | 116 | 151 | 189 | 177 | 147 | 283 | 373 | 242 | 188 | | 2333 |
| | 32-90 | 169 | 176 | 52 | 145 | 156 | 180 | 171 | 216 | 205 | 175 | 291 | 307 | 214 | 196 | | 2653 |
| $n_d$ | 7-15 | 21 | 10 | 0 | 3 | 11 | 14 | 18 | 36 | 21 | 23 | 55 | 58 | 40 | 38 | | 347 |
| | 16-31 | 36 | 30 | 11 | 22 | 10 | 22 | 25 | 29 | 21 | 25 | 52 | 60 | 41 | 38 | | 422 |
| | 32-90 | 36 | 32 | 15 | 32 | 20 | 24 | 23 | 35 | 21 | 25 | 54 | 53 | 40 | 39 | | 449 |

**Table 4** *Total number of catches (N), number of non-zero catches taken at days with more than one non-zero haul (n) and number of such days ($n_d$) for all combinations of year and length-group.*

which is roughly the area covered for 1997-99. The resulting estimates are given in the second line of Table 5. We see that the differences are levelled out to some degree, but still $\hat{D}$ is largest the three last years. With even further reduction of the area as defined in the last line of Table 5, $\hat{D}$ for both 1997 and 1998 is still significantly higher than $\hat{D}$ for 1994 and 1995 with a 10% significance level.

| area\year | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 |
|---|---|---|---|---|---|---|
| unrestricted | 0.823 | 0.401 | 0.985 | 1.949 | 1.682 | 1.263 |
| 14<lon<39, 70<lat<75 | 0.858 | 0.761 | 1.330 | 1.949 | 1.733 | 1.388 |
| 20<lon<35, 71<lat<74 | 0.666 | 1.136 | 1.453 | 1.875 | 2.086 | 1.247 |

**Table 5** *Estimates of $\hat{D}$ of (2.3) based on all observations of 7-15 cm fish (first line) and on observations from restricted area only (last lines).*

Figure 8 b) shows confidence intervals for $\hat{b}_1$ from the regression of $D$ on depth, and the dependency of diurnal variation on depth also seems to vary from year to year. In fact all but 4 of the confidence intervals in Figure 8 b) contains zero. For length group 7-15 cm all estimates are very close to zero except for those of the years 1995, 1996 and 1998, which all are significantly different from zero. Further, it does not seem to be any systematic connection between dependency on depth as measured by $\hat{b}_1$ in Figure 8 b) and the strength of the diurnal variation as measured by $\hat{D}$ in Figure 8 a). Hence, it is expected that excluding the years 1995, 1996 and 1998 from the total data material in Figures 6 and 7, the depth dependencies would become much less severe. In the two leftmost plots of Figure 9 this is

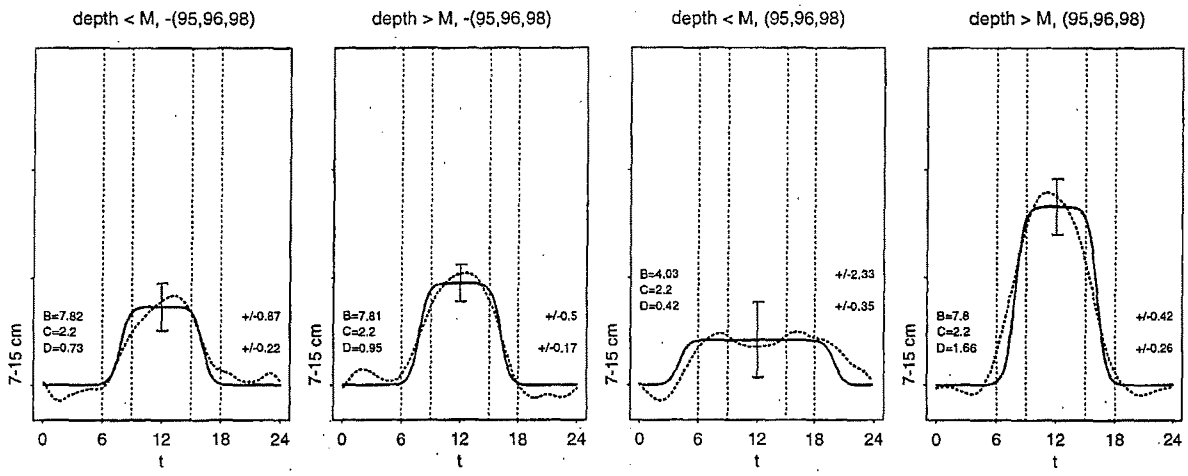


**Figure 9** *The two leftmost plots are the same as the two leftmost plots of Figure 6 for length group 7-15 cm, but with the data from 1995, 1996 and 1998 excluded. The two rightmost plots are based on data from 1995, 1996 and 1998 only.*

