ICES Journal of Marine Science, 58: 1272–1285. 2001
doi:10.1006/jmsc.2001.1112, available online at http://www.idealibrary.com on IDEAL®

Correcting for the effect of daylight in abundance estimation of juvenile haddock (*Melanogrammus aeglefinus*) in the North Sea: an application of kriging with external drift

Jacques Rivoirard and Kai Wieland

Rivoirard, J., and Wieland, K. 2001. Correcting for the effect of daylight in abundance estimation of juvenile haddock (*Melanogrammus aeglefinus*) in the North Sea: an application of kriging with external drift. – ICES Journal of Marine Science, 58: 1272–1285.

Kriging with external drift allows for the estimation of a spatial variable when this is driven by an external parameter, through a response function only known up to constants. This is advantageous when the effect of the parameter exists or is postulated but is not known precisely. A postulated day/night effect on catch rates in trawl survey data can be accounted for even when the day and night levels are poorly known. Similarly, the effect of time of day on catch rates can be accounted for supposing, for instance, that it varies as a cosine but with unknown coefficients. The methods are illustrated on catches of age 1 to 3 haddock in the North Sea from the first quarter International Bottom Trawl Survey (IBTS) 1983–1997, where daylight effects exist without being precisely known. A cross-validation on data values is used to measure the improvement of the methods over Ordinary Kriging. It reveals excessive variations in the parameters of individual annual variograms. Using a generic variogram appears an improvement, though not changing the global abundance. The results of kriging with external drift are compared to Ordinary Kriging, IBTS standard indices and the assessment made by the International Council for the Exploration of the Seas (ICES), in terms of global abundance and mortality coefficients. The level of agreement with the ICES assessment was similar for the abundance indices obtained by the different methods. This indicates that the IBTS standard indices are remarkably robust against sampling irregularities. Nonetheless, External Drift Kriging resulted in higher indices than the IBTS standard ones, notably for the 1-group. External Drift Kriging is capable of compensating successfully for daylight effects and provides a valuable tool for the calculation of survey-based abundance indices.

Keywords: abundance indices, cross-validation, diurnal variation, external drift, generic variogram, geostatistics, kriging, North Sea haddock.

Received 31 January 2001; accepted 7 July 2001.

Jacques Rivoirard: Centre de Géostatistique, Ecole des Mines de Paris, 35 rue Saint-Honoré, F-77305 Fontainebleau-Cedex, France; tel: (+33) 1 64 69 47 64; fax: (+33) 1 64 69 47 05; e-mail: rivoi@cg.ensmp.fr. Kai Wieland: Greenland Institute of Natural Resources, PO Box 570, DK-3900 Nuuk, Greenland; e-mail: wieland@natur.gl

Introduction

Differences in catch rates have been observed between day and night for gadoids in the Barents Sea (Aglen *et al.*, 1999; Engås and Soldal, 1992) as well as in the North Sea (Ehrich and Gröger, 1989; Wieland, 1998). Diurnal variability in bottom trawl catches affects the quality of the survey indices, especially if the diurnal effects vary from year to year, e.g. due to alterations of the ratio and changes in the location of day and night hauls. In practice, however, it may be difficult to quantify the response of the catch rates to different daylight levels from survey data as is required for a proper correction to be applied (Hjellvik *et al.*, 2001).

The International Bottom Trawl Survey (IBTS) is a coordinated, multi-vessel survey that has been conducted in the North Sea in the first quarter of the year since the mid-1960s [see Heessen *et al.* (1997) for a comprehensive description]. Trawling is preferably conducted by day but a substantial number of hauls are taken outside the daylight period in order to achieve full area coverage in the short survey time.

1054–3139/01/061272+14 $35.00/0

The main objective of the IBTS is to provide recruitment estimates and tuning data for the ICES assessments of several commercially important fish stocks. However, standard abundance indices by age group are routinely calculated in a way that does not account for spatial distribution patterns nor a possible bias due to differences in catch rates between day and night.

A comparison of quarterly IBTS indices has indicated that the catchability of gadoids, in particular age 1, is lower in the first quarter of the year and that it varies between years (ICES, 1998). This has been attributed to area effects but to some extent also to the fact that differences in catch rates between day and night are more pronounced for the 1-group than for the older ages (Wieland, 1998).

Recently, daylight effects have been included in a statistical analysis of IBTS data using Generalized Additive Models for herring (Clarke and Simmonds, 2000) and for gadoids (Jarre, Clarke and Lundgren, unpublished results). Moreover, a geostatistical technique that allows for a correction of daylight effects without knowing the exact diurnal variation of the catch rates has been proposed to map and estimate the abundance of age 2 North Sea haddock by Wieland and Rivoirard (in press). This method is kriging with external drift and was initially designed to map a geological horizon, accounting for the shape given by seismic data (Chilès and Delfiner, 1999).

The present study extends the estimation of North Sea haddock abundance to ages 1 and 3, allowing a comparison of results from age to age. Special attention is also paid to estimation variances, including a cross-validation to evaluate the performance of the methods.

Material and methods

IBTS data

Age-disaggregated catches (in numbers per hour trawling) by haul were obtained for the first quarter surveys 1983–1997 from the ICES IBTS Database. Together with the catch data information was received on a single haul basis, e.g. country, vessel, shooting position, time of day and a day/night code. The domain (area) chosen corresponds to the IBTS standard area for haddock except that the 200 m isobath was taken as the limit in the north and the east and the entire Skagerrak was excluded because of missing data in nine of the 15 years (Figure 1). This domain, which comprises an area of about 107 628 nmi^2, had about 200–300 data points available in each year.

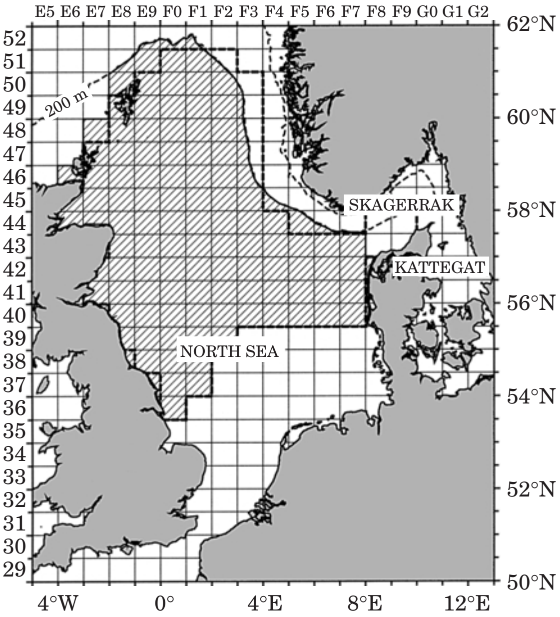

Figure 1. Map of the North Sea with ICES statistical rectangles (referenced on left and top axes), the IBTS standard area for haddock (limits indicated by the thick broken line) and the domain used in the present study (shaded area).

