Can the precision of bottom trawl indices be increased by using simultaneously collected acoustic data? The Barents Sea experience

Vidar Hjellvik, Dag Tjøstheim, and Olav Rune Godø

Abstract: Acoustic data are recorded continuously during the winter survey for demersal fish in the Barents Sea. This paper presents a method for using the information from the acoustic recordings between trawl stations in an attempt to increase the precision of the trawl estimate. The method is related to the double-sampling regression estimation, in which information from a frequently sampled auxiliary variable (e.g., acoustics) that is correlated with the main variable (e.g., trawl) is used for the purpose of increasing the precision in the estimate of the population mean of the main variable. The version presented here allows for additional explanatory variables and for autocorrelation in the main and auxiliary variables. However, when applied to the Barents Sea data, only a minor variance reduction is obtained. The main reasons for this are a high autocorrelation in the acoustic data and a relatively low correlation between trawl and acoustics on trawl stations. Another unexpected result is that the acoustic density during trawling is significantly higher than between trawl stations.

Résumé : Durant l'échantillonnage d'hiver des poissons démersaux de la mer de Barents, il se fait un enregistrement continu des données acoustiques. Notre travail présente une méthode pour utiliser les enregistrements acoustiques entre les stations de chalutage pour augmenter la précision des estimations obtenues par chalutage. La méthode est reliée à l'estimation par régression à double échantillonnage, dans laquelle l'information obtenue sur une variable auxiliaire fréquemment échantillonnée (par ex., les données acoustiques) qui est en corrélation avec la variable principale (par ex. les données de chalutage) sert à améliorer la précision de l'estimation de la moyenne de la population de la variable principale. La version que nous présentons permet l'insertion de variables explicatives additionnelles et l'existence d'autocorrélation dans les variables principale et auxiliaires. Cependant, son utilisation avec les données de la mer de Barents ne produit qu'une faible réduction de la variance. Les raisons principales en sont une forte autocorrélation dans les données acoustiques et une corrélation relativement faible entre les données acoustiques et celles de chalutage dans les stations de chalutage. Un autre résultat inattendu est que la densité acoustique durant le chalutage est significativement plus importante qu'entre les stations de chalutage.

[Traduit par la Rédaction]

Introduction

The standard bottom trawl surveys comprise the principal fisheries independent data source for estimating abundance of commercial demersal fish in European waters. These surveys use samples taken from very small areas (tow tracks) as representative of much larger inhomogeneous areas (strata), which can lead to a high variance in the resultant index. It is now possible, and in many cases routine, to collect acoustic data simultaneously during the surveys. Combining trawl and acoustic data would be one of the most cost-effective ways of improving these abundance estimates. As an illustration, in the Norwegian winter survey in the Barents Sea, trawl hauls with a duration of 1.5 nautical miles (n.mi) are typically taken every 20 n.mi, whereas acoustic sampling is done continuously along the survey track.

Received 10 August 2006. Accepted 24 May 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 9 October 2007.
J19479

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

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

It is an obvious idea (see, e.g., Ona et al. 1991; Cachera et al. 1999) to try to use the additional information in the acoustic data to increase the precision of the trawl estimate. This was the objective of the European Union (EU) funded project CATEFA (Combining Acoustic and Trawl surveys to Estimate Fish Abundance), in which several approaches were tried: GAMs (generalized additive models) (Beare et al. 2004), artificial neural networks (Neville et al. 2004), geostatistics (Bouleau et al. 2004), and fuzzy logic (Mackinson et al. 2005). However, little was achieved in terms of increased precision of the trawl index. The purpose of the present paper is to seek to explain and understand these results. This is done by using statistical techniques for combined sampling to demonstrate that the potential for an improvement in precision is, in fact, small. Our data are restricted to the Barents Sea winter survey, but we believe that the conclusions of the analysis have a more general validity because the results follow from features of the data that seem to be present in many combined acoustic and trawl surveys, i.e., variable and

 doi:10.1139/F07-101

often weak correlation between trawl and acoustics in addition to high autocorrelation of acoustic samples. As a consequence of the analysis, it will also emerge how strong the dependence should be between these two variables to obtain substantial improvement.

Our approach is based on standard statistical methods for combining information from two sets of data. Such methods are essential for the technique of double sampling (Cochran 1977, chapter 12). We will utilize this technique but do not impose the somewhat restrictive finite population sampling designs assumed in that book. Instead, we use a model that is more conventional in one respect (based on an infinite population) but more general in another. The increased generality is obtained by allowing for data coming from stationary time series or transect series and by permitting additional explanatory variables. We will continue to use the appellation “double sampling”, but now this will just refer to a situation in which an auxiliary variable (acoustic density) is measured together with a main variable (trawl catch) at all locations where the main variable is observed and, in addition, at many other locations. If the main and the auxiliary variables are correlated, the idea is that information from the additional samples of the auxiliary variable should increase the precision of the estimate of the main variable.

Material and methods

The Barents Sea winter survey

This is a combined bottom trawl and acoustic survey for cod and haddock that has been undertaken annually since 1981 (Jakobsen et al. 1997). As the data before 1997 have a lower resolution, only data from 1997–2002 have been used. The horizontal resolution of these data is 1 n.mi, and the vertical resolution is 10 m (surface-related layers). In addition, the acoustic density near the bottom is available as bottom-related layers with 1 m vertical resolution. Because the horizontal resolution of the acoustic data is 1 n.mi and a trawl station typically covers 1.5 n.mi, there is not an exact match between the area covered by the two sample types. A rough estimate of the acoustic density corresponding to each trawl station has been calculated using a weighted average of the acoustic samples that overlap with the trawl station. This is described in detail in Hjellvik et al. (2003) together with a method for converting the surface-related layers to bottom-related layers. In Hjellvik et al. (2003), the correlation between trawl catches and acoustic densities was found to be highest when the acoustic density was accumulated from the bottom up to about 50 m above the bottom, and this part of the acoustic density has been used here.

The trawl catches have been transformed into ENASC (equivalent NASC (nautical area scattering coefficient, $m^2 \cdot (n.mi)^{-2}$)) values as

$$\mathrm{ENASC} = \sum_L [(C_L \sigma_L)/(d w_L)]$$

where C_L, σ_L, and w_L denote catch in number, scattering cross section, and effective fishing width, respectively, for length group L, and d is the towed distance (for details, see Hjellvik et al. 2003). An overview of the survey methodology is given in Jakobsen et al. (1997).

The correlation between trawl and acoustics is essential in combining them. As the data are very skewed (approximately lognormally distributed), outliers can heavily influence the correlation estimates, and all of the analyses are therefore carried out on log-transformed data.

Typically, trawl stations are taken every 20th n.mi. However, the representativeness of a catch can be destroyed by clay in the codend, torn net, or other uncontrolled factors. Such catches are not used in the analysis. Moreover, for logistic reasons, the number of acoustic samples between trawl stations may be different from 20. The survey transect and the distribution of the number of between-station samples for the 2002 survey is shown (Fig. 1).

Double sampling with independent data

Let T_i and A_i denote the trawl catch (ENASC) and acoustic density (NASC), respectively, at trawl station i (notation is given in List of symbols). We want to estimate expected trawl catch $\mu(T)$ per square nautical mile in a certain region of size Ω, say. The average trawl catch $\overline{T}$ is an obvious estimate of $\mu(T)$.

The idea of double sampling is that the more frequent between-station observations A_j can be used to increase the precision of $\overline{T}$ as an estimate of $\mu(T)$ supposing that A_i and T_i are correlated. First, we explain the effect of double sampling in the simple situation in which there is a standard linear regression relationship

$$(1) \qquad T_i = \alpha + \beta A_i + u_i, \quad i = 1, \ldots, n_{\mathrm{ON}}$$

between A_i and T_i. Here $\{u_i\}$ are independent identically distributed error variables with expected value $E(u_i) = 0$ and n_{ON} is the number of trawl stations. We assume A_i, $i = 1, \ldots, n_{\mathrm{ON}}$, to be independent and independent of $\{u_i\}$. Observations for T_i are available for $i = 1, \ldots, n_{\mathrm{ON}}$, whereas for A_j we have observations $j = 1, \ldots, n_{\mathrm{ALL}}$, where n_{ALL} is the number of acoustic samples and is considerably larger than n_{ON}.

To explain the essentials of the method in a simple setting, first, we assume that A_j, $j = 1, \ldots, n_{\mathrm{ALL}}$ are independent. This assumption is clearly not realistic as the acoustic samples are close in time and in space. In addition, both A_j and T_i will be influenced by spatial trends. More realistic assumptions will be introduced in the next section. We are interested in the expected value $\mu(T)$ of T using a combined index estimate $\hat{\mu}(T) = I_C$ composed of both the n_{ON} joint observations of (T_i, A_i) at trawl stations and the n_{ALL} minus n_{ON} acoustic observations taken between stations. From eq. 1, we have $\mu(T) = \alpha + \beta\mu(A)$ and $\overline{T}_{\mathrm{ON}} = \alpha + \beta\overline{A}_{\mathrm{ON}} + \overline{u}_{\mathrm{ON}}$, where $\overline{T}_{\mathrm{ON}}$, $\overline{A}_{\mathrm{ON}}$, and $\overline{u}_{\mathrm{ON}}$ are averages of on-station observations. Subtracting we have

