ICES Journal of Marine Science

ICES Journal of Marine Science (2016), 73(9), 2208–2226. doi:10.1093/icesjms/fsw054

Original Article

Evaluating total uncertainty for biomass- and abundance-at-age estimates from eastern Bering Sea walleye pollock acoustic-trawl surveys

Mathieu Woillez[1,2†]*, Paul D. Walline[1], James N. Ianelli[1], Martin W. Dorn[1], Christopher D. Wilson[1], and Andre E. Punt[2]

[1]*Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115, USA*
[2]*School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195, USA*

**Corresponding author: tel: +33 229008565; fax: +33 229008547; e-mail: mathieu.woillez@ifremer.fr*
[†]*Present address: Ifremer, Sciences et Technologies Halieutiques, CS 10070, 29280 Plouzané, France.*

Woillez, M., Walline, P. D., Ianelli, J. N., Dorn, M. W., Wilson, C. D., and Punt, A. E. Evaluating total uncertainty for biomass- and abundance-at-age estimates from eastern Bering Sea walleye pollock acoustic-trawl surveys. – ICES Journal of Marine Science, 73: 2208–2226.

Received 17 September 2015; revised 5 February 2016; accepted 13 March 2016; advance access publication 5 May 2016.

A comprehensive evaluation of the uncertainty of acoustic-trawl survey estimates is needed to appropriately include them in stock assessments. However, this evaluation is not straightforward because various data types (acoustic backscatter, length, weight, and age composition) are combined to produce estimates of abundance- and biomass-at-age. Uncertainties associated with each data type and those from functional relationships among variables need to be evaluated and combined. Uncertainty due to spatial sampling is evaluated using geostatistical conditional (co-) simulations. Multiple realizations of acoustic backscatter were produced using transformed Gaussian simulations with a Gibbs sampler to handle zeros. Multiple realizations of length frequency distributions were produced using transformed multivariate Gaussian co-simulations derived from quantiles of the empirical length distributions. Uncertainty due to errors in functional relationships was evaluated using bootstrap for the target strength-at-length and the weight-at-length relationships and for age – length keys. The contribution of each of these major sources of uncertainty was assessed for acoustic-trawl surveys of walleye pollock in the eastern Bering Sea in 2006 – 2010. This simulation framework allows a general computation for estimating abundance- and biomass-at-age variance – covariance matrices. Such estimates suggest that the covariance structure assumed in fitting stock assessment models differs substantially from what careful analysis of survey data actually indicate.

Keywords: acoustic-trawl surveys, biomass and abundance estimates, geostatistics, stock assessment, uncertainty evaluation.

Introduction

A critical feature of all modern stock assessment approaches is characterization of uncertainty (Prager *et al.*, 2003; Shertzer *et al.*, 2008; Hanselman, 2009). Basic to this is reliable estimation of survey uncertainty. The acoustic-trawl (AT) survey estimates of total numbers (age 2 and older) are assumed to have a coefficient of variation (CV) of 20% in assessment models currently in use for eastern Bering Sea (EBS) walleye pollock (*Gadus chalcogrammus*) despite variance estimates (e.g. Williamson and Traynor, 1996; Honkalehto *et al.*, 2012) from one-dimensional geostatistical transitive analyses (Petitgas, 1993; Bez, 2002) that suggest sampling CVs between 3.9 and 8.8% over 2006–2010. Further, somewhat arbitrarily, an effective sample size for the age-composition data is specified to be proportional to the number of trawl-hauls conducted each year, and it is assumed that the age-composition data follow a multinomial distribution (Ianelli *et al.*, 2011). However, a careful analysis of uncertainty in the abundance- and biomass-at-age estimates derived from the acoustic survey is possible and could be included in assessment models.