The number of night hauls in the IBTS varied substantially between years as well as between different parts of the survey area. In the central and northern North Sea the proportion of night hauls increased from about 14% to 33% during the past decade (Wieland and Rivoirard, in press). It should also be noted that the hauls were not equally distributed over time of day and that the night hauls were predominantly taken in the early morning and early evening while samples from the middle of the night were relatively rare. Because the night hauls were not located uniformly within the North Sea, the effect of daylight was analysed by selecting close (<20 nmi) pairs of day/night hauls from the same vessel and same year (Rivoirard, 2000). It was not possible to establish a quantitative relationship between catches of juvenile haddock and different daylight levels approximated either by time of day or sun elevation. Nonetheless, a correlation was observed, with larger catches in the middle of the day, and the ratios of mean night and mean day catches proved that the average catch rate is usually lower at night than by day for all age groups considered. While such effects were statistically significant they explain only a very small part (typically about 1%) of the variance. For instance, the part of sample variance explained by day/night is 0.024, 0.01 and 0.004 from ages 1 to 3, although the mean of night catches is about half of the mean of day catches. Hence a strong contrast exists between the amplitude of the effect on the mean catch level and the small proportion of the variability that it explains.

Geostatistical analysis

A geostatistical analysis usually involves two steps: (a) the analysis of spatial structure, e.g. through the calculation and fitting of a variogram; and (b) the use of this structure, e.g. in the application of a linear method of spatial prediction such as kriging [see e.g. Chilès and Delfiner (1999) for a comprehensive description of geostatistics, and Rivoirard *et al.* (2000) for the application to estimating fish abundance].

Variography

The variogram describes the spatial dependence of data measuring the half variability (semi-variance) between data points as a function of their distance. In the present study no indication of spatial anisotropy was found and experimental variograms were computed from the sample points as an average over all directions according to the classical formula:

$$\gamma^*(h) = 0.5\frac{1}{N(h)}\sum_{i=1}^{N(h)}\left(Z(x_i+h)-Z(x_i)\right)^2 \qquad (1)$$

where N(h) is the number of experimental pairs of data $(Z(x_i), Z(x_i+h))$ separated by a distance h. Distance intervals with lag increment of 15 nmi and a tolerance of ± 7.5 nmi were used. The experimental variograms were normalised by the sample variance in order to avoid possible numerical instability when using squares of high values in the model fitting.

Variogram models, $\gamma(h)$, were fitted with a semi-automatic, weighted-least-squares procedure by minimising:

$$\sum_{i=1}^{H} N(h_i)\left(\gamma(h_i)-\gamma^*(h_i)\right)^2 \qquad (2)$$

where H is the number of distance intervals. The variogram models were basically fitted to a maximum distance of 300 nmi with a two-component structure: a nugget effect, representing unresolved small-scale variation and observation error; and/or a spherical or a linear component.

Kriging

Kriging is a linear estimator accounting for the spatial structure and the geometrical configuration. It is optimal in the sense that it is unbiased and it minimises the variance of the estimation error. There are different types of kriging depending on the hypotheses made on the stochastic process Z(x) representing the target variable as a function of location x. We suppose here that the spatial structure is given by increments $Z(x+h) - Z(x)$ having a semi-variance which depends only on the distance vector h, equal to the variogram model $\gamma(h)$.

For Ordinary Kriging, the process is supposed to have a constant mean $E[Z(x)]=m$, or to have increments with mean 0. The estimator Z_0^{OK} of $Z_0=Z(x_0)$ from the $Z_i=Z(x_i)$ can be written as:

$$Z_0^{OK} = \sum \lambda_i Z_i \qquad (3)$$

with weights satisfying:

$$\sum \lambda_j = 100\% = 1 \qquad (4)$$

This condition ensures that the estimator is unbiased whatever the value of the mean, which may be unknown. The kriging weights, which minimise the variance of the estimation error (then called kriging variance), are found by solving a system of linear equations using the spatial information given by the variogram. Similar developments allow for the direct estimation of the average over a domain.

In kriging with external drift, the mean of the process is supposed to be linearly related to an external variable f(x) known everywhere, e.g.:

$$E[Z(x)] = m(x) = a\ f(x) + b \qquad (5)$$

Imposing the conditions on the weights:

$$\sum \lambda_i = 1 \text{ and } \sum \lambda_i\ f(x_i) = f(x_0) \qquad (6)$$

yields kriging with external drift:

$$Z_0^{ED} = \sum \lambda_i Z_i \qquad (7)$$

which is unbiased whatever the values of a and b. The kriging weights are then obtained by minimising the variance of the estimation error from:

$$\begin{cases} \sum_j \lambda_j \gamma(x_i - x_j) + \mu_0 + \mu_1 f(x_i) = \gamma(x_i - x_0) \quad \text{whatever } i \\ \sum_i \lambda_i = 1 \\ \sum_i \lambda_i f(x_i) = f(x_0) \end{cases} \qquad (8)$$

where μ_0 and μ_1 are Lagrange parameters constraining the weights to satisfy the two conditions given in Equation (6). The kriging variance is given by:

$$\sigma_K^2 = \sum_i \lambda_i \gamma(x_i - x_0) + \mu_0 + \mu_1 f(x_0) \qquad (9)$$

The estimation of the average Z(V) over a domain V, equal to the average of the estimates of its points, can be obtained directly by solving the system:

$$\begin{cases} \sum_j \lambda_j \gamma(x_i - x_j) + \mu_0 + \mu_1 f(x_i) \\ \quad = \frac{1}{|V|}\int_V \gamma(x_i - x)dx \qquad \text{whatever } i \\ \sum_i \lambda_i = 1 \\ \sum_i \lambda_i f(x_i) = \frac{1}{|V|}\int_V f(x)dx \end{cases} \qquad (10)$$

The kriging variance of this, which is not the average of the kriging variances of the points, is:

$$\sigma_K^2 = \sum_i \lambda_i \frac{1}{|V|} \int_V \gamma(x_i - x)\,dx - \frac{1}{|V|^2} \int_V \int_V \gamma(x - y)\,dx\,dy + \mu_0 + \mu_1 \frac{1}{|V|} \int_V f(x)\,dx \tag{11}$$

In the present analysis kriging with external drift is used to account for a catch level either different between day and night or varying with time of day. The mean of the process Z(x) is then allowed to vary between day or night, or with time.

In the first case (referred to as External Drift Kriging with day/night indicator) we have:

$$E[Z(x)] = m(x) = a\ 1_D(x) + b \tag{12}$$

where $f(x) = 1_D(x)$ is a day/night indicator, equal to 1 by day and to 0 at night. Thus the mean is supposed to be equal to b at night and to (a+b) by day, although neither a, nor b nor (a+b) are assumed to be known. Suppose that we want to estimate the target variable by day, then the conditions imposed on the weights imply firstly that the sum of all weights is 1 and secondly that the sum of day weights is 1. It follows that the estimator is a weighted average of day samples with weights summing to 1 and of night samples with weights summing to 0:

$$Z_0^D = \sum_{\text{day samples}} \lambda_i^D Z_i^D + \sum_{\text{night samples}} \lambda_j^N Z_j^N \tag{13}$$

with:

$$\sum_{\text{day samples}} \lambda_i^D = 1 \quad \text{and} \quad \sum_{\text{night samples}} \lambda_j^N = 0$$

The estimator is then scaled to the day-sample values, but with a correction coming from the variations between the night-sample values. The level of night values itself or any constant added to them would disappear.