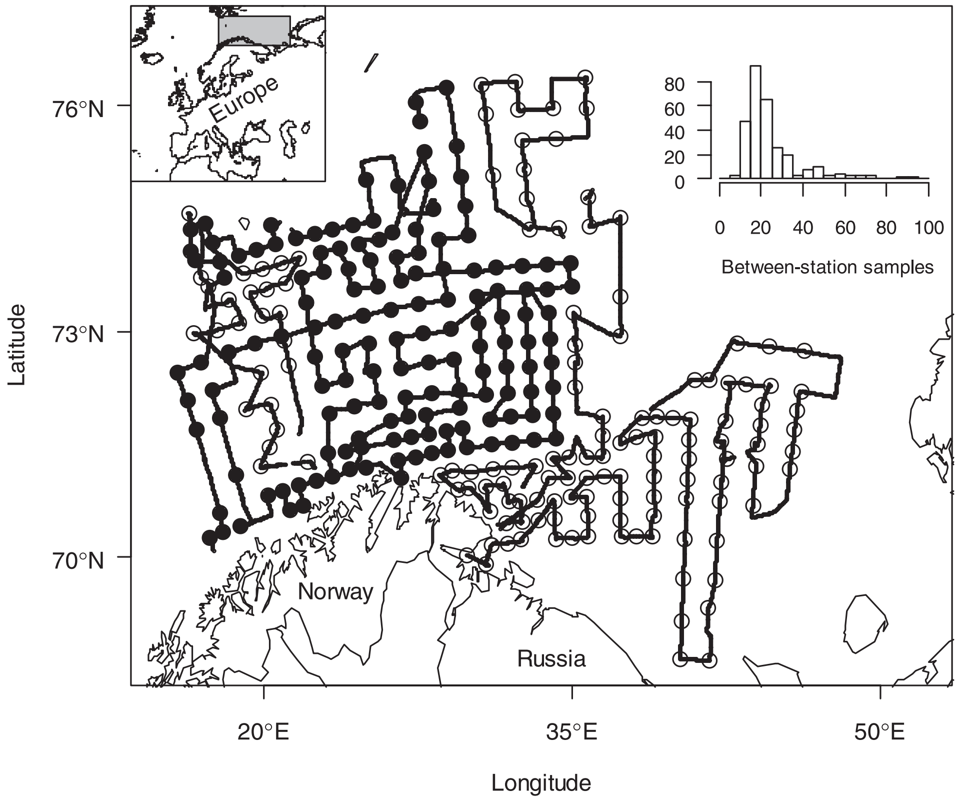
$$\mu(T) - \overline{T}_{\mathrm{ON}} = \beta[\mu(A) - \overline{A}_{\mathrm{ON}}] - \overline{u}_{\mathrm{ON}}$$

Because $E(u_i) = 0$, the best estimate of the mean $\overline{u}_{\mathrm{ON}}$ of the regression residuals is zero, and it follows that $\mu(T)$ can be estimated by

$$\hat{\mu}(T) = I_c = \overline{T}_{\mathrm{ON}} + \hat{\beta}[\hat{\mu}(A) - \overline{A}_{\mathrm{ON}}]$$

The expected value $\mu(A)$ is estimated by using all of the acoustic observations, and hence

Fig. 1. Acoustic samples (small points, hardly distinguishable as they are so dense) and trawl stations taken by *Johan Hjort* (open circles) and *G.O. Sars* (solid circles) for the 2002 survey. Upper right corner: histogram of the number of acoustic samples between trawl stations for both vessels.



(2) $$I_c = \hat{\mu}(T) = \bar{T}_{\mathrm{ON}} + b_{\mathrm{ON}}(\bar{A}_{\mathrm{ALL}} - \bar{A}_{\mathrm{ON}})$$

where b_{ON} is the standard estimator of the regression coefficient β in eq. 1 based on the on-station joint observations.

Using standard properties of the regression model of eq. 1, the variance of I_C (see Appendix A for details) is

(3) $$\mathrm{var}(I_c) = (1 - \rho_{A,T}^2)\frac{\mathrm{var}(T_i)}{n_{\mathrm{ON}}} + \rho_{A,T}^2\frac{\mathrm{var}(T_i)}{n_{\mathrm{ALL}}} \le \frac{\mathrm{var}(T_i)}{n_{\mathrm{ON}}} = \mathrm{var}(\bar{T})$$

where $\rho_{A,T}$ is the correlation between trawl and acoustics. Note that eq. 3 is really an approximation (see Appendix A) that is only valid if n_{ON} is large (fulfilled in our case) and if $\{T_i\}$ and $\{A_j\}$ are series of "independent" samples from the survey area Ω. It will be generalized in the next subsection to take autocorrelation and possible explanatory variables into account. If T_i is measured in, e.g., kilograms per square nautical mile, ΩI_C is now an estimate of the total biomass in the survey area, and with the assumptions used in this subsection, it follows from eq. 3 that it will have lower variance than the trawl-only estimate if $\rho_{A,T} \neq 0$.

The correlation between trawl and acoustics is usually positive ($\beta > 0$), and hence usually $b_{\mathrm{ON}} > 0$, and the combined index I_C is adjusted upwards or downwards compared with the trawl-only index $\bar{T}$ according to whether or not $\bar{A}_{\mathrm{ALL}} > \bar{A}_{\mathrm{ON}}$.

Generalized double sampling

In a generalized version, trawl catches and acoustics are first modelled as regression functions of explanatory variables using the GAM (generalized additive model) approach, and the variance of the combined index (corresponding to eq. 3) is then calculated from the residuals of the regressions. This is done to take into account the explanatory power of covariables and to remove trends due to spatial variation, for instance. The residuals obtained from the chronologically ordered observations along the transect will be treated as time series. They are approximately stationary along the transect as opposed to the trawl and acoustic series themselves, which will contain trends. As long as the parallel sections of the transect are sufficiently far apart in time or in space, this will be true irrespective of the particular spatial sampling design chosen for the survey. The analogue for a trawl-only index is to fit a regression model and calculate the error of the predicted mean trawl catch from the regression error variance. The models fitted to the trawl and acoustic observations are

(4) $$T_i = f_T(x_{i1}, x_{i2}, \ldots, x_{ik}) + z_i, \quad i = 1, \ldots, n_{\mathrm{ON}}$$

(5) $$A_j = f_A(y_{j1}, y_{j2}, \ldots, y_{jl}) + \varepsilon_j, \quad j = 1, \ldots, n_{\mathrm{ALL}}$$

where f_T and f_A are functions of the explanatory variables $x_1, \ldots, x_k$ and $y_1, \ldots, y_l$ and where the residuals z_i, $i = 1, \ldots, n_{\mathrm{ON}}$, and ε_j, $j = 1, \ldots, n_{\mathrm{ALL}}$, are assumed to be stationary and autocorrelated. This autocorrelation signifies the continued presence of spatial dependence after trend removal along the transect. It is weak (from 0 to 0.4) for the trawl data, because the trawl stations are far apart and strong (0.8–0.95) for the acoustic data. The explanatory variables do not have to be the same for trawl and acoustics, although it seems reasonable that they are. Note that in fact we have two acoustic data sets, the entire set $\{A_j, j = 1,\ldots,n_{\mathrm{ALL}}\}$ covering both trawl stations and the tracks between them and the on-station set $\{A_i, i = 1,\ldots,n_{\mathrm{ON}}\}$. In the entire set, each observation is the NASC averaged over 1 n.mi, whereas in the on-station set, each observation is a weighted average of the two or three acoustic samples overlapping with the corresponding trawl station. For example, if trawl station i has a length of 1.5 n.mi and is covered by A_j (20%), A_{j+1} (100%), and A_{j+2} (30%), then $A_i = (0.2A_j + A_{j+1} + 0.3A_{j+2})/1.5$ (cf. Hjellvik et al. 2003). Acoustic residuals for the trawl stations are defined as

$$(6)\qquad e_i = A_i - f_A(y_{i1}, y_{i2}, \ldots, y_{il}), \quad i = 1, \ldots, n_{\mathrm{ON}}$$

where f_A is the function in eq. 5. In practice we use estimated residuals $\hat{z}_i$, $\hat{e}_i$, and $\hat{\varepsilon}_j$ corresponding to estimates of f_T and f_A. Finally, as for eq. 2, the combined index is calculated as

$$(7)\qquad I_c = \overline{T}_{\mathrm{ON}} + b_{\mathrm{ON}}(\overline{\varepsilon}_{\mathrm{ALL}} - \overline{e}_{\mathrm{ON}})$$

where $b_{\mathrm{ON}} = \hat{\beta}$ is the estimated regression coefficient in the regression $z_i = \alpha + \beta e_i + u_i$, $i = 1,\ldots,n_{\mathrm{ON}}$, i.e., (z_i, e_i) are playing the role of (T_i, A_i) in eq. 1. Here, because these are on-station quantities, z_i, $i = 1,\ldots,n_{\mathrm{ON}}$, e_i, $i = 1,\ldots,n_{\mathrm{ON}}$, and u_i, $i = 1,\ldots,n_{\mathrm{ON}}$, are weakly autocorrelated variables, and $\{e_i\}$ is assumed to be independent of $\{u_i\}$.