done, and as can be seen the confidence intervals for the two groups in fact overlap. In the two rightmost plots, where only data from 1995, 1996 and 1998 are used, there is a very distinct difference between the catches taken at deep water and those taken at shallow water. The regression line based on data from 1995, 1996 and 1998 only, corresponding to that of the upper left plot of Figure 7, is $\hat{D}_{d(i)} = -1.86 + 0.0103x$, i.e. $\hat{b}_1$ is almost twice as large as when all data are included. The 95% confidence interval for $\hat{b}_1$ is [0.0077 , 0.0129]. The corresponding polynomial curve is $\hat{D}_{d(i)} = -1.61 + 0.00822x + 0.0000039x^2$, which gives an almost straight line also for the reduced data set. When we do the regression on all years except 1995, 1996 and 1998, we get $\hat{b}_1 = 0.00189 \pm 0.00194$, i.e. not significant with a 5% level.

Autumn: As can be seen from Figure 1 the time at which the autumn surveys are conducted, and thus the light conditions (cf. Table 1), varies substantially from year to year. In the rightmost lower plot of Figure 7 we see that for 32-90 cm fish the diurnal variation, as measured by $D$, changes from positive at the darkest days to negative at the lightest days. Thus we expect a positive $\hat{D}$ for this length group on surveys undertaken late in the autumn and a negative $\hat{D}$ in surveys done in July and August, and as can be seen from Figure 10 a) this is roughly the tendency. For all years with negative point estimates of $D$ in Figure 10 (except for 1994), the median survey time in the rightmost plot of Figure 1 occurs in July or August. Correspondingly, for all years with positive point estimates of $D$ (except for 1985), the median surveytime occurs in September. The yearly numbers of catches on which the confidence intervals are based, are shown in the lower part of Table 4.

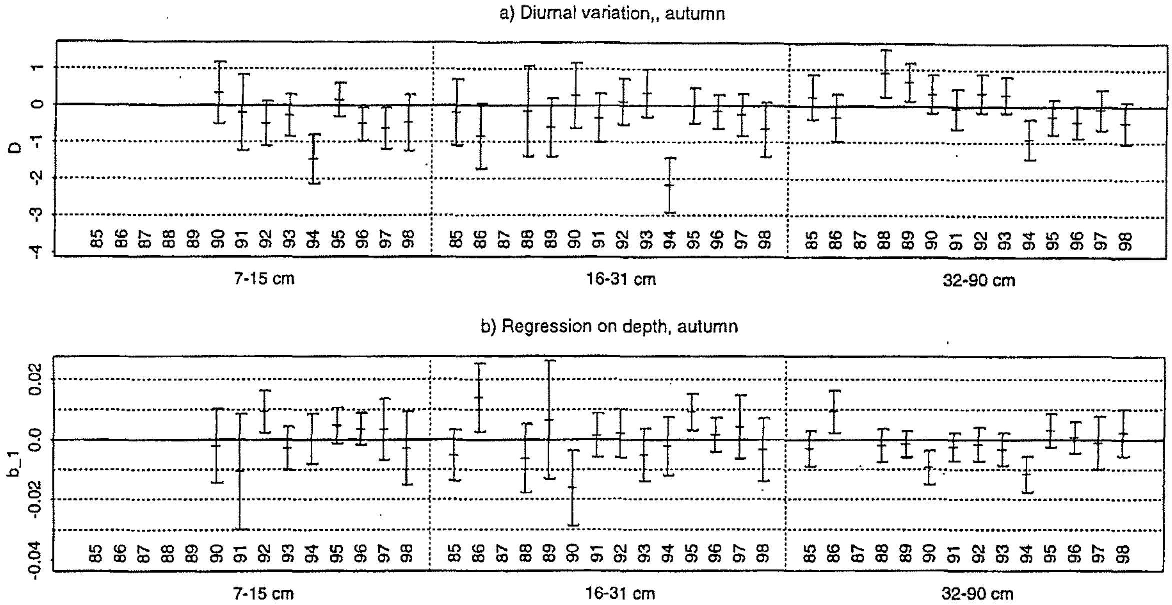


**Figure 10** *95% confidence intervals for D (top part) of model (2.2) and for $b_1$ in (2.7) (bottom) for the autumn survey. The point estimates are indicated with horizontal bars. Combinations of year and length groups with less than 40 degrees of freedom are not included.*

We also see from Figure 10 that the diurnal variation in 1994 was very strong, especially for 16-31 cm fish, with the biggest catches taken at night. This can not be explained neither by the time span nor the geographical span of the survey, which both are quite similar to the previous surveys.