A comprehensive analysis of uncertainty is not straightforward for AT surveys, because a variety of data types, including the acoustic backscatter recorded by the echosounder along the vessel track and biological information such as fish length, weight and age obtained from trawl samples, need to be combined to estimate abundance- or

biomass-at-age (Simmonds and MacLennan, 2005). Indeed, whereas sampling only takes place at some spatial locations (acoustic transects and trawl stations), abundance- or biomass-at-age estimates are obtained by integrating predicted values for all cells of a grid covering the survey area. Consequently, uncertainty related to spatial sampling error (i.e. error made when predicting values between sample locations), measurement error (i.e. error made when measuring values at sample locations, i.e. the nugget effect) and error on the functional links needs to be evaluated and accounted for. Earlier works have focused either on the measurement error assuming statistical distributions for acoustic parameters generated using Monte-Carlo simulations (e.g. Rose *et al.*, 2000; Demer, 2004; O'Driscoll, 2004), or on spatial sampling and measurement errors estimated using spline-based generalized additive models (Zwolinski *et al.*, 2009), bootstrap (Simmonds *et al.*, 2009) or geostatistical simulations (Gimona and Fernandes 2003; Walline 2007; Simmonds *et al.*, 2009; Woillez *et al.*, 2009) to combine trawl and acoustic data. To our knowledge, only one attempt has been made to address both measurements, spatial sampling, and functional links error; classic bootstrap was used to evaluate errors on the functional links and measures, while block and circular bootstraps were used to characterize spatial sampling errors (Løland *et al.*, 2007).

A simulation framework based on geostatistical simulations and bootstrap techniques is proposed here to more fully quantify sources of uncertainty in EBS walleye pollock AT surveys. This work builds on geostatistical simulations developed for EBS walleye pollock (Walline, 2007). Transformed Gaussian geostatistical simulations are used to capture the uncertainty associated with the acoustic backscatter (Woillez, 2007; Woillez *et al.*, 2009), while uncertainty associated with the age data is taken into account through simulation of the spatial distribution of fish lengths, combined with bootstrap sampling of the age–length data. Earlier studies (e.g. Rivoirard *et al.*, 2000; Gimona and Fernandes, 2003; Walline, 2007; Woillez *et al.*, 2009) used only a single summary statistic (i.e. mean length or root mean squared length) to represent the length frequency distribution. Age data were not considered, and only uncertainty of total abundance or biomass were provided, except for Woillez *et al.* (2009), who partitioned the estimate of total abundance to fish age by simulating proportions-at-age as an additional set of realizations. Use of a single statistic to describe a length frequency distribution is only appropriate when distributions are narrow and unimodal. This might be the case for some small pelagic fish such as herring (Woillez *et al.*, 2009), but not for EBS walleye pollock (Honkalehto *et al.*, 2012). Here, we outline a more general set of summary statistics, quantiles of length data, to develop a multivariate geostatistical model. Uncertainties are accounted for due to variability in the relationships between target strength and length, weight and length, and age and length. Additionally, spatial sampling error associated with acoustic and length frequency data is evaluated. Finally, the uncertainty for the total abundance and biomass and for abundance- and biomass-at-age is evaluated by combining errors from all sources. The relative contribution of each source of uncertainty is evaluated as well as consequences for stock assessment outcomes.

Material and methods

Data

Summer AT surveys of walleye pollock in the EBS have been conducted by the Alaska Fisheries Science Center of the National Oceanographic and Atmospheric Administration (NOAA) since 1979. The present study used data collected over 5 years aboard the NOAA ships *Miller Freeman* (2006) and *Oscar Dyson* (2007–2010). The survey's primary objective was to collect acoustic and trawl information to estimate midwater walleye pollock abundance and distribution (Honkalehto *et al.*, 2012). Data from these surveys are used to estimate the parameters of a population dynamics model for EBS walleye pollock to determine total biomass, total abundance, abundance- and biomass-at-age (Ianelli *et al.*, 2011).

The survey design involves systematic parallel transects oriented in a generally north–south direction, and spaced 20 miles apart, with a random starting point. Sampling is conducted over a 2-month period in summer, and covers the EBS shelf. The southern end of each transect is either nearshore (Alaska Peninsula or Aleutian Islands) or beyond the shelf break in water 500–1000 m deep. The northern ends of transects are fixed based on historical pollock distributions, in water 50–85 m deep, but are extended if pollock are observed. Only pollock distributed in US waters are considered in this study. Calibrated acoustic backscatter data are recorded using a Simrad ER60 echosounder operating at 38 kHz. Nautical area backscattering coefficient (s_A, following the notation of MacLennan *et al.*, 2002) visually attributed to walleye pollock is calculated for each 0.5 nautical mile length segment and averaged over 2.5 nautical miles to reduce the skewness and the percentage of zeros in the acoustic backscatter variable (Figure 1a). The following evaluation of uncertainty (e.g. estimation variance) is only relevant at the physical area associated to the individual acoustic data values. Then, regular trawl haul samples are used to identify acoustic targets in addition to biological samples from which length and age data are collected (Figure 1b). These methods follow procedures detailed generally in Simmonds and MacLennan (2005) and specifically in Honkalehto *et al.* (2012).

The acoustic backscattering from juvenile pollock was differentiated from that of adults in some years. The separation was based on 2D morphology and the depth of the pollock aggregations as seen on the echograms. Data for juveniles and adults were analysed separately here because the trawl hauls used to characterize the length frequency distribution of juveniles and adults differ, and because both groups can occur at the same location, but in different depth layers, with juvenile aggregations often found closer to the sea surface, and adult aggregations closer to the seabed.

Evaluating the various sources of uncertainty

A simulation framework was developed to evaluate the uncertainty associated with the trawl and acoustics data. The uncertainty arising from spatial sampling was evaluated for the acoustic backscatter and for the trawl length frequency using Gaussian geostatistical conditional (co-) simulations. Both variables were assumed independent, and simulated independently with Gaussian simulations, taking into account spatial correlation in the univariate case of the acoustic backscatter, and the spatial cross-correlation in the multivariate case of the length frequency. The uncertainties arising from errors in the functional relationship between target strength and length, age and length, and weight and length were evaluated for each year separately using bootstrapping, except for target strength-at-length, for which bootstrap sampling from the same relationship was used in all years. Thus, the age–length and weight-length relationships were assumed to change among years but not the target strength-at-length relationship. To help the understanding, a flow diagram of the simulation framework is presented in Figure 2, and the derivation of some

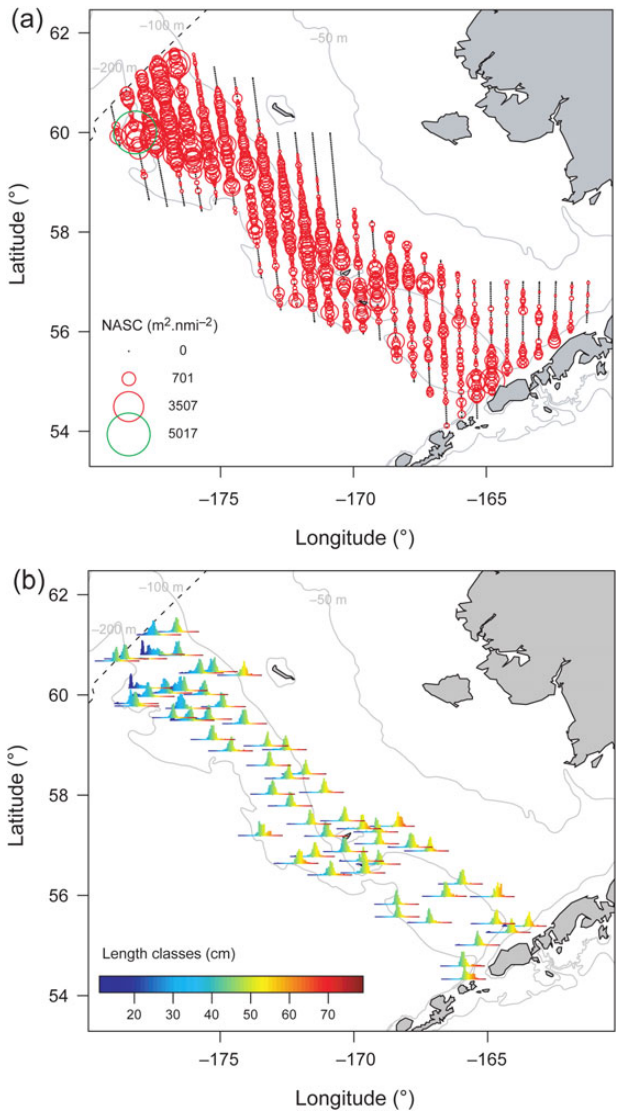

Figure 1. Spatial distribution of (a) acoustic backscatter (m^2·nautical $mile^{-2}$) attributed to walleye pollock and (b) walleye pollock length frequencies in summer 2006. Length classes range from 11 to 79 cm.

supporting equations is provided in Annexes. The various steps of the simulation procedure are detailed hereafter.

Geostatistical conditional simulation of the acoustic backscatter

An important step for this procedure is to simulate values of acoustic backscatter with realistic spatial variability (i.e. that inferred from the sampled locations) for all cells of a grid covering the whole surveyed region, conditional on the values observed at the sampled locations from the acoustic survey (Figure 2). To do so, transformed Gaussian geostatistical simulation with a Gibbs sampler to handle the zero values was used to create multiple realizations of acoustic backscatter (Woillez *et al.*, 2009).

In this approach, the idea is to find a hidden Gaussian random field $Y(x)$ (x denoting a point in two-dimensional), from which is a transformation of the original acoustic backscatter spatial random field $Z(x)$. This means that multiple realizations of $Z(x)$ can be obtained through back-transforming a set of realizations of $Y(x)$. However, many zeros values are observed in the acoustic backscatter, meaning that the $Z(x)$ process has a positive probability of producing a zero value (unlike a Gaussian distribution, which is continuous), and this must be taken into account in the simulation procedure.

Practically, the simulation procedure follows these steps. First, a normal score transformation of the acoustic backscatter data $Z(x_i)$, $i = 1, \ldots, n$ indexing the sample locations, is performed. However, because of the high proportion of zeros, the resulting variable is a lower-cut Gaussian $Y^+(x_i)$. This variable corresponds to a truncated Gaussian distribution but with an additional spike at y_c, where y_c is determined from the proportion of zeros. The inverse transformation between $Z(x)$ and $Y(x)$ (also named a Gaussian anamorphosis function) relates the acoustic backscatter spatial random field $Z(x)$ and the Gaussian random field $Y(x)$. The fitted model for $Y(x)$ corresponds to 0 for values below y_c, and to a simple piece-wise linear interpolation function for values above y_c (i.e. where acoustic values are not zero). Then, a Gibbs sampler was applied to $Y^+(x_i)$ to simulate the values for $Y(x)$ where the sampled value was less than y_c (Lantuéjoul, 2002; Woillez, 2007). This is an iterative process on the set of samples to be modified, where the Gaussian value at a given point is simulated conditionally to the Gaussian values at all other data points using the variogram model of $Y(x)$ (see Annex1). However, the experimental variogram of $Y(x)$ is not available, because of the zeros in the original variable $Z(x)$. Hence, the variogram model of $Y(x)$ must be fitted indirectly using the experimental variogram of the lower-cut Gaussian $Y^+(x_i)$. As developed in Annex 2, there an explicit link between the covariance function (i.e. the variogram) of $Y(x)$ and the covariance function of $Y^+(x)$. Thus, the idea is to find the variogram model parameters of $Y(x)$ so that the variogram model of $Y^+(x)$ fits the experimental variogram of $Y^+(x_i)$. Once the variogram model of $Y(x)$ is indirectly fitted, the Gibbs sampler is performed, and classical conditional simulation is performed using the values of $Y(x_i)$ for conditioning. This amounts to simulating Gaussian values at the unsampled cell locations over the whole surveyed area, so that it reflects the spatial variability (i.e. that inferred from the sampled locations) that must exist between all the unsampled cell locations and it honours the data values known at the sample locations (Chilès and Delfiner, 1999). The conditioning is straightforward in the Gaussian case and involves adding a spatially independent simulated kriging error to the kriging estimator (Chilès and Delfiner, 1999). This means that, conditionally to the data, the mean of the conditional simulations is equal to the kriging:

$$E[Y_{CS}(x)|Y(x_i), \quad i = 1, \ldots, n] = Y^K(x) \tag{1}$$

and the variance of the conditional simulations equals the variance of the non-conditional simulated kriging error, also equals the kriging variance:

$$\begin{aligned} &\mathrm{var}[Y_{CS}(x)|Y(x_i), \\ &\quad i = 1, \ldots, n] = \mathrm{var}(Y_{NCS}(x) - Y^K_{NCS}(x)) = \sigma^2_K(x). \end{aligned} \tag{2}$$

Multiple realizations of $Y(x)$ were produced and back-transformed using the anamorphosis model (see above) to obtain a set of simulations of the original acoustic spatial random field $Z(x)$.

Geostatistical conditional co-simulation of the length frequency distribution

The aim is to simulate length frequency distributions with realistic spatial variability for all cells of a grid covering the whole surveyed