The index I_C of eq. 7 would in general be different from, but analogous to, the index I_C of eq. 2. If $\overline{\varepsilon}_{\mathrm{ALL}} > \overline{e}_{\mathrm{ON}}$, the acoustic residuals between stations are on average larger than those on station, and if $b_{\mathrm{ON}} > 0$, as is usually the case, the combined index I_C is adjusted upwards as compared with the trawl-only index $\overline{T}$. Ignoring the uncertainty in fitting eqs. 4 and 5, i.e., in estimating f_T and f_A, and considering the explanatory variables (e.g., longitude and latitude) as nonrandom, it follows that $\mathrm{var}(T_i) = \mathrm{var}(z_i)$ and the variance of the trawl-only index is $\mathrm{var}(\overline{z})$ (the variance of the mean error of eq. 4), whereas the variance of the combined index is given by $\mathrm{var}(I_C) = \mathrm{var}[\overline{z}_{\mathrm{ON}} + b_{\mathrm{ON}}(\overline{\varepsilon}_{\mathrm{ALL}} - \overline{e}_{\mathrm{ON}})]$. Using the derivation in Appendix A, we have

$$(8)\qquad \mathrm{var}(I_c) = (1 - \rho_{e,z}^2)\,\mathrm{var}(\overline{z}_{\mathrm{ON}}) + \beta^2\,\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})$$

where again this is an approximation based on the simplifying assumptions made above and in Appendix A. Note that in the case of independent z_is and ε_is, the index I_C of eq. 7 is completely analogous to I_C of eq. 2. Moreover, eqs. 8 and 3 are identical in this situation, which is easy to check using standard regression results. The change in variance obtained by the combined index eq. 7 compared with the trawl-only index $\overline{T} = \overline{T}_{\mathrm{ON}}$ is

$$(9)\qquad \Delta_V = 1 - \frac{\mathrm{var}(I_c)}{\mathrm{var}(\overline{T})} = 1 - \frac{(1-\rho_{e,z}^2)\,\mathrm{var}(\overline{z}_{\mathrm{ON}}) + \beta^2\,\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})}{\mathrm{var}(\overline{z}_{\mathrm{ON}})}$$

$$= \rho_{e,z}^2 - \beta^2 \frac{\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})}{\mathrm{var}(\overline{z}_{\mathrm{ON}})} = \rho_{e,z}^2\left(1 - \frac{\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})\,\mathrm{var}(z_i)}{\mathrm{var}(\overline{z}_{\mathrm{ON}})\,\mathrm{var}(e_i)}\right)$$

where we used $\beta^2 = \rho_{e,z}^2\,\mathrm{var}(z_i)/\mathrm{var}(e_i)$. Note that Δ_V will be small if $\rho_{e,z}^2$ is small.

The fact that ε_j, $j = 1,\ldots,n_{\mathrm{ALL}}$, are autocorrelated must be taken into account when calculating $\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})$ in eq. 8. This can be done analytically or by autoregressive bootstrap if ε_j follows an autoregressive process, or by block bootstrap for more complicated ε_j processes like those of the survey data (see Appendix A for details). The calculation of $\mathrm{var}(\overline{z}_{\mathrm{ON}})$ can be done similarly, but here the first-order autoregressive analytic approximation is usually accurate because the autocorrelation generally is much weaker.

Checking the variance formula by simulation

The approach described in the preceding subsection consists of two stages: the fitting of a GAM regression to the data as in eqs. 4 and 5 and the application of the generalized double-sampling technique to the residuals from this fit, resulting in eqs. 7 and 8. As has been mentioned, there are some approximations involved in eq. 8. First, there are the approximations in the derivation in Appendix A. Second, the uncertainty has been ignored in the estimates $\hat{f}_A$ and $\hat{f}_T$ of f_A and f_T in eqs. 4 and 5 when computing $\mathrm{var}(I_C)$ of eq. 8. We have therefore investigated the accuracy of the variance eqs. 8 and 9 by three sets of simulation experiments, two of them to some extent mimicking the 2002 survey. In all of the simulations, eq. A4 is used to calculate $\mathrm{var}(\overline{z}_{\mathrm{ON}})$, and the block bootstrap with block length 100 and 100 bootstrap replicates is used for $\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})$.

We simulate residual processes $\{z_j\}$ and $\{\varepsilon_j\}$ for trawl and acoustics, respectively, as first-order autoregressive processes (cf. Appendix A):

$$(10)\qquad \varepsilon_j = a\varepsilon_{j-1} + u_j + w_j,\; u_j \sim N(0,\sigma_u^2), w_j \sim N(0,\sigma_w^2), \quad j = 1,\ldots,n_{\mathrm{ALL}}$$

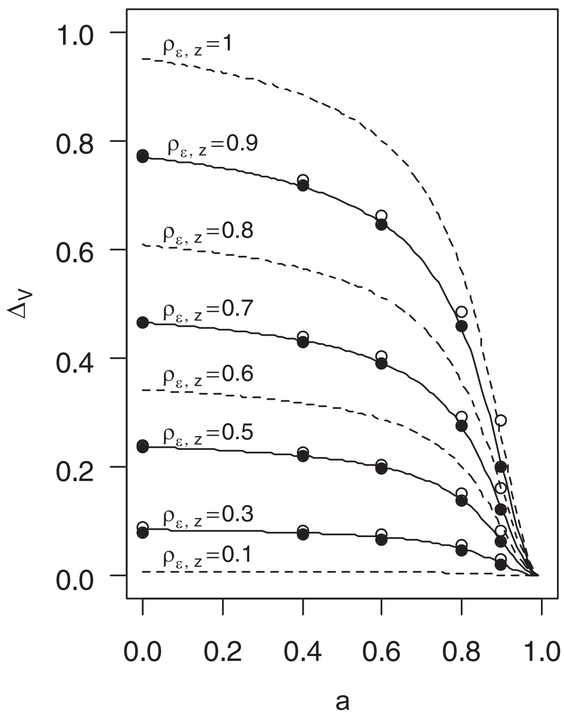
$$z_j = az_{j-1} + v_j + w_j,\; v_j \sim N(0,\sigma_v^2), \qquad j = 1,\ldots,n_{\mathrm{ALL}}$$

where the sequences of independent normally distributed random variables $\{u_j\}$, $\{v_j\}$, and $\{w_j\}$ are independent. The correlation between acoustics and trawl is generated by the common variable $\{w_j\}$, where a higher value of σ_w^2 leads to higher correlation (cf. eq. A8 in Appendix A). For convenience, we first generate one trawl residual z_j for each acoustic residual ε_j, and then discard all of the z_js except the ones that correspond to on-station data. That is, we use the z_js with $j \in I_{\mathrm{ON}}$, where I_{ON} denotes the set of samples that are on-station. Below we refer to this subset as $\{z_i\}$.

In the first experiment, we use $n_{\mathrm{ALL}} = 5000$ and one trawl residual for each 20th acoustic residual, i.e., $I_{\mathrm{ON}} = \{10, 30, \ldots, 4990\}$. This makes the simulated variance reductions directly comparable to theoretical variance reductions in eq. 9. Five values of a (0, 0.4, 0.6, 0.8, and 0.9) and four values of $\rho_{e,z}$ (0.3, 0.5, 0.7, 0.9) are used. Equations 7–9 are applied directly on the $\{z_i\}$ and $\{\varepsilon_j\}$ processes with no GAM involved, i.e., $\overline{T}_{\mathrm{ON}}$ and $\overline{e}_{\mathrm{ON}}$ in eq. 7 are replaced with $\overline{z}_{\mathrm{ON}}$ and $\overline{\varepsilon}_{\mathrm{ON}}$, respectively.

In the second experiment, the main change is that instead of using equally spaced trawl stations, we use the same spacing as in the 2002 survey. We use $n_{\mathrm{ALL}} = 7383$ and $n_{\mathrm{ON}} = 287$, as in the 2002 survey, and I_{ON} now denotes the set of samples for which A_i in the real data overlap with a trawl station (only one i for each station). The correlations and variances in the simulation are as for the demersal trawl and acoustic data in 2002, i.e., $a = \rho_\varepsilon(1) = 0.8673$, $\rho_{e,z} = 0.3841$,

Fig. 2. Curves: theoretical variance reduction (Δ_V) calculated using eqs. 9 and A4, when the acoustic samples $\{\varepsilon_j, j = 1,\ldots,n_{\mathrm{ALL}}\}$ and the trawl samples $\{z_i, i = 1,\ldots,n_{\mathrm{ON}}\}$, $n_{\mathrm{ON}} = n_{\mathrm{ALL}}/20$, are first-order autoregressive processes as in eq. 10 with autocorrelations a and a^{20}, respectively, and the correlation $\rho_{\varepsilon,z}$ is as indicated in the figure. Solid and open circles are simulated variance reductions as described in the main text.



$\sigma_z = 0.8116$, and $\sigma_e = 0.9664$, obtained by using $\sigma_u = 0.3917$, $\sigma_w = 0.2517$, and $\sigma_v = 0.2992$ (cf. eq. A8 in Appendix A). This gives an autocorrelation of about 0.12 for on-station z_is.

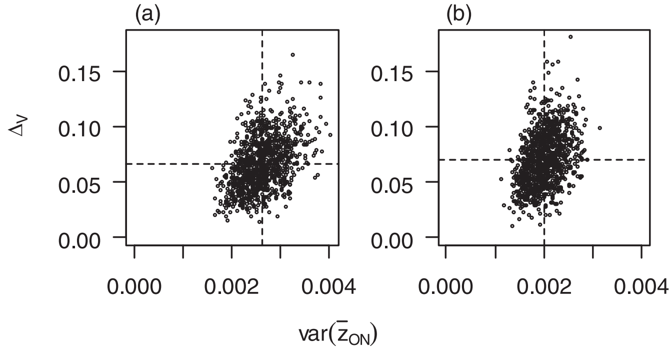
In the third experiment, in addition to the unequally spaced trawl stations, the GAM aspect is included, both in the data generation and in the index calculation. Acoustic samples A_j and trawl catches T_i are generated as

$$(11)\qquad A_j = f_A(\mathrm{lat}_j, \mathrm{lon}_j) + \varepsilon_j, \quad j = 1,\ldots,n_{\mathrm{ALL}}$$

$$T_i = f_T(\mathrm{lat}_i, \mathrm{lon}_i) + z_i, \quad i = 1,\ldots,n_{\mathrm{ON}}$$

where the complete set $\{\varepsilon_j\}$ from eq. 10 is used to generate acoustic samples on and between stations, and $\{z_i\}$ (the on-station zs) from eq. 10 are used to generate trawl catches on stations. The functions f_A and f_T are taken to be the GAMs fitted to the 2002 data of demersal acoustics and catch, respectively, lat_i and lon_i are the latitudes and longitudes of the trawl stations that year, and lat_j and lon_j are the latitudes and longitudes of the acoustic samples. Next, a GAM was fitted to the simulated data, obtaining the estimates $\hat{f}_A$ and $\hat{f}_r$. The combined index and its variance were then calculated according to eqs. 7 and 8, replacing z_i by $\hat{z}_i = T_i - \hat{f}_T(\mathrm{lat}_i, \mathrm{lon}_i)$, $i = 1,\ldots,n_{\mathrm{ON}}$, and ε_j by $\hat{\varepsilon}_j = A_j - \hat{f}_A(\mathrm{lat}_j, \mathrm{lon}_j)$, $j = 1,\ldots,n_{\mathrm{ALL}}$, where again the latitudes and longitudes are taken from the 2002 survey.