When it comes to the dependency of diurnal variation on depth, the estimates of $b_1$ ranges from significantly positive to significantly negative, but in most cases $\hat{b}_1$ is not significantly different from zero.

# 5 Adjusting the catches for diurnal variation

Because there is a large year-to-year variation in diurnal effects for small fish in the winter survey (cf. Section 4), the appropriate adjustment will vary from year to year as well. For the abundance indices which are based on the catches themselves rather than their logarithms, the differences between years will be even larger. To get a rough idea of the size of the adjustment, if we use a pure threshold model with the transition taking place at $t = 8$ and at $t = 16$, two thirds of the catches are night time catches and have to to be adjusted upwards. Moreover, looking at the median, this imply that $M(Y_{\text{adj}}) \approx \frac{1}{3}M(Y) + \frac{2}{3}\{M(Y) + D\} = M(Y) + \frac{2}{3}D$. Taking antilogarithms we get $M(X_{\text{adj}}) \approx \exp\{M(Y)\}\exp(\frac{2}{3})^D \approx 2^D M(X)$.

Following the algorithm described in Section 2.3, using 100 bootstrap replicas, we get the

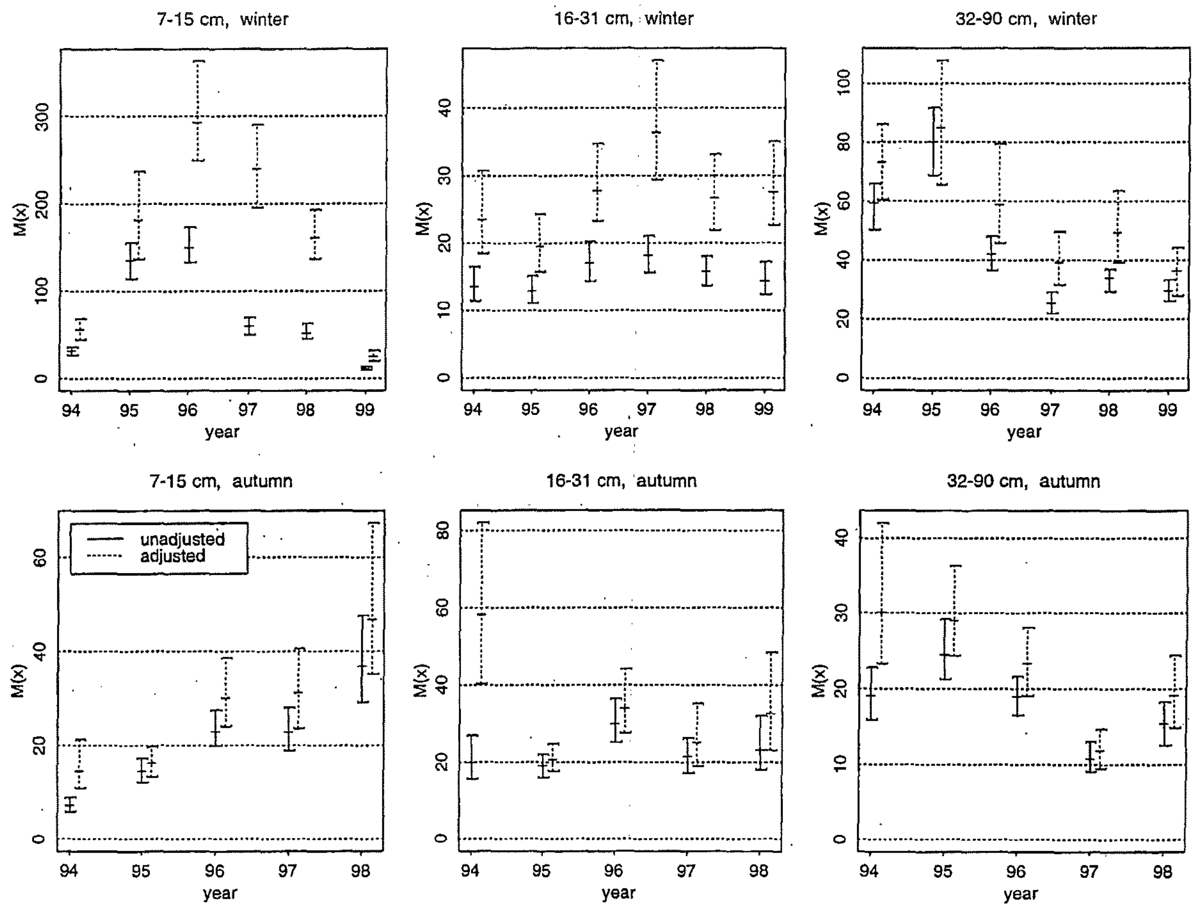


**Figure 11** *90% bootstrapped confidence intervals for the medians of the non-zero catches before (solid lines) and after adjusting for diurnal variation. The medians of the bootstrap distributions are indicated with horizontal bars.*

results shown in Figure 11 for the years 1994 to 1999. The models used are LT with $C = 2.2$ and $B = 8$ for the winter data and ST for the autumn data. Note that we get confidence intervals for the unadjusted $Y$-s from $Y_i^\star = \hat{\mu}_{d(i)} + \hat{f}(t_i) + \epsilon_i^\star$. As can be seen, the results for the winter survey agree well with the rough approximation above. For example, for small fish in 1996 and 1997 where $D \approx 1$ and $D \approx 2$, respectively, the adjusted catches are approximately 2 and 4 times the size of the unadjusted catches, respectively. We also see that for all length