In the second case, referred to as External Drift Kriging with time of day, we assume that the level of catch varies with sunlight as a cosine function of time:

$$E[Z(x)] = m(x) = a \cos\left(2\pi \frac{t - 12}{24}\right) + b \tag{14}$$

but neither a nor b, which are independent of x, are assumed known. This case must be considered as a theoretical situation since, while having a 24-hour cycle and an extreme in the middle of the day, this continuous function may be different from the actual but unknown response. The conditions on the weights [Equation (6)] become:

$$\sum \lambda_i = 1$$

and

$$\sum \lambda_i \cos\left(2\pi \frac{t_i - 12}{24}\right) = \cos\left(2\pi \frac{t_0 - 12}{24}\right) \tag{15}$$

supposing that we want to estimate the target variable at time t_0 (e.g. 12:00), and so the weights are modulated according to time of samples and time of estimation.

External Drift Kriging with day/night indicator and with time of day has been used here to estimate the global abundance respectively by day and at noon, when the catchability is expected to be maximal. For each year the estimation has been performed directly for the whole domain, from all data, according to Equation (10) and with the kriging variance given by Equation (11).

In principle the variogram to be used in kriging with external drift should represent the variability of the residual part of the variable, excluding the drift, which is usually a problem as the coefficients of the drift are considered unknown. In our case, however, the variability explained by daylight on samples is so small that this problem was neglected.

Cross-validation

Cross-validation is used to both evaluate and compare the performance of estimators, by estimating successively each data point from other data points and comparing its estimated value to the original one (e.g. Chilès and Delfiner, 1999). The square of the mean of the errors observed on all points was negligible compared to their variance. The performance of the different forms of kriging was therefore assessed using the mean square error (MSE), which includes both terms. A method was then considered as performing better than another if it resulted in a lower MSE. For a given method, on the other hand, MSE can be compared to its prediction, the mean of kriging variances at all points, to see how well the estimation method predicts uncertainty.

Comparison with IBTS standard indices and ICES assessment

IBTS standard indices of age 1 to 3 haddock were taken from ICES (1999) and from previous ICES annual reports from the first quarter IBTS. Stock numbers-at-age were taken from the most recent ICES assessment (ICES, 2000). The ICES standard assessment is, however, not strictly independent from the first quarter IBTS

Table 1. Variogram models (normalized by sample variance) for age 1 to 3 haddock, first quarter IBTS 1983-1997. GOF: Goodness of fit statistic (see Rivoirard *et al*., 2000 for definition).

Year	Model components	Nugget effect	Linear component	Spherical component		GOF (%)
				Sill	Range (nmi)	
Age 1						
1983	Nugget+spherical	0.88	—	0.21	267	0.40
1984	Nugget+spherical	0.57	—	0.51	110	1.86
1985	Nugget+spherical	0.65	—	0.47	228	0.23
1986	Nugget+spherical	0.68	—	0.42	128	0.97
1987	Nugget+spherical	0.60	—	0.58	240	0.37
1988	Nugget	1.03	—	—	—	2.94
1989	Nugget+spherical	0.53	—	0.68	154	0.59
1990	Spherical	—	—	1.07	43	1.77
1991	Nugget+spherical	0.55	—	0.65	123	1.14
1992	Nugget+spherical	0.62	—	0.48	166	0.60
1993	Nugget+linear	0.70	0.0016	—	—	0.67
1994	Nugget+spherical	0.60	—	0.51	74	3.67
1995	Nugget+spherical	0.40	—	0.47	144	0.78
1996	Nugget+spherical	0.82	—	0.24	99	1.45
1997	Nugget+linear	0.64	0.0024	—	—	0.61
Age 2						
1983	Nugget+spherical	0.27	—	0.72	48	1.58
1984	Spherical	—	—	1.04	31	1.19
1985	Nugget+spherical	0.51	—	0.63	95	0.51
1986	Nugget+spherical	0.63	—	0.45	158	0.41
1987	Nugget+spherical	0.65	—	0.45	197	0.76
1988	Nugget+spherical	0.53	—	0.58	143	1.54
1989	Nugget+spherical	0.51	—	0.56	122	0.79
1990	Spherical	—	—	1.05	41	1.19
1991	Nugget+spherical	0.38	—	0.76	105	1.18
1992	Nugget+spherical	0.72	—	0.31	95	1.83
1993	Nugget+spherical	0.47	—	0.54	88	1.28
1994	Nugget+spherical	0.42	—	0.79	266	1.07
1995	Nugget+spherical	0.61	—	0.54	250	1.00
1996	Nugget+spherical	0.58	—	0.54	228	2.00
1997	Nugget+spherical	0.39	—	0.70	48	2.84
Age 3						
1983	Nugget+spherical	0.09	—	0.81	45	0.77
1984	Nugget+linear	0.80	0.0012	—	—	1.61
1985	Nugget+spherical	0.75	—	0.43	103	1.03
1986	Nugget+spherical	0.80	—	0.39	99	1.48
1987	Nugget+spherical	0.75	—	0.26	203	2.21
1988	Nugget+spherical	0.58	—	0.46	112	0.67
1989	Nugget+spherical	0.43	—	0.72	188	0.98
1990	Nugget+spherical	0.03	—	1.01	45	0.74
1991	Nugget+spherical	0.61	—	0.56	67	1.44
1992	Nugget+spherical	0.70	—	0.32	92	0.90
1993	Nugget+spherical	0.64	—	0.33	77	2.68
1994	Nugget+spherical	0.50	—	0.51	133	1.52
1995	Nugget+spherical	0.08	—	1.30	289	3.49
1996	Spherical	—	—	1.03	30	1.85
1997	Nugget+spherical	0.74	—	0.31	153	2.26

because the IBTS data are used among data from other fleets for tuning and recruitment estimation. Hence assessments in which the IBTS was excluded from the tuning fleets, but with all other settings identical to the standard assessment, were used for comparison.

To prove the effect introduced by the different forms of kriging on the abundance indices for subsequent age groups of the same cohort, instantaneous coefficients of total mortality (Z) were calculated using the standard equation for exponential decay:

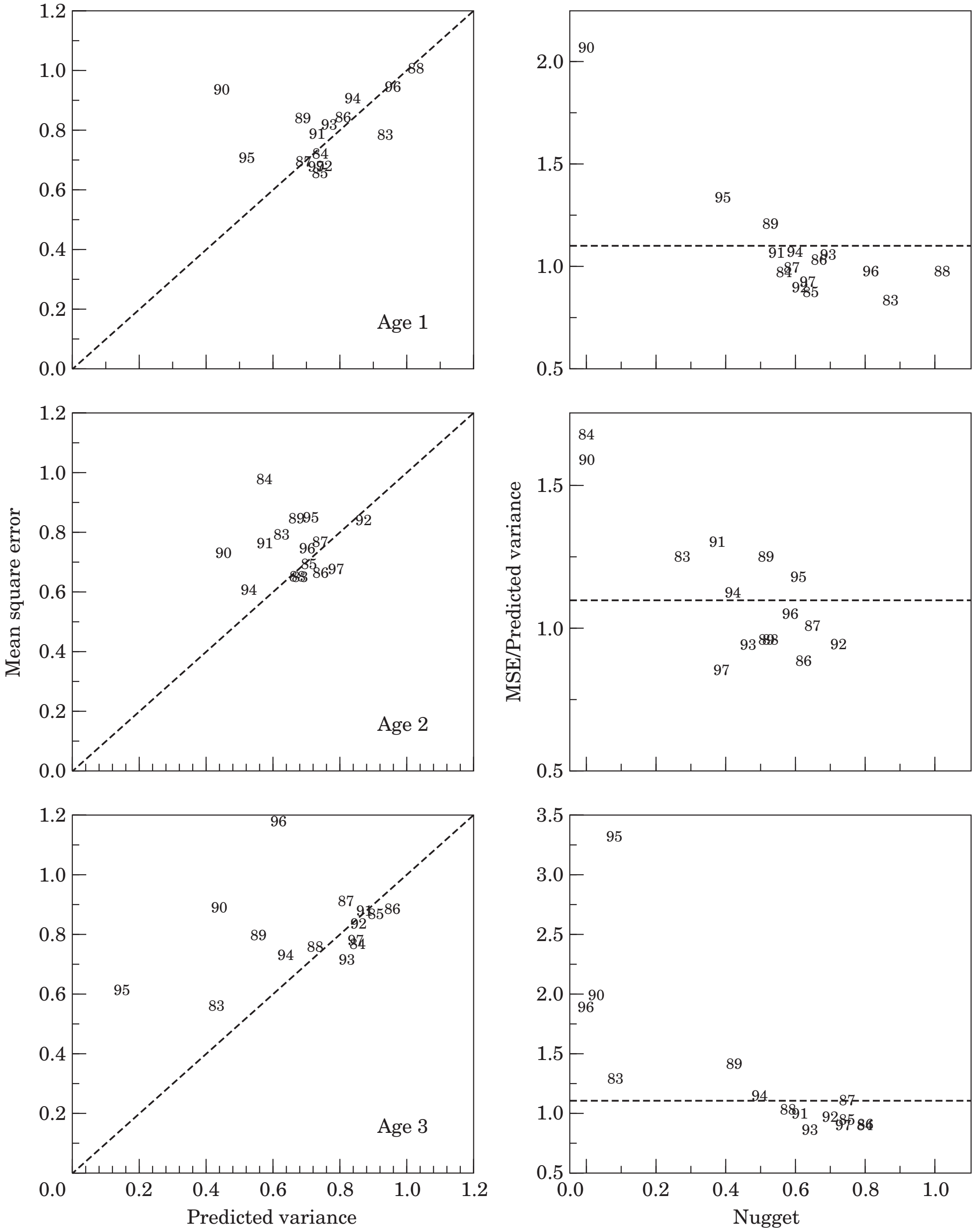

Figure 2. Cross-validation of Ordinary Kriging for age 1 to 3 haddock, using individual annual variograms. Left: mean square error (MSE) vs. predicted variance (mean kriging variance), both normalised by sample variance. Right: ratio between these two vs. nugget component of variogram. Years are indicated.

$$Z=\ln\left(N_{\mathrm{age,\,year}}/N_{\mathrm{age+1,\,year+1}}\right) \tag{16}$$

where N is either an abundance index (IBTS standard indices and geostatistical estimates) or the stock number-at-age (from an ICES assessment).

Results

Table 1 gives the parameters of fitted variogram models including both day and night catches of age 1, 2, and 3 haddock for the years 1983–1997. The number of pairs

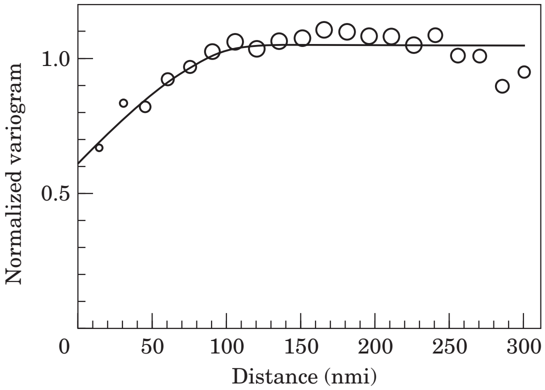

Figure 3. Generic variogram fit for haddock age 1 (nugget effect: 0.61, spherical component: 0.44, range: 120 nmi) based on a mean over all years. Symbol size is proportional to number of pairs ranging from about 3 700 to 27 700.

of samples increased rapidly with distance as is typical for omnidirectional variograms. The highest number of pairs of samples always occurred for distances between 100 and 200 nmi where most of the variograms reached their maximum level of variability. In some years satisfactory model fits were only obtained if the maximum distance included was reduced. This was accepted because a good fit near the origin was most critical for further analysis and the remaining distance range was still close to the half of the maximum distance over the domain, which can be regarded as the distance of reliability (Journel and Huijbregts, 1978). The values of the goodness-of-fit statistic (Rivoirard *et al.*, 2000; Fernandes and Rivoirard, 1999) were reasonably low, ranging between 0.4 and 3.7%. Nearly all of the resulting variogram models consisted of a nugget, between 0.03 and 0.90 and a spherical component with a range between 30 and 290 nmi.

The results from cross-validation of the different types of kriging indicated that External Drift Krigings with day/night indicator and with time of day perform better than Ordinary Kriging, particularly for younger haddock. For a given age and year the MSE, when estimating successively each data by the other data, was typically about 80% of sample variance with a difference between External Drift Krigings and Ordinary Kriging of about 2.5% of this variance for age 1, and 1% or less for ages 2 and 3. While such differences may look small they are consistent with the small part of the variability explained by the influence of daylight on samples. The differences in the predicted variances, the mean kriging variance, for the different methods were very small. The variances for the two forms of kriging with external drift were practically the same while the variance for Ordinary Kriging was typically 0.5% lower. Hence ignoring day/night or time-of-day variability leads to a lower predicted variance and so gives the illusion of better precision. The relationship between MSE and predicted variance was similar for all methods and is depicted in Figure 2 in the example of Ordinary Kriging. The ratio of MSE on predicted variance is often around 1. Given the intrinsic instability of such statistics this indicates a satisfying prediction of the variances. However the ratio can occasionally exceed a value of 1.5, indicating a far too optimistic prediction of the variance of data points. Interestingly the low variances appear to be related to zero or small nugget effects and, more generally, a decreasing relation is observed between the ratio between MSE and kriging variance and the nugget effect. This showed that the nugget and other parameters of the variogram, as fitted from the individual annual variograms, were too dispersed.

A generic variogram model, by definition the same for all years, was used for comparison in cross-validation. The mean of annual variograms normalised by the sample variance was computed for each age. Since the differences between the three age groups were small, a unique model fit to the variogram of age 1 was used for all ages. This consisted of a nugget effect of 0.61 and a spherical component of 0.44 with range 120 nmi (Figure 3). Comparisons of cross-validation results with those obtained with the individual annual variograms showed on average a MSE lower by about 1% and the prediction of the variance was improved. However, the comparisons between the different kriging methods, based on the mean relative differences in MSE or predicted variance, were not changed. The use of the generic variogram to map an annual abundance is likely to modify locally the spatial distribution of the abundance. However, it did not significantly change the estimates of global abundance of haddock age 1 to 3 and their coefficients of variation.