Fig. 3. Variance reduction (Δ_V) as a function of var($\bar{z}_{\mathrm{ON}}$) for (*a*) simulation experiment 2 without any GAM involved and (*b*) experiment 3 with GAM involved. Broken vertical and horizontal lines show the average var($\bar{z}_{\mathrm{ON}}$) and variance reduction, respectively.



Results

Simulations

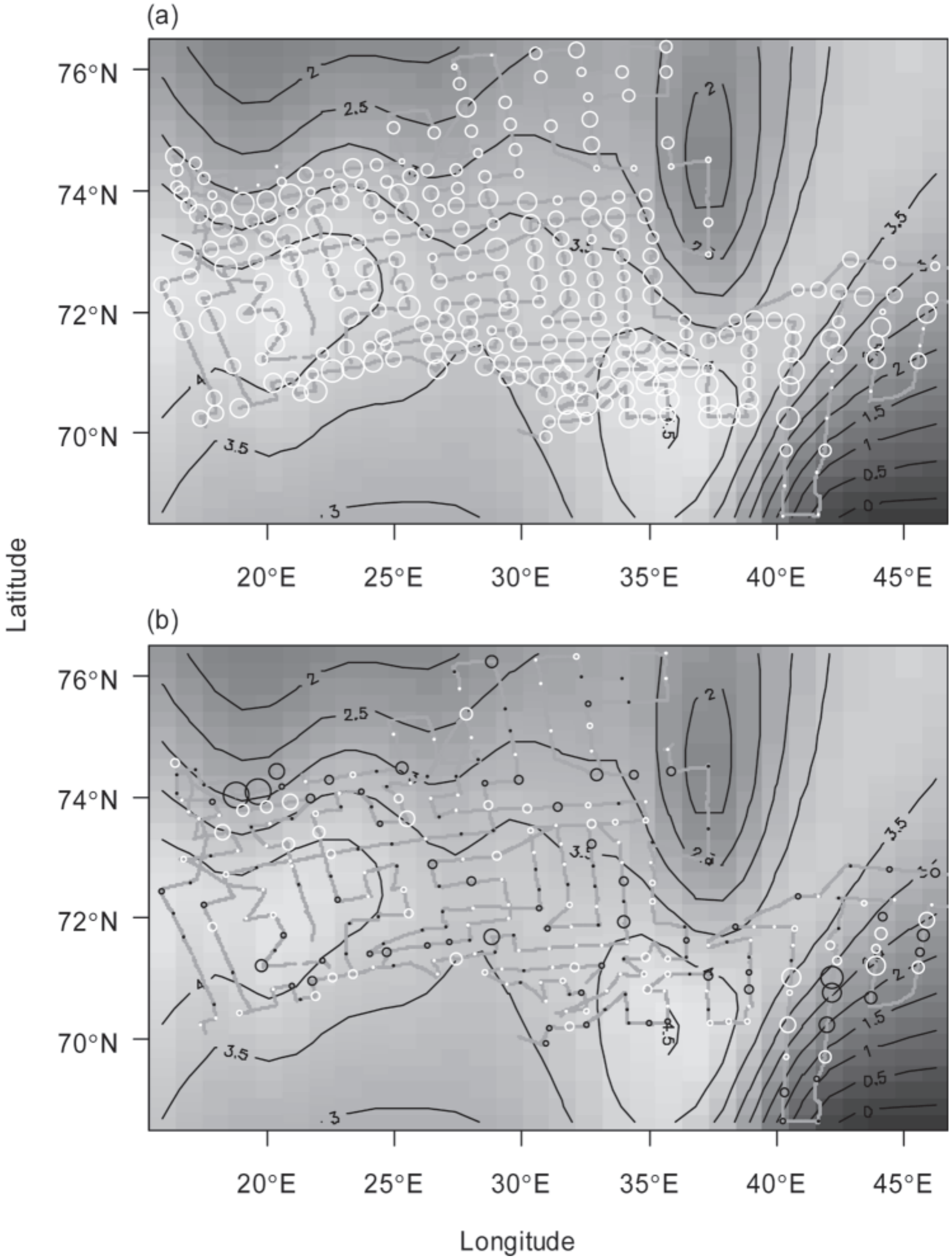
Because the interpretation of the results for the real data depends on the accuracy of eqs. 8 and 9, we present the results from the simulation experiments first (depicted in Figs. 2 and 3).

In the first experiment, the trawl samples $\{z_i, i = 1,\ldots,n_{\mathrm{ON}}\}$ and acoustic samples $\{\varepsilon_j, j = 1,\ldots,n_{\mathrm{ALL}}\}$ are described by the first-order autoregressive models of eq. 10, and the autocorrelations of $\{\varepsilon_j\}$ and $\{z_i\}$ are given by a and a^{20}, respectively. In this case, the theoretical variance reductions predicted by eqs. 8 and 9 are shown as curves (Fig. 2), with each curve having a fixed value for the correlation $\rho_{\varepsilon,z}$ between on-station trawl and acoustics. Equations A4 and A5 with the known values of a, $\sigma_X^2 = \sigma_\varepsilon^2$, and $\sigma_Y^2 = \sigma_z^2$ were used to calculate var($\bar{\varepsilon}_{\mathrm{ALL}}$) and var($\bar{z}_{\mathrm{ON}}$) in eq. 9. It is seen that the variance reduction decreases rapidly with increasing autocorrelation a for $a > 0.8$.

The results from the simulations in the first experiment are shown as points (Fig. 2). Each solid point shows the empirical variance reduction for 100 000 simulations for a given combination of autocorrelation a (0, 0.4, 0.6, 0.8, or 0.9) and trawl–acoustics correlation $\rho_{\varepsilon,z}$ (0.3, 0.5, 0.7, or 0.9). This is based on computing $\bar{z}_{\mathrm{ON}}$ and $I_C = \bar{z}_{\mathrm{ON}} + b_{\mathrm{ON}}(\bar{\varepsilon}_{\mathrm{ALL}} - \bar{e}_{\mathrm{ON}})$ for each realization and then computing $\Delta_V = 1 - s_I^2/s_z^2$, where s_I^2 and s_z^2 are the empirical variances of the 100 000 I_Cs and $\bar{z}_{\mathrm{ON}}$s, respectively. No approximation assumption or knowledge about the processes $\{\varepsilon_j\}$ and $\{z_i\}$ were used in the computations. The solid points are very close to the theoretical curves, demonstrating that the variance reduction formulas of eqs. 8 and 9 have high accuracy.

When working with real data, we have to estimate the quantities in eq. 9. Conventional estimates were used for $\rho_{\varepsilon,z}$, and β, and eq. A4 was used to estimate var($\bar{z}_{\mathrm{ON}}$) with var(z_i) and a estimated from the data. The block bootstrap was employed to estimate var($\bar{\varepsilon}_{\mathrm{ALL}}$) in order to have a robust estimation procedure that is not dependent on a particular model chosen for $\{\varepsilon_j\}$. The estimation was carried out for each of 1000 simulated realizations (because var(I_C) is much more time consuming to compute than I_C, Δ_V was only calculated for 1000 of the 100 000 realizations). For each realiza-

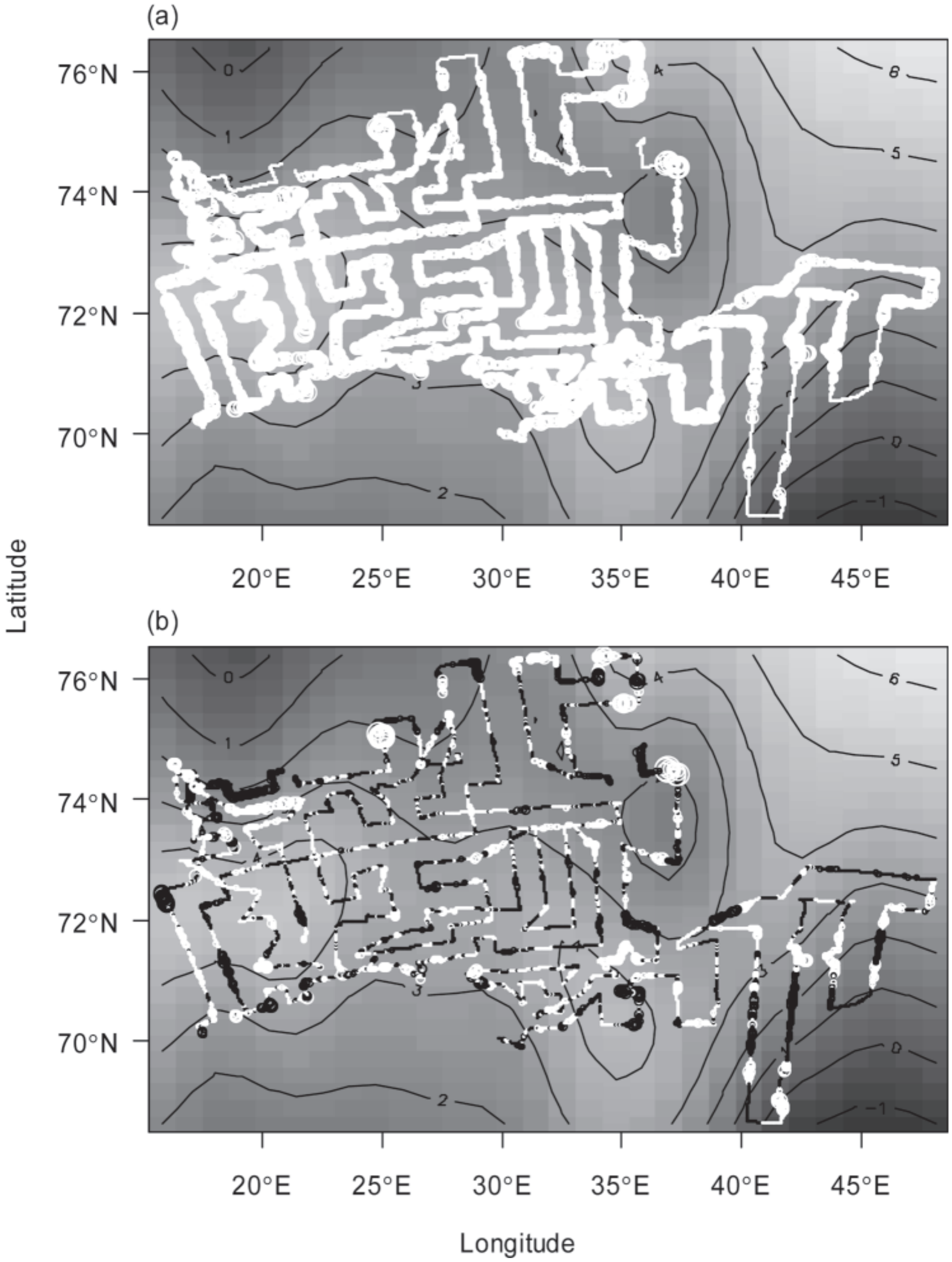
Fig 4. The GAM fitted to demersal trawl catches in 2002. The background shading indicates the level of the fitted surface $f_T(\mathrm{lon}_i, \mathrm{lat}_i)$, with light shading corresponding to a high level. In (*a*), the log-transformed trawl catches are proportional to the diameter of the circles. In (*b*), the absolute value of the residuals from the GAM fit is proportional to the diameter of the circles. Black and white circles indicate negative and positive residuals, respectively, and the underlying grey lines indicate the survey track.