groups in the winter survey there are only five cases where the confidence intervals overlap, four of which occur for the 32-90 cm length group. In the most extreme case (7-15 cm 1997), the lower limit of the adjusted interval of $M(x)$ is almost three times as large as the upper limit of the unadjusted interval. For the autumn survey the effect of adjusting is much less severe, with the exception of 1994, where all the three cases of non-overlapping confidence intervals for the autumn survey occur. Regressing $D_{d(i)}$ on depth and adjusting for each day gives almost identical results, which is reasonable since the nonregressed $D$ approximately equals the average of all the regressed values $D_{d(i)}$.

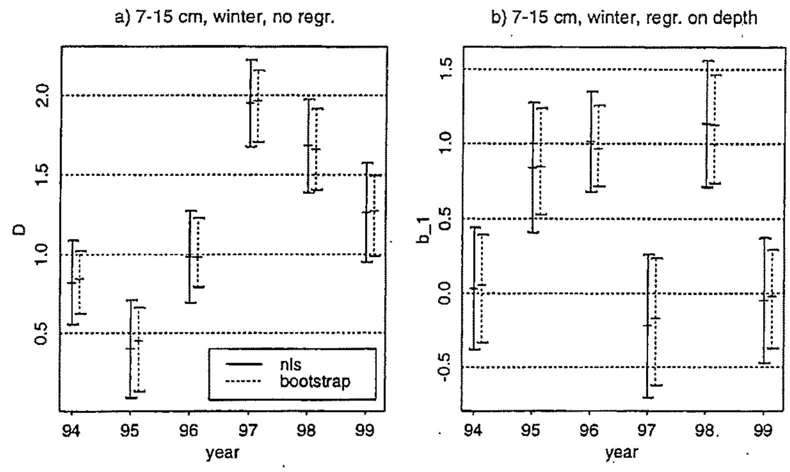


**Figure 12** *Solid lines: 90% confidence intervals for D (a) and for $b_1$ in the simple linear regression case (b) as estimated from the standard errors yielded by* `nls`. *Dotted lines: corresponding intervals from the bootstrap algorithm described in Section 2.3.*

The bootstrap procedure also gives confidence intervals for $D$, or in the regression case, for $b_1$. As can be seen from Figure 12 these tends to be slightly narrower than those yielded by `nls`, but the ratio between the two changes little.

# 6 Summary conclusions

This paper explored the capability of various statistical models to reflect diurnal variation in trawl catches. A smooth threshold model appeared to perform best on the 1985-1999 winter survey data on cod from the Barents Sea, whereas a sine model performed best on the autumn survey data. Further, a bootstrapping technique was designed to adjust trawl catches by size group. The diurnal effect on cod catches is extensive and variable from season to season and from year to year, particularly for the small sized fish, and adds considerable noise to the data if not taken into account. Further studies on other species and areas are needed to explore the general applicability of the models and their performance when applied in the estimation process of a survey time series.