For all three age groups the relative difference in the global estimation of mean abundance between External Drift Kriging with the day/night indicator and Ordinary Kriging increased significantly ($p<0.05$) with the proportion of night hauls, but with some variation between the years (Figure 4). Moreover, a closer and highly significant ($p<0.001$) negative correlation of the relative difference between these two forms of kriging with the ratio of mean night and mean day catch was found (Figure 4). The deviations between the two forms of kriging of more than 15% correspond to a high proportion of night samples, i.e. >20%, together with a mean night catch below 80% of the mean day catch. The relative differences between these two forms of kriging were positive in all years except in 1986 for age 3, when the largest catch was taken at night close to a day haul and received a negative weight in External Drift Kriging. On average the effect of External Drift Kriging with day/night indicator was higher for the 1-group (12.6%)

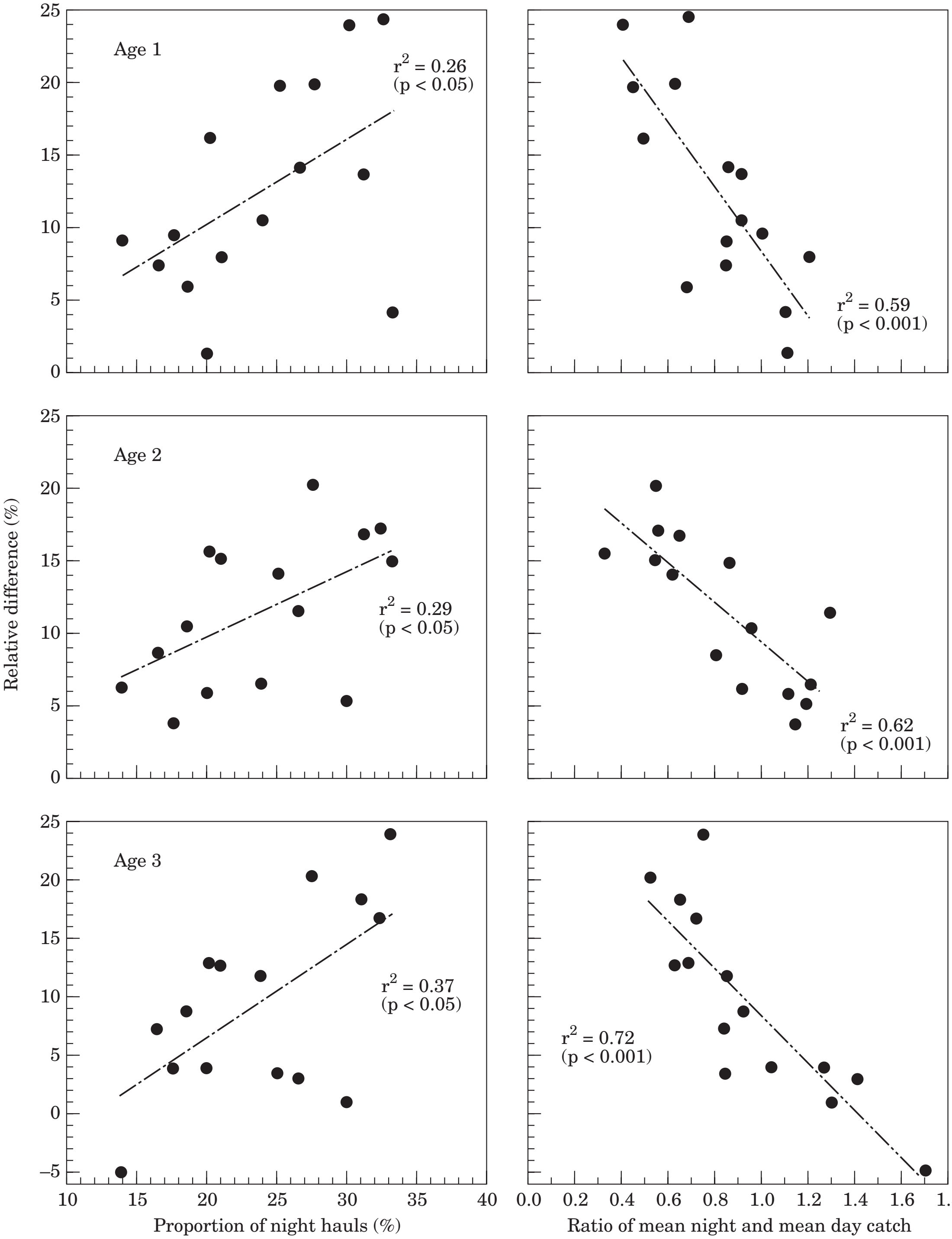

Figure 4. Relative differences between mean abundance of age 1 to 3 haddock in the first quarter IBTS 1983–1997 estimated by Ordinary Kriging and External Drift Kriging with day/night indicator. The relative difference is expressed with respect to External Drift Kriging, i.e. $(Z_0^{ED} - Z_0^{OK})/Z_0^{ED}$, and is plotted vs. the portion of night hauls and vs. the ratio of mean night and day catch rates from all stations in the domain.

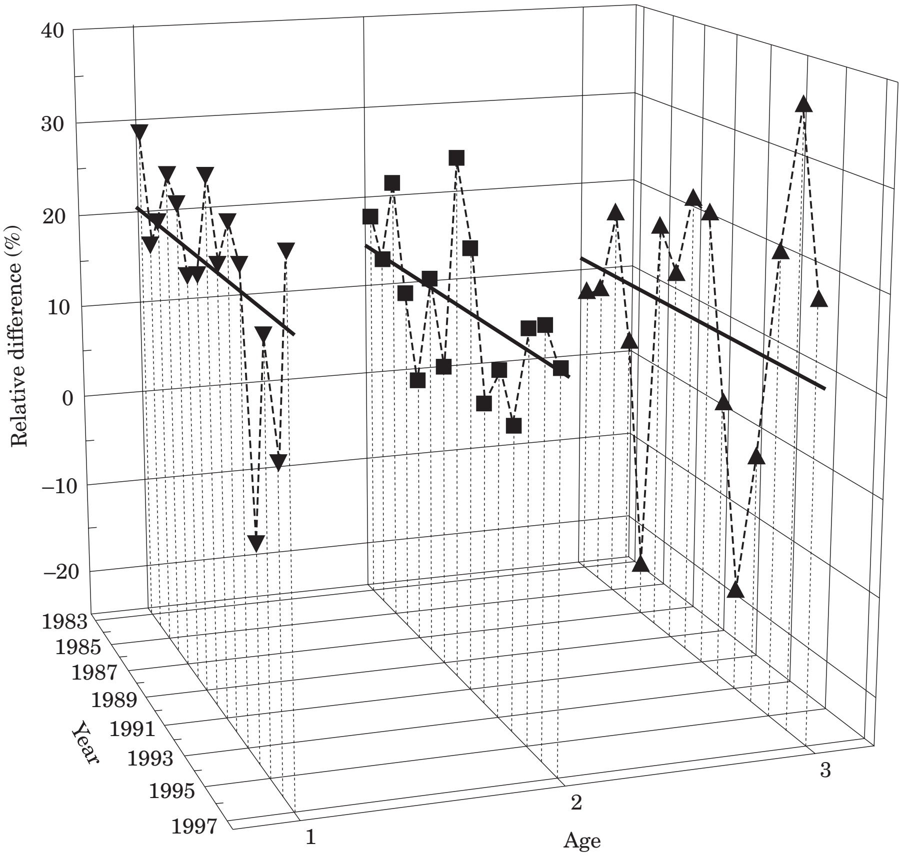

Figure 5. Relative difference between indices of mean abundance for age 1 to 3 haddock in the first quarter IBTS 1983–1997 estimated by External Drift Kriging with time of day and Ordinary Kriging. The relative difference is expressed with respect to External Drift Kriging, i.e. $(Z_0^{ED} - Z_0^{OK})/Z_0^{ED}$. Solid lines refer to mean values.

than for age 2 (11.4%) and age 3 (9.5%). External drift with time of day resulted generally in a further increase in the abundance estimates which, however, were occasionally below those obtained by Ordinary Kriging (Figure 5). This was the case for age 1 in 1994 and for age 3 in 1987 and in 1993, where high catches were taken a long time before or after noon. The average relative difference compared to Ordinary Kriging was again highest for the 1-group amounting to 20.6 % and decreased to 14.7 and 11.4 % for age 2 and 3, respectively. These variations are similar but more pronounced that those obtained from the day/night indicator.

In every case the kriging variance of the abundance was lowest for Ordinary Kriging and highest for External Drift Kriging with time of day (typically 0.3, 0.4, and 0.6% of the sample variance for the three types of kriging). This resulted in coefficients of variation on the estimates that were on average equal to 11% for Ordinary Kriging, 13% for External Drift Kriging with day/night indicator, and amounted to 16% for External Drift Kriging with time of day.

Figure 6 compares the time-series of mean abundance estimated by the three different forms of kriging with the IBTS standard index and the most recent ICES assessment for age 1, 2, and 3 haddock. For all age groups the Ordinary Kriging estimates were very close to the standard indices in most years. The estimates obtained by External Drift Kriging with day/night indicator substantially exceeded the IBTS standard indices in a couple of years, in particular for the age 1 and 2 in the 1990s. External Drift Kriging with time of day resulted in higher abundance indices than External Drift Kriging with day/night indicator for age 1 but not so much for age 2 and 3. Consequently the difference between External Drift Kriging with time of day and the IBTS standard index was even more pronounced for the 1-group.

Despite some considerable differences in their level all the survey-based abundance indices revealed similar

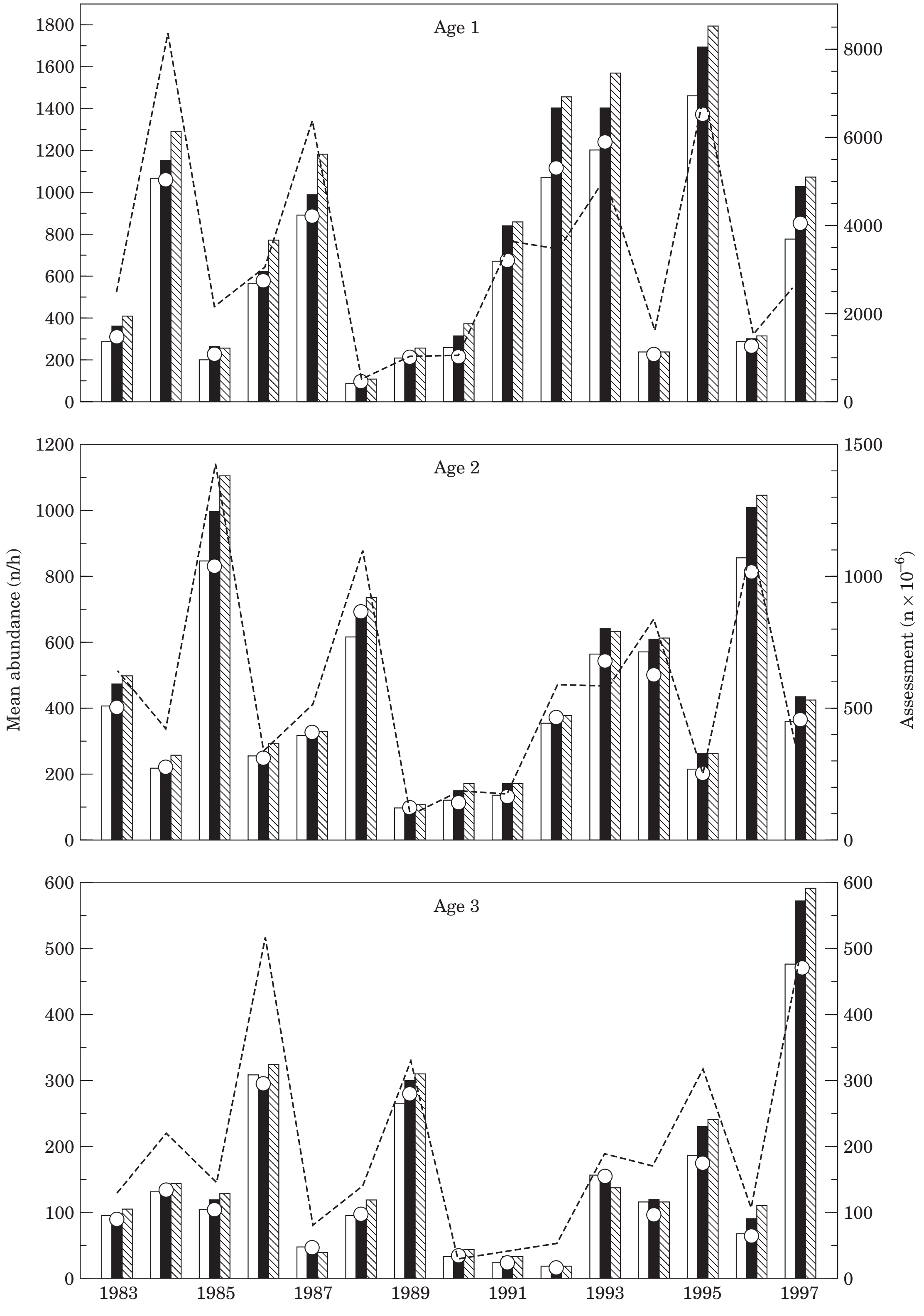

Figure 6. Times-series of abundance indices for age 1 to 3 haddock, first quarter IBTS 1983–1997, estimated by Ordinary Kriging (open bars), by External Drift Kriging with day/night indicator (solid bars) and by External Drift Kriging with time of day (shaded bars). The IBTS standard index (open circles, source: ICES, 1999) and the most recent ICES assessment (broken line, source: ICES, 2000) are given for comparison.

Table 2. Coefficients of determination (r^2) for the correlation of first quarter IBTS abundance indices with ICES assessments of age 1, 2, and 3 haddock, 1983–1997

Index	Assessment with IBTS*			Assessment without IBTS**		
	Age 1	Age 2	Age 3	Age 1	Age 2	Age 3
IBTS standard index	0.71	0.92	0.90	0.71	0.91	0.89
Ordinary Kriging	0.74	0.90	0.91	0.73	0.90	0.91
Day/night as external drift	0.65	0.89	0.85	0.64	0.88	0.84
Time of day as external drift	0.70	0.91	0.86	0.69	0.90	0.86

*: ICES, 2000; **: assessment provided by Stuart Reeves (pers. comm., Marine Laboratory Aberdeen).

trends over time, and hence correspondence with assessment results differed only marginally, irrespective of whether the ICES standard assessment or the assessment in which the IBTS had been excluded from the tuning is considered (Table 2).

Figure 7 shows instantaneous coefficients of total annual mortality for age 1 and age 2 haddock calculated from the survey-based abundance indices and from the ICES assessments for the years 1983–1996. The mortality coefficients based on External Drift Kriging with time of day deviated somewhat from the other abundance indices which were very similar to each other. In particular, for the 1-group, External Drift Kriging with time of day resulted in mortality coefficients that were higher (0.42 on average) than the other survey-based estimates (0.34–0.35 on average). All of the survey-based mortality coefficients showed much stronger fluctuations than in the ICES assessment. The unrealistic low and frequently negative mortality coefficients for the 1-group obtained from the first quarter IBTS indicate that this age group has not reached its maximum catchability at that time of the year. In contrast, the survey-based mortality coefficients for the 2-group with average values ranging between 1.16 (Ordinary Kriging) and 1.19 (External Drift Kriging with time of day) were close to the mean from the ICES assessment of 1.13.

Discussion

Several studies conducted in small, well-defined areas have shown that bottom trawl catch rates of juvenile gadoids are lower at night than by day (Aglen *et al.*, 1999; Engås and Soldal, 1992; Ehrich and Gröger, 1989; Wieland, 1998). In the case of haddock, it has further been demonstrated that this effect, although decreasing with increasing fish size (Korsbrekke and Nakken, 1999), can lead to a serious bias of survey indices up to age 3 (Engås and Soldal, 1992). Hjellvik *et al.* (2001) reported quantitative relationships between catch and time-of-day as well as between catch and sun elevation for cod in the Barents Sea. While a significant day/night or time effect exists in the IBTS data analysed in our study uncertainties in their quantification do not allow a direct rescaling of catches according to day/night or time (Rivoirard, 2000).

Kriging with external drift is a flexible means of estimating a variable when the drift is driven by another variable with unknown coefficients. In the present study the drift was modeled either from a day/night indicator, or from time of day, i.e. it was assumed that catches have a different level by day as opposed to night, or according to time. However, neither the expected level of day-catches and night-catches nor the level and amplitude of the cosine response function to time needed to be known. The assumption on the temporal variation of catch is thus weak, particularly using the day/night indicator. In this case kriging is scaled on day sample values: night catches only add a correction depending on their increments and their locations relative to day catches. Compared to the single Ordinary Kriging from day samples only this can be responsible for local differences, which can be high in the vicinity of night samples. However the estimated abundance over the whole domain is practically the same. This is due to the relatively good coverage of the domain by the day samples. By contrast, using the same method of External Drift Kriging with day/night indicator to estimate the abundance at night produces results that are different from Ordinary Kriging using night samples only. It is noteworthy that the difference between External Drift Kriging estimates of the catch by day or at noon and those obtained by Ordinary Kriging was positive with few exceptions. This means that a day-night effect, or a time effect with a night level lower than the day level, was corroborated even for years in which the mean catch at night was higher than the mean catch by day. However, the difference from Ordinary Kriging increased with the proportion of night hauls and was most pronounced in years where the ratio between the mean of the night samples and the mean of the day samples was low.

The kriging variance quantifies the variance of the error on the estimation of abundance due to the spatial coverage, as predicted by the geostatistical model. That

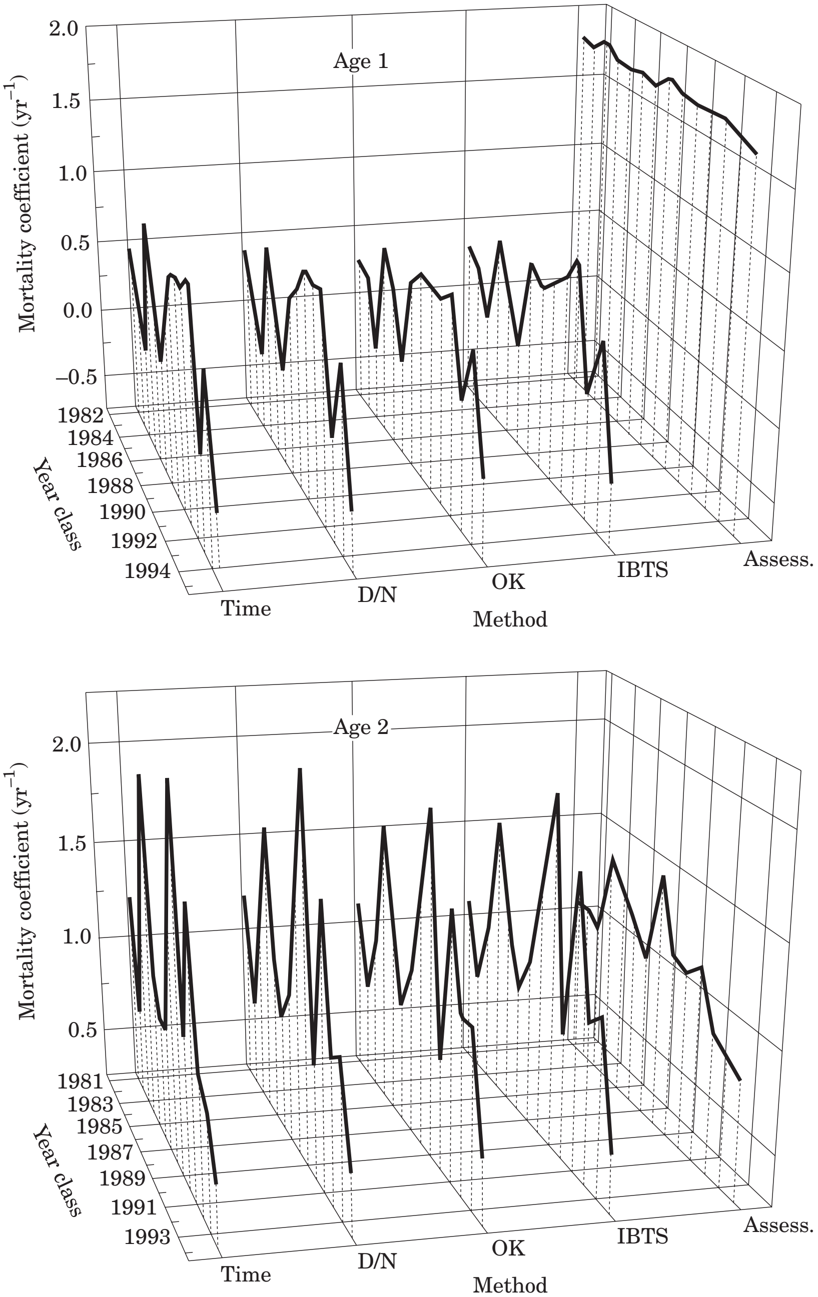

Figure 7. Coefficients of total mortality for haddock age 1 (year classes 1982–1995) and age 2 (year classes 1981–1994) based on ICES assessment (Assess.), IBTS standard indices (IBTS), and abundance indices obtained by Ordinary Kriging (OK), External Drift Kriging with day/night indicator (D/N) and External Drift Kriging with time of day (Time).

Ordinary Kriging variance was lower than External Drift Kriging variance does not mean that Ordinary Kriging is a better estimator, but rather that the estimation is less precise if a daylight effect exists. This apparent paradox makes sense. Supposing, for instance, that the variogram is pure nugget, the weights of night hauls in External Drift Kriging with day/night indicator (summing to zero and all equal) would be zero. In this case External Drift Kriging would give the average of the day hauls only, while Ordinary Kriging would be the average of all hauls. Ignoring a daylight effect if it exists yields a too optimistic precision.

In the case of the pure nugget effect, External Drift Kriging with time of day amounts to fitting by least-squares a cosine function centred at noon on the data values. Due to the lack of hauls in the middle of the night, External Drift Kriging with time of day can be unstable, even when there is a spatial structure. Of course the existence of outliers, which makes Ordinary Kriging itself unstable, may exaggerate the instability of External Drift Kriging. In a similar analysis on herring in the North Sea, Simmonds and Rivoirard (2000) reported unusually high values in the abundance index and in Ordinary Kriging for ages 2 and 3 in 1988. This instability was caused by very few extreme catch values by day. Curiously, using External Drift Kriging to estimate the abundance at noon would have exaggerated this instability while the estimation over the full 24 hours of the day stabilised the time-series.

When kriging each data point from the others using the individual annual variograms in cross-validation, predicted variances appear to be too small compared to mean square errors when the nugget was low. However the variogram describes the half square error when estimating the value at location x by a value at a varying distance h apart, and so fitting directly a model on the experimental variogram is a particular type of cross-validation, consisting in estimating each data value by one neighbouring data at varying distance, rather than by kriging from other data. Where conflict occurs, there is no reason to privilege cross-validation using kriging over direct variogram fitting when dealing with only one data set (e.g. one year). However the relationship observed over a series of 15 years shows that the variograms for the individual years, and in particular their nugget effect, exhibit too much variation from year to year. Additionally, although the variogram models the spatial structure, the nugget effect includes the variability in time that exists locally which is why the nugget could be expected not to be 0 or too low. The use of a generic variogram model, the same for all years by definition, was found to perform better in cross-validation. Practically it did not change the estimates of global abundance and their coefficients of variation.

External Drift Kriging with day/night indicator and with time of day, respectively, did not alter the level of correspondence between the abundance indices and the ICES assessment: the trend over years remained unchanged. On the other hand, the correction introduced by External Drift Kriging was most pronounced for age 1 and its effect decreased rapidly for the two subsequent age groups, in particular when time of day was used as external variable. This is well in accordance with the results of Korsbrekke and Nakken (1999) who reported a strong decrease in the diurnal variation of catch rates with size for haddock in the Barents Sea. The increase in the abundance indices for age 1, however, was not large enough to obtain much more reasonable mortality estimates than from the standard IBTS indices. This indicates that additional factors other than daylight effects are involved in the low and varying catchability of the 1-group in the first quarter of the year.

Despite the recommendation of the IBTS Working Group for daytime trawling the proportion of night hauls in the northern North Sea has increased to a high level and several rectangles have frequently been sampled exclusively at night during the 1990s. The estimates of mean abundance obtained by External Drift Kriging with a day/night indicator as well as with time of day exceeded the IBTS standard values in most of years and in particular for all years from 1990. This most likely stemmed from a systematic difference in the catch rates due to daylight effects. It should be mandatory that the proportion of night hauls be kept to a minimum and that the sampling of a given rectangle exclusively at night be avoided in future surveys as long as no correction for daylight effects is adopted. External Drift Kriging as used here re-scales in its own manner night catches to day level and is thus capable of compensating for daylight effects even if a quantitative relationship is not known. It can therefore be regarded as a valuable alternative for the calculation of survey-based abundance indices.

Acknowledgements

The European Union provided financial support (DG XIV study No. 97/0009, MIQES – The use of multivariate data for improving the quality of survey-based stock estimation in the North Sea). Stuart Reeves, Marine Laboratory Aberdeen, provided the assessment data, in which the first quarter IBTS was removed from the set of tuning fleets. The analysis was carried out with an S-Plus library for geostatistical analysis compiled by Nicolas Bez and Jacques Rivoirard (Centre de Géostatistique Fontainebleau, France). Two reviewers provided a series of valuable comments and their constructive criticism is greatly acknowledged.

References

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

Chilès, J-P., and Delfiner, P. 1999. Geostatistics: Modeling spatial uncertainty. John Wiley & Sons, New York. 695 p.

Clarke, L., and Simmonds, E. J. 2000. Spatio-temporal models of North Sea herring. ICES 2000/K: 09.

Ehrich, S., and Gröger, J. 1989. Diurnal variation in catchability of several fish species in the North Sea. ICES 1989/B: 39.

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

Fernandes, P. G., and Rivoirard, J. 1999. A geostatistical analysis of the spatial distribution and abundance of cod, haddock and whiting in North Scotland. Quantitative Geology and Geostatistics, 10: 201–212.

Heessen, H. J. L., Dalskov, J., and Cook, R. M. 1997. The International Bottom Trawl Survey in the North Sea, Skagerrak and Kattegat. ICES 1997/Y: 31.

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

ICES. 1998. Report of the Study Group on the Evaluation of the Quarterly IBTS Surveys. ICES 1998/D: 4.

ICES. 1999. Report of the International Bottom Trawl Survey in the North Sea, Skagerrak and Kattegat in 1998: Quarter 1. ICES 1999/D: 8.

ICES. 2000. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and the Skagerrak. ICES. 2000/ACFM: 7.

Journel, A. G., and Huijbregts, C. J. 1978. Mining geostatistics. Academic Press, London. 600 p.

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

Rivoirard, J. 2000. Testing the effects of vessel, gear and daylight on catch data from the International Bottom Trawl Survey in the North Sea. ICES 2000/K: 30.

Rivoirard, J., Simmonds, J., Foote, K. G., Fernandes, P., and Bez, N. 2000. Geostatistics for estimating fish abundance. Blackwell Science Ltd, Oxford, UK. 206 p.

Simmonds, J., and Rivoirard, J. 2000. Vessel and day/night effects in the estimation of herring abundance and distribution from the IBTS surveys in North Sea. ICES 2000/K: 32.

Wieland, K. 1998. Spatial distribution pattern generating processes in the International Bottom Trawl Survey in the North Sea. DFU-rapport nr. 60–98.

Wieland, K., and Rivoirard, J. In press. A geostatistical analysis of IBTS data for age 2 North Sea haddock (*Melanogrammus aeglefinus*) considering daylight effects. Sarsia.