tion, the estimated variance reduction Δ_V was computed using estimated quantities in eq. 9 (the averages over the 1000 realizations are shown as open circles in Fig. 2). Again, we obtained a very good correspondence with the theoretical curves except for the combination of high *a*s and high values of the correlation $\rho_{\varepsilon,z}$, where Δ_V was overestimated. If var($\bar{\varepsilon}_{\mathrm{ALL}}$) was calculated using the autoregressive bootstrap instead of the block bootstrap, there was no overestimation, but this is not a realistic approach for the real data. Increasing the block length to 500 and (or) the number of bootstrap replicas to 200 did not help. However, we do not think the overestimation represents much of a problem as we seldom observe pairs of values (a, $\rho_{\varepsilon,z}$) in the real data for which the overestimation in Fig. 2 is substantial.

The simulated variance reductions for the last two experiments mimicking the 2002 survey as described in the Material and methods section are shown (Fig. 3). On average, the variance reduction for the residual process with no GAM involved (Fig. 3*a*) is 0.067, whereas the theoretical variance reduction of Fig. 2 for the corresponding correlation and autocorrelation (a = 0.876 and $\rho_{\varepsilon,z}$ = 0.348) is 0.041 when there is one trawl station for each 20th acoustic sample. However, with n_{ALL} = 7383 and n_{ON} = 287, there is on average one trawl station for approximately each 26th acoustic

Fig 5. The GAM fitted to the demersal acoustics in 2002. The background shading indicates the level of the fitted surface $f_A(\mathrm{lon}_j, \mathrm{lat}_j)$, with light shading corresponding to a high level. In (*a*), the log-transformed echo abundance is proportional to the diameter of the circles. In (*b*), the absolute value of the residuals from the GAM fit is proportional to the diameter of the circles. Black and white circles indicate negative and positive residuals, respectively.



sample, which corresponds to a theoretical variance reduction of 0.055. The difference of 0.012 between this and the simulated variance reduction is as one would expect because of the tendency our method has to overestimate Δ_V for high autocorrelations (open circles in Fig. 2). Thus, the fact that the trawl stations are unequally spaced in experiment 2 (cf. the histogram in Fig. 1) does not seem to have much impact on the results. Including the GAM lead to a change in average Δ_V of only 5% to 0.070 (Fig. 3*b*).

We also did the simulations in the last two experiments with the first-order model of eq. 10 replaced by a 12th-order autoregressive model fitted to the 2002 demersal data. This resulted in a decrease of 0.01 in the average Δ_V both with and without the GAM included.

Real data

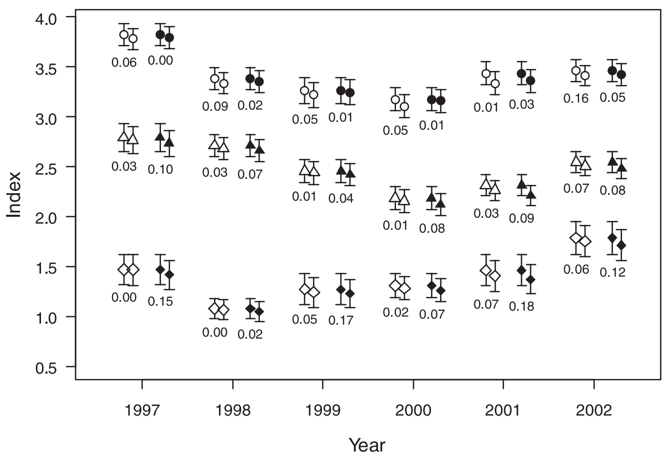
The data were log-transformed and modelled with latitude and longitude as explanatory variables, i.e.,

$$\log(T_i + 1) = f_T(\mathrm{lat}_i, \mathrm{lon}_i) + z_i$$

$$\log(A_j + 1) = f_A(\mathrm{lat}_j, \mathrm{lon}_j) + \varepsilon_i$$

where f_T and f_A are estimated nonparametrically by GAMs. A constant of 1 is added to the raw data in the log transfor-

Fig. 6. Trawl-only indices (left in each pair) and combined indices (right in each pair) for demersal fish (circles), cod (triangles), and haddock (diamonds). Solid symbols indicate scrutinized acoustics (cod, in the case of demersal catch). The data are log-transformed. The error bars indicate ±2 standard deviations, calculated from eq. 9 for the combined index. The numbers indicate variance reduction.



mation because then zero observations remain zero observations after the transformation, and the residuals become approximately normally distributed. The GAM fit to the demersal trawl catches in 2002 is shown together with $\log(T_i + 1)$ and z_i (Fig. 4). The corresponding plots for the acoustics are also provided (Fig. 5). The contours of the surfaces are similar for trawl and acoustics, and for both variables, there is a tendency to have the largest residuals (in absolute value) close to the borders of the survey area. This is reasonable as the scarcity of data here makes the GAM fit less reliable. A careful study of Figs. 4 and 5 reveals that the GAM is quite effective in removing spatial trends, and actually, the GAM fit results in a variance reduction close to 50% for 2002 (cf. Fig. 8 in the Discussion). A strong autocorrelation along the transect remains for the acoustic residuals (Fig. 5).

In some of the acoustic series, there are periods with no fish and hence $\varepsilon_j \approx \varepsilon_{j-1}$. Because of this nonstationarity in the $\{\varepsilon_j\}$ series, block bootstrapping (see Appendix A) with 1000 replications and a relatively long block length of 100 was used to calculate $\mathrm{var}(\bar{\varepsilon}_{\mathrm{ALL}})$, whereas eq. A4 was employed to calculate $\mathrm{var}(\bar{z}_{\mathrm{ON}})$. The variance reduction Δ_V computed using eq. 9 was on average about 0.06, ranging from 0 to 0.18 (Fig. 6). For cod and haddock, in all years, Δ_V was largest when the scrutinized acoustics was used (0.10 on average), and for demersal fish, it was typically largest when the demersal acoustics was used. Note also that in almost all cases, the combined index has a lower value than the trawl-only index. The variance reduction clearly increases with the correlation between trawl and acoustics (Fig. 7*a*), but more unexpectedly, it also tends to increase with increasing autocorrelation (Fig. 7*b*). This seems to be because high correlation between trawl and acoustics tends to occur together with high autocorrelation in acoustics (Fig. 7*c*). For a fixed correlation, the variance reduction decreases with increasing autocorrelation.

Discussion

The variance reduction obtained for the Barents Sea bottom trawl survey estimate by using acoustics as an auxiliary variable in the combined sampling approach was relatively small (typically less than 10% on the log scale). This is due to the high autocorrelation in the acoustic residuals (which means that each acoustic sample does not carry much independent additional information) combined with a relatively low and variable correlation between trawl and acoustics, and the observed variance reductions are approximately as one would expect (Fig. 2). For a lag-one autocorrelation around 0.9, the variance reduction for first-order autoregressive processes is quite small for moderate correlations (Fig. 2). For higher-order autoregressive processes, the situation is even worse because the autocorrelation decreases less rapidly as the lag increases. For the Barents Sea data, the autocorrelation varies a lot between years and acoustic fish categories. The lag-20 autocorrelation varies from about 0.04 (demersal 1998) to about 0.41 (haddock 2002). The fact that low acoustic autocorrelation is generally associated with low correlation between trawl and acoustics and vice versa does not help. It may indicate that there is a latent variable or trend with which both the acoustics and trawl data are positively correlated.

Because the scrutinized acoustic density is more strongly correlated with the trawl catches than is the total acoustic density, one would expect the best results when using the former as an auxiliary variable. However, because the scrutinizing process is based to some extent on the species distribution in the trawl catches, there is, in principle, a possibility that the scrutinized acoustic density is biased (for example, if there is a high percentage of cod in the catch at a given trawl station, much of the acoustic echo abundance in the vicinity of that station may be allocated to cod). This bias could yield a too high trawl–acoustics correlation and hence a too high estimated variance reduction. Moreover, the scrutinized data have slightly higher autocorrelation. The use of total acoustic density would in this sense be safer.

The effect of simplifying assumptions

The time series approach is a simplification with regards to the Barents Sea data, because the spatial aspect is ignored. However, by considering the residuals after fitting a GAM, using latitude and longitude as explanatory variables, most of the large-scale spatial trends should be removed, and the autocorrelation along the survey transect is the dominating one. The reduction in total variation obtained by including latitude and longitude as explanatory variables is much larger than the additional reduction obtained by including acoustics (Fig. 8; see Mackinson et al. (2005) for corresponding results for North Sea surveys). Also, the combined index in eqs. 4–9 is computed for log-transformed data, yielding log indices, and it is not trivial to get a back-transformed combined index with an appropriate uncertainty measure. On the other hand, "real" trawl-only indices with uncertainty estimates can be calculated using the GAM approach, for example, in combination with bootstrapping.

Another simplification is the treatment of the acoustic residuals within one year as a contiguous time series, despite

Fig. 7. Variance reduction (Δ_V) plotted against (*a*) correlation between trawl and acoustics and (*b*) autocorrelation in acoustic residuals and (*c*) autocorrelation plotted against trawl–acoustics correlation. Each point corresponds to one of the pairs in Fig. 6. Circles, triangles, and diamonds represent demersal fish, cod, and haddock, respectively. Open and solid symbols represent demersal and scrutinized acoustics (cod, in the case of demersal catch), respectively.

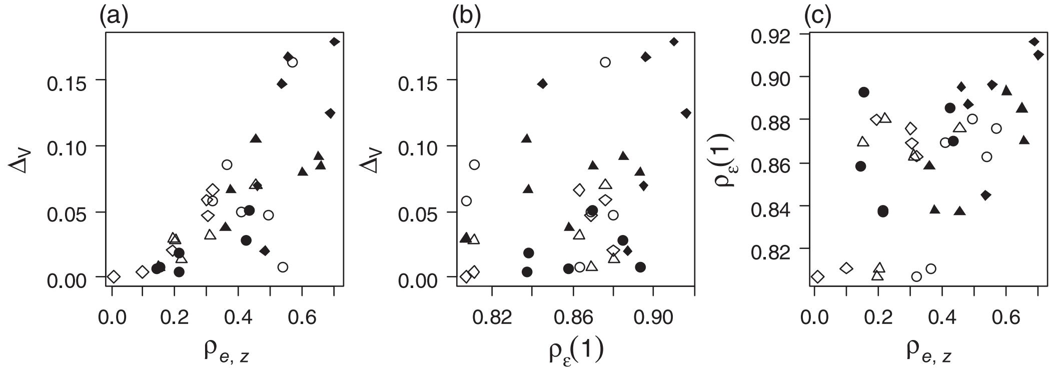
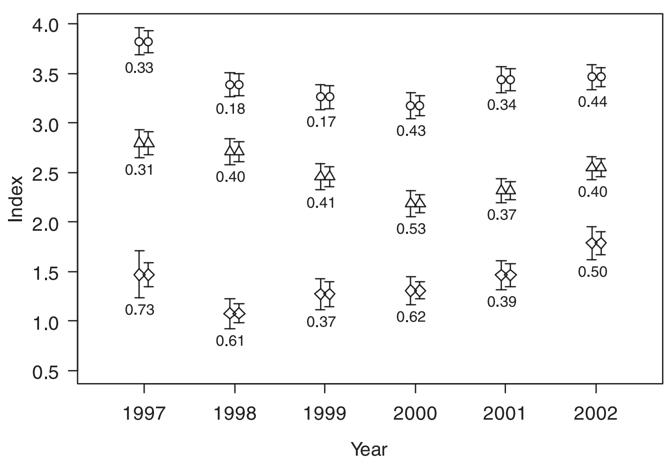


Fig. 8. Mean catch $\overline{T} \pm 2$ bootstrapped standard errors for demersal fish (circles), cod (triangles), and haddock (diamonds). The standard errors are calculated by bootstrapping the catches themselves (left in each pair) or residuals from a fitted GAM with latitude and longitude as explanatory variables (right in each pair). Numbers: variance reduction defined as $\Delta_V = 1 - \mathrm{var}(\overline{z})/\mathrm{var}(\overline{T})$, where $\{z_i\}$ are the residuals from the GAM fit.



the fact that two or three vessels are involved. There are typically some discontinuities in the series for each vessel as well. However, the number of discontinuities is small compared with the total number of samples, and taking these into account would probably not have improved the results.

Further, the on-station and between-station acoustic observations are not subsamples of the same process. The elementary sampling distance unit (EDSU) is 1 n.mi, and the between-station samples cover 1 EDSU each, whereas the on-station acoustic densities are weighted averages of two or three successive EDSUs overlapping the towed distance of 1.5 n.mi. However, setting the on-station acoustic density equal to the density of the first EDSU covered by the trawl station did not improve the variance of the combined index. Adjusting trawl catches and acoustic densities for diurnal variation did not lead to appreciable changes either. If the on-station acoustic abundance for each trawl station had been integrated exactly over the trawling distance, then the trawl–acoustics correlations would probably have been slightly higher, but it is dubious whether the variance reduction would increase much. By increasing the EDSU to, for example, 2 n.mi, one would obtain a lower autocorrelation in the acoustic data, but the number of acoustic samples would decrease correspondingly, so the variance reduction would not improve. The tendency of overestimation for high autocorrelations (cf. the simulation experiments) would, however, be smaller.

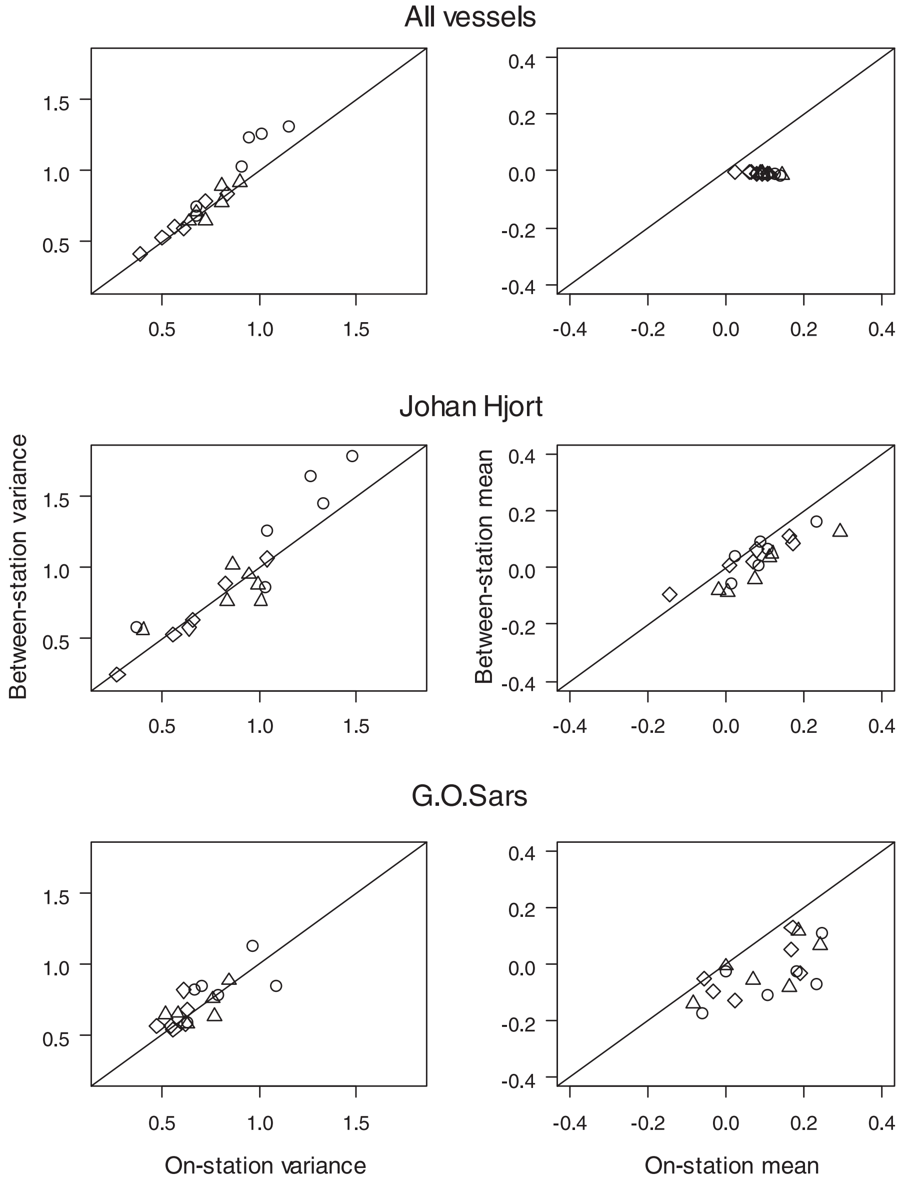
Our analysis is based on stationarity, and this assumption seems roughly to be fulfilled as regards cod and demersal fish, but in the haddock time series, there are periods with no fish ($A_j = 0$) and hence $\varepsilon_j \approx \varepsilon_{j-1}$. Removing these periods, however, did not influence the autocorrelation much.

On- and between-station differences

The properties of the acoustic residuals between stations are different from those on stations (Fig. 9). Overall, the variance is slightly higher between stations, whereas the means between stations are considerably lower. On average, the variance is about 7% higher between stations, and to check for the effect that this has on the estimated variance reduction, we did a simulation experiment in which the between-station εs in eq. 11 were multiplied by $\sqrt{1.07}$. For $\rho_{\varepsilon,z} = 0.6$, this led to a reduction in Δ_V from 0.079 to 0.062.

In this study, we have seen that the combined index almost always yields a lower value than the trawl-only index. The reason for this is that the on-station means, particularly for *G.O. Sars*, are higher than the between-station means (Fig. 9; in the upper right plot, the between-station means are close to zero because $\overline{\hat{\varepsilon}}_{\mathrm{ALL}} = 0$ due to properties of the GAM, and $\overline{\hat{\varepsilon}}_{\mathrm{BETW}} \approx \overline{\hat{\varepsilon}}_{\mathrm{ALL}}$ because the on-station ε_js are few). Taken together with the lower variance on station, this indicates that during trawling the vessel (in particular *G.O. Sars*) may in some sense act as a fish-attracting device (cf. Røstad et al. 2006). The findings in Handegard and Tjøstheim (2005) support this hypothesis. An alternative explanation is that the behavior of fish alerted by, e.g., a trawling vessel will change towards a more aimed and coordinated swimming pattern resulting in a reduction in average tilt angle. The result will be an increased average target strength as is observed during trawling. It should be noted that this runs contrary to some earlier results (Ona and Godø 1990): based on 134 trawl stations from 1985 to 1986, taken by vessels other than *Johan Hjort* and *G.O. Sars*, the acoustic density

Fig. 9. On- and between-station variances (left plots) and means (right plots) of ε_j. Each point corresponds to one of the acoustic categories and years in Fig. 6 (circles, triangles, and diamonds represent demersal fish, cod, and haddock, respectively). The solid lines indicate $y = x$ (equal variance and mean). In the upper plots, all samples from each survey are used; in the lower plots, only samples taken by *Johan Hjort* or *G.O. Sars* are used.



was lower during trawling than before or after. We consider the results illustrated in Fig. 9 to be important and requiring further investigation.

The higher on-station means support the findings of Bez et al. (2007). They demonstrated a generally high consistency between on-station and between-station acoustics for several surveys, including the Barents Sea winter survey, using tools like the variogram and a global index of collocation. However, the acoustic density in the bottom layers tended to be higher during trawling than just before or after trawling. In the current analysis, we have also used the acoustic density in the bottom layers (up to 50 m above the bottom), but for the whole water column the echo abundance for demersal fish is on average 12% lower before and after trawling than during trawling for *G.O. Sars* and 1% lower for *Johan Hjort* (N.O. Handegard and V. Hjellvik, unpublished data).

Alternative combination methods?

In conclusion, the precision of the Barents Sea bottom trawl survey index cannot be increased much by using the simultaneously collected acoustic data, at least not using standard statistical techniques as suggested in this paper. Moreover, from eqs. 9 and A4, such a result can always be expected in the case where one has high autocorrelation of the acoustics and quite low correlation between acoustics and trawl catches. It seems to be difficult to get rid of the autocorrelation for the acoustics. If there is an autocorrelation of 0.8, a correlation of about 0.9 is required between trawl and acoustics to obtain a 50% variance reduction (on the log scale if the data are log-transformed).

An alternative approach could be to employ the acoustic data to allocate trawl effort in an adaptive sampling design (Ona et al. 1991; Everson et al. 1996). Further, there is no doubt that the acoustic data do contain information not present in the trawl data. For example, the bottom trawl does not catch the fish distributed in the upper part of the water column, whereas they are detected by the acoustics. Thus, an intuitive way of combining trawl and acoustic data is to add the acoustic density above the effective fishing height of the trawl to the trawl catch. As is well known, this is not an easy task. The effective fishing height of the trawl is higher than the physical height of the trawl, because fish situated higher in the water column are to some extent disturbed and move downwards to the bottom where they are caught by the trawl. However, the details of this mechanism are not well understood (Handegard and Tjøstheim 2005; Hjellvik et al. 2003). Both survey methods are, for example, sensitive to diurnal variation and vertical distribution phenomena (Hjellvik et al. 2002, 2004), but such systematic influential factors may affect the relationship in density estimates in a manner that we are not yet able to take into account (Godø 1994). Also, the length and species composition of fish is typically not the same close to the bottom as higher in the water column. Using the composition in the trawl catch to allocate the acoustic backscatter to species and length groups will therefore bias the combined estimate. It thus appears that an efficient combination of trawl and acoustic estimates of density still needs improved quantitative understanding of the complex behavioural processes influencing trawl and acoustic efficiency.

Acknowledgements

We thank the European Union for financial support through the CATEFA (Combined Acoustic and Trawl data for Estimating fish Abundance) project (Q5RS-2001-02038). We are grateful for a number of very valuable suggestions and comments from the referees and an associate editor on an earlier version. In particular, the GAM fitting of all, not only the on-station part, of the acoustic data in eq. 5 is due to a suggestion from one of the referees.

References

Beare, D.J., Reid, D.G., Greig, T., Bez, N., Hjellvik, V., Godø, O.R., Bouleau, M., van der Kooij, J., Neville, S., and Mackinson, S. 2004. Positive relationships between bottom trawl and acoustic data. ICES CM 2004/R:24.

Bez, N., Reid, D., Neville, S., Vérin, Y., Hjellvik, V., and Gerritsen, H.D. 2007. Acoustic data collected during and between bottom trawl stations: consistency and common trends. Can. J. Fish. Aquat. Sci. **64**: 166–180.

Bouleau, M., Bez, N., Reid, D.G., Godø, O.R., and Gerritsen, H. 2004. Testing various geostatistical models to combine bottom trawl catches and acoustic data. ICES CM 2004/R:28.

Cachera, S., Massé, J., and Vérin, Y. 1999. How the use of acoustics during bottom trawl surveys may provide more accurate abundance indices: an application to IBTS surveys carried out in the southern North Sea. ICES CM 1999/J:12.

Cochran, W.G. 1977. Sampling techniques. Wiley, New York.

Everson, I., Bravington, M., and Goss, C. 1996. A combined acoustic and trawl survey for efficiently estimating fish abundance. Fish. Res. **26**: 75–91.

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

Handegard, N.O., and Tjøstheim, D. 2005. When fish meet a trawling vessel: examining the behaviour of gadoids using a free-floating buoy and acoustic splitbeam tracking. Can. J. Fish. Aquat. Sci. **62**: 2409–2422.

Hjellvik, V., Godø, O.R., and Tjøstheim, D. 2002. Diurnal variation in bottom trawl survey catches: does it pay to adjust? Can. J. Fish. Aquat. Sci. **59**: 33–48.

Hjellvik, V., Michalsen, K., Aglen, A., and Nakken, O. 2003. An attempt at estimating the effective fishing height of the bottom trawl using acoustic survey recordings. ICES J. Mar. Sci. **60**: 967–979.

Hjellvik, V., Godø, O.R., and Tjøstheim, D. 2004. Diurnal variation in acoustic densities: why do we see less in the dark? Can. J. Fish. Aquat. Sci. **61**: 2237–2254.

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 1991/Y:17.

Mackinson, S., van der Kooij, J., and Neville, S. 2005. The fuzzy relation between trawl and acoustic surveys in the North Sea. ICES J. Mar. Sci. **62**: 1556–1575.

Neville, S., Hjellvik, V., Mackinson, S., and van der Kooij, J. 2004. Using artificial neural networks to combine acoustics and trawls in the Barents and North Seas. ICES CM 2004/R:05.

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.

Ona, E., Pennington, M., and Vølstad, J.H. 1991. Using acoustics to improve the precision of bottom trawl indices of abundance. ICES CM 1991/D:13.

Røstad, A., Kaartvedt, S., Klevjer, T.A., and Melle, W. 2006. Fish are attracted to vessels. ICES J. Mar. Sci. **63**: 1431–1437.

List of symbols

Symbol	Meaning
T	Trawl catch
A	Acoustic sample
n	Number of samples
i, ON	Subscripts used for on-station trawl and acoustics
j, ALL	Subscripts used for all acoustic samples
z	Residuals from regression model fitted to trawl catches (eq. 4)
ε	Residuals from regression model fitted to all acoustic samples (eq. 5)
e	On-station acoustic “residuals” (eq. 6)
β	True regression coefficient between trawl and acoustics
b_{ON}	Estimated regression coefficient between trawl and acoustics
$\rho_{X,Y}$	Correlation between X and Y
$\rho_X(k)$	Lag k autocorrelation of X
I_C	Combined index (eq. 7)
Δ_V	Variance reduction
L	Block length in block bootstrap
*	Superscript indicating bootstrap sample

Appendix A

Derivation of eq. 3

Using eq. 1 twice and eq. 2, elementary algebra yields

$$\text{(A1)}\quad \begin{aligned}\hat{\mu}(T)-\mu(T) &= I_C-\mu(T)\\ &= \overline{T}_{\mathrm{ON}}+b_{\mathrm{ON}}(\overline{A}_{\mathrm{ALL}}-\overline{A}_{\mathrm{ON}})-\alpha-\beta\mu(A)\\ &= \alpha+\beta\overline{A}_{\mathrm{ON}}+\overline{u}_{\mathrm{ON}}+b_{\mathrm{ON}}(\overline{A}_{\mathrm{ALL}}-\overline{A}_{\mathrm{ON}})-\alpha-\beta\mu(A)\\ &= \beta[\overline{A}_{\mathrm{ALL}}-\mu(A)]+\overline{u}_{\mathrm{ON}}+(b_{\mathrm{ON}}-\beta)(\overline{A}_{\mathrm{ALL}}-\overline{A}_{\mathrm{ON}})\end{aligned}$$

When n_{ON} gets large, the last product term in the last line is clearly of smaller order and can be neglected as both $(b_{\mathrm{ON}}-\beta)$ and $(\overline{A}_{\mathrm{ALL}}-\overline{A}_{\mathrm{ON}})$ go to zero approximately at the same rate as the first two terms. Because of the independence of A_j, $j = 1,\ldots,n_{\mathrm{ALL}}$ and u_i, $i = 1,\ldots,n_{\mathrm{ON}}$, with this approximation we can write

$$\text{(A2)}\quad \mathrm{var}(I_C)=\beta^2\frac{\mathrm{var}(A_i)}{n_{\mathrm{ALL}}}+\frac{\mathrm{var}(u_i)}{n_{\mathrm{ON}}}$$

(see, e.g., Hogg and Tanis 2006, p. 211). It is a well-known result in standard linear regression (see, e.g., Hogg and Tanis 2006, p. 211) that $\mathrm{var}(u_i) = (1-\rho^2_{A,T})\mathrm{var}(T_i)$ and $\beta^2\mathrm{var}(A_i) = \rho^2_{A,T}\mathrm{var}(T_i)$, from which eq. 3 follows.

Derivation of eq. 8

For the derivation of eq. 8, we replace T by z, $\overline{A}_{\mathrm{ALL}}$ by $\overline{\varepsilon}_{\mathrm{ALL}}$, and $\overline{A}_{\mathrm{ON}}$ by $\overline{e}_{\mathrm{ON}}$ in the above derivation of eq. 3 and obtain from eq. A1 with $I_C' = \overline{z}_{\mathrm{ON}}+b_{\mathrm{ON}}(\overline{\varepsilon}_{\mathrm{ALL}}-\overline{e}_{\mathrm{ON}})$,

$$I_C'-\mu(I_C')=\beta[\overline{\varepsilon}_{\mathrm{ALL}}-\mu(\varepsilon)]+\overline{u}_{\mathrm{ON}}+(b_{\mathrm{ON}}-\beta)(\overline{\varepsilon}_{\mathrm{ALL}}-\overline{e}_{\mathrm{ON}})$$

Using exactly the same approximation as in the derivation of eq. A2 and using independence of $\{\varepsilon_i\}$ and $\{u_i\}$, we can write

$$\mathrm{var}(I_C)=\mathrm{var}(I_C')=\beta^2\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})+\mathrm{var}(\overline{u}_{\mathrm{ON}})$$

If the on-station observations are taken so far apart that they can be considered independent, the standard result (Hogg and Tanis 2006, p. 211) quoted in the derivation of eq. 3 holds, i.e.,

$$\text{(A3)}\quad \mathrm{var}(\overline{u}_{\mathrm{ON}})=\frac{\mathrm{var}(u_i)}{n_{\mathrm{ON}}}=(1-\rho^2_{e,z})\frac{\mathrm{var}(z_i)}{n_{\mathrm{ON}}}$$

If $\{u_i\}$, $\{e_i\}$, and $\{z_i\}$ are allowed to be weakly autocorrelated, but with $\{u_i\}$ independent of $\{e_i\}$, eq. A3 can be replaced by

$$\mathrm{var}(\overline{u}_{\mathrm{ON}})=(1-\rho^2_{e,z})\,\mathrm{var}(\overline{z}_{\mathrm{ON}})$$

and eq. 8 follows.

Estimating the variance of the mean of an autocorrelated variable

In the variance reduction formula in eq. 9 we need to compute $\mathrm{var}(\overline{\varepsilon}_{\mathrm{ALL}})$ and $\mathrm{var}(\overline{z}_{\mathrm{ON}})$ for autocorrelated $\{z_i\}$ and $\{\varepsilon_j\}$. Generally, if the variables X_i, $i = 1,\ldots,n$, follow a first-order autoregressive process, the calculation of the variance of the mean is straightforward. If

$$X_i = aX_{i-1}+u_i$$

where the u_is are independent, zero mean, and identically distributed with $\mathrm{var}(u_i) = \sigma_u^2$, we have $\mathrm{var}(X_i) = \sigma_X^2 = \sigma_u^2(1-a^2)^{-1}$ and

$$\text{(A4)}\quad \mathrm{var}(\overline{X})=\frac{\sigma_X^2}{n^2}\left\{n+2\sum_{i=1}^{n-1} i a^{n-i}\right\}$$

(see, e.g., Brockwell and Davis 1996, chapters 2.3 and 2.4). In the special case when $a = 0$, it follows that $\mathrm{var}(\overline{X}) = \sigma_X^2/n$, and as $a \to 1$, we have $\mathrm{var}(\overline{X}) \to \sigma_X^2$, these two extremes having very different effects on the variance reduction.

Let $Y_i = X_{20i}$ for $i = 1,\ldots,n_{20}$, where $n_{20} = n/20$. In other words, the sequence $\{Y_i\}$ contains every 20th observation of $\{X_i\}$. Then we have

$$\text{(A5)}\quad \mathrm{var}(\overline{Y})=\frac{\sigma_Y^2}{n_{20}^2}\left\{n_{20}+2\sum_{i=1}^{n_{20}-1} i a^{20(n_{20}-i)}\right\}$$

For higher-order autoregressive processes, the variance of the mean is more complicated to calculate analytically as a function of the autoregressive coefficients. It can be easily expressed (Brockwell and Davis 1996, p. 56) as a function of the autocovariances, but estimates of higher-order autocovariances are inaccurate. A better alternative is to use the autoregressive bootstrap as follows. Fit an autoregressive model to $\{X_i\}$ by estimating the autoregressive order and the autoregressive coefficients. Then, create m bootstrapped time series realizations $X^*_{1,k},\ldots,X^*_{n,k}$, $k = 1,\ldots,m$, using the estimated autoregressive model. For each bootstrap realization, calculate the sample mean $\bar{X}^*_k$ and then estimate $\mathrm{var}(\bar{X})$ by $(m-1)^{-1}\sum_{k=1}^{m}(\bar{X}^*_k - \bar{X}^*)^2$, where $\bar{X}^* = m^{-1}\sum_{k=1}^{m}\bar{X}^*_k$.

If it is difficult to fit an autoregressive time series model to the data, then the block bootstrap (Politis et al. 1999) is an option assuming the time series to be essentially stationary (Politis et al. 1999, p. 101). There are several ways to block-bootstrap time series data. We have applied the method described in Politis et al. (1999, section 3.9). Given a set of observations, $X_1,\ldots,X_n$, generate a bootstrap replica by drawing with replacement n/L contiguous blocks of length L from $X_1,\ldots,X_n$. For example, with $n = 9$ and $L = 3$, the blocks may be X_7, X_8, X_9 and X_2, X_3, X_4 and X_6, X_7, X_8. Then, generate m bootstrap replicas, and for each bootstrap replica $\{X^*_k\}$, $k = 1,\ldots,m$, calculate the sample mean $\bar{X}^*_k$. Finally, estimate $\mathrm{var}(\bar{X})$ by $(m-1)^{-1}\sum_{k=1}^{m}(\bar{X}^*_k - \bar{X}^*)^2$.

Simulation of two time series with a given correlation and autocorrelation

In the simulation experiments, we need to generate time series having specific correlation and autocorrelation properties. To simulate two time series $\{\varepsilon_j\}$ and $\{z_j\}$ with a given correlation $\rho_{\varepsilon,z}$, autocorrelations $\rho_\varepsilon(1) = \rho_z(1) = a$ and variances σ^2_ε and σ^2_z, we can use the model (again we refer to Brockwell and Davis 1996, chapters 2.1–2.4, for correlation properties of autoregressive and ARMA time series)

$$\varepsilon_j = a\varepsilon_{j-1} + u_j + w_j, \quad u_j \sim N(0,\sigma^2_u), \quad w_j \sim N(0,\sigma^2_w), \quad j = 1,\ldots,n_{\mathrm{ALL}}$$

$$z_j = a z_{j-1} + v_j + w_j, \quad v_j \sim N(0,\sigma^2_v), \quad j = 1,\ldots,n_{\mathrm{ALL}}$$

which gives a correlation

$$\text{(A6)} \qquad \rho_{\varepsilon,z} = \mathrm{corr}(\varepsilon_j, z_j) = \frac{(1-a^2)\sigma_{\varepsilon,z}}{\sqrt{(\sigma^2_u + \sigma^2_w)(\sigma^2_v + \sigma^2_w)}}$$

where $\sigma_{\varepsilon,z} = \mathrm{Cov}(\varepsilon_j, z_j) = E(\varepsilon_j z_j)$. Simulated realizations are generated by using a Gaussian random number generator to generate time series $\{u_j\}$, $\{v_j\}$, and $\{w_j\}$. In order to do this, we need to specify σ^2_u, σ^2_v, and σ^2_w. Because $E(\varepsilon_j z_j) = E(\varepsilon_{j-1} z_{j-1})$, we have

$$E(\varepsilon_j z_j) = a^2 E(\varepsilon_j z_j) + \sigma^2_w$$

which means that

$$\text{(A7)} \qquad \sigma_{e,z} = \frac{\sigma^2_w}{1-a^2}$$

Putting eq. A7 into eq. A6 we get

$$\text{(A8)} \qquad \rho_{\varepsilon,z} = \frac{\sigma^2_w}{\sqrt{(\sigma^2_u + \sigma^2_w)(\sigma^2_v + \sigma^2_w)}} = \frac{\sigma^2_w}{(1-a^2)\sigma_\varepsilon\sigma_z}$$

To generate time series with a given σ^2_ε, σ^2_z, and $\rho_{\varepsilon,z}$ under the restriction $\rho_{\varepsilon,z} < \min(\sigma_z/\sigma_\varepsilon, \sigma_\varepsilon/\sigma_z)$, choose

$\sigma^2_w = \sigma_\varepsilon\,\sigma_z\,\rho_{\varepsilon,z}(1-a^2)$, $\sigma^2_u = \sigma^2_\varepsilon(1-a^2) - \sigma^2_w$, and $\sigma^2_v = \sigma^2_z(1-a^2) - \sigma^2_w$

References

Brockwell, P.J., and Davis, R.A. 1996. Introduction to times series and forecasting. Springer, New York.
Hogg, R.V., and Tanis, E.A. 2006. Probability and statistical inference. 7th ed. Pearson Education, London.
Politis, D.N., Wolf, M., and Romano, J.P. 1999. Subsampling. Springer, New York.