Due to the difficulties in handling diurnal variability of survey trawl catches quantitatively, the problem is normally solved or minimised through application of certain rigid survey strategies. These are all built on certain assumptions, and studies of the validity of these assumptions are within the scope of this paper. For example, are day catches always higher and more representative for the population than night catches? Is an even distribution of day and night hauls by stratum from year to year enough to avoid a disturbing effect of diurnal patterns in survey time series? For the Barents Sea, where we found that the high-catch period lasts for approximately 8 hours each day in winter, surveys by day only is not a feasible strategy, due to the high cost of keeping advanced research vessels unoccupied 60-70 % of the time. Further, a major objective of this survey is to assess recruiting year classes. The analysis in this paper shows that small fish are most exposed to diurnal variability. Further, we demonstrate that it is possible to reduce the problem through modelling and compensation of the diurnal effect. This seems to be a preferable approach for several reasons: Firstly, catch rates and compositions all over the survey area become more comparable. Further, these catches are used as input in the conversion of acoustic density to fish abundance from the simultaneous acoustic survey and, hence, these estimates are similarly biased as the bottom trawl survey. Also, with this approach the survey can be designed more independent of time of day.

Does this paper contribute to the theme sessions main objectives: application of acoustics to bottom trawl survey? For semi-demersal fish like cod, with vertical distribution and migration, diurnal variation in trawl catches can be caused by changes in efficiency of the gear or variation in the amount of fish available to the trawl (or a combination of the two). Acoustic methodology can supply information to assess effects of diurnal changes in vertical distribution on catches (Aglen 1996), although also acoustic information can be affected by diurnal variability (Michalsen et al. 1996). The models explored in this paper are designed on the basis of trawl data. However, the methodology can easily be adjusted to other data with diurnal characteristics like acoustic densities. When combining data from acoustic and bottom trawl surveys to improve survey estimates of total density, the results from this paper demonstrate the importance of controlling diurnal effects. For example, if the low night catches of small fish is caused by vertical migration (see Aglen et al. 1999), seriously biased estimates will emerge, when applying night catches from bottom trawl on pelagic acoustic night recordings. If, in addition, acoustic densities have a diurnal cycle (Michalsen et al.1996), disentangling causes and quantifying effects become very difficult. The diurnal problem is most accentuated for small fish, and quantitative information on this fish is available from scientific surveys only. It is thus even more important to establish methodology to assess their abundance properly. We suggest that modelling and adjustment of diurnal influence should be an integral part of methodologies which try to combine acoustic - bottom trawl methodologies. In conclusion, applying acoustics for improving bottom trawl surveys demands basic knowledge and understanding of behavioural dynamics of the fish studied, and quantitative models to manage these behavioural characteristics. In particular, diurnal effects are apparently important for the species and area studied here and the developed models may potentially serve as an important tool for further development of methodology for combining information from acoustics and bottom trawl surveys.

# References


Aglen, A. 1996. Impact of fish distribution and species composition on the relationship between acoustic and swept-area estimates of fish density. *ICES Journal of Marine Science, 53*: 501-506.

Aglen, A., Engås, A., Huse, I., Michalsen, K., and Stensholt. B. 1999. How vertical fish distribution may affect surveys results. *ICES Journal of Marine Science, 56*: 345-360.

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.

Doubleday, W. G. 1981. Manual on Groundfish Surveys in the Northwest Atlantic. *NAFO Scientific Council Studies No. 1981*: 56p

Engås, A., and Soldal, A.V. 1992. Diurnal variations in bottom trawl catches of cod and haddock and their influence on abundance indices. *ICES Journal of Marine Science, 49*: 89-95.

Godø, O.R. 1994. Factors affecting reliability of groundfish abundance estimates from bottom trawl surveys. *In* A. Fern and S. Olsen [eds.] pp 166-199. *Marine Fish Behaviour in Capture and Abundance Estimation.* Fishing News Books. Oxford.

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 1997/Y:17*: 1-26.

Korsbrekke, K. and Nakken, O. 1999. Length and species dependent diurnal variation in catch rates in the Norwegian Barents Sea bottom trawl surveys. *ICES Journal of Marine Science, 56*: 284-291.

Michalsen, K., Godø, O.R. and Fernø, A. 1996. Diel variation in the catchability of gadoids and its influence on the reliability of abundance indices. *ICES Journal of Marine Science, 53*: 389-395.

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

Pennington, M. and Vølstad, J.H. 1991. Optimum Size of Sampling Unit for Estimating the density of Marine Populations. *Biometrics 47*: 717-723.

Venables, W.N. and Ripley, B.D. 1997. *Modern Applied Statistics with S-Plus. 2nd edition.* Springer, New York.

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

Walsh, S. 1988. Dial variability in trawl catches of juveniles and adult yellowtail flounder on the Grand Banks and the effect on assessment. *N. Amer. J. Fish. Man. 8*: 373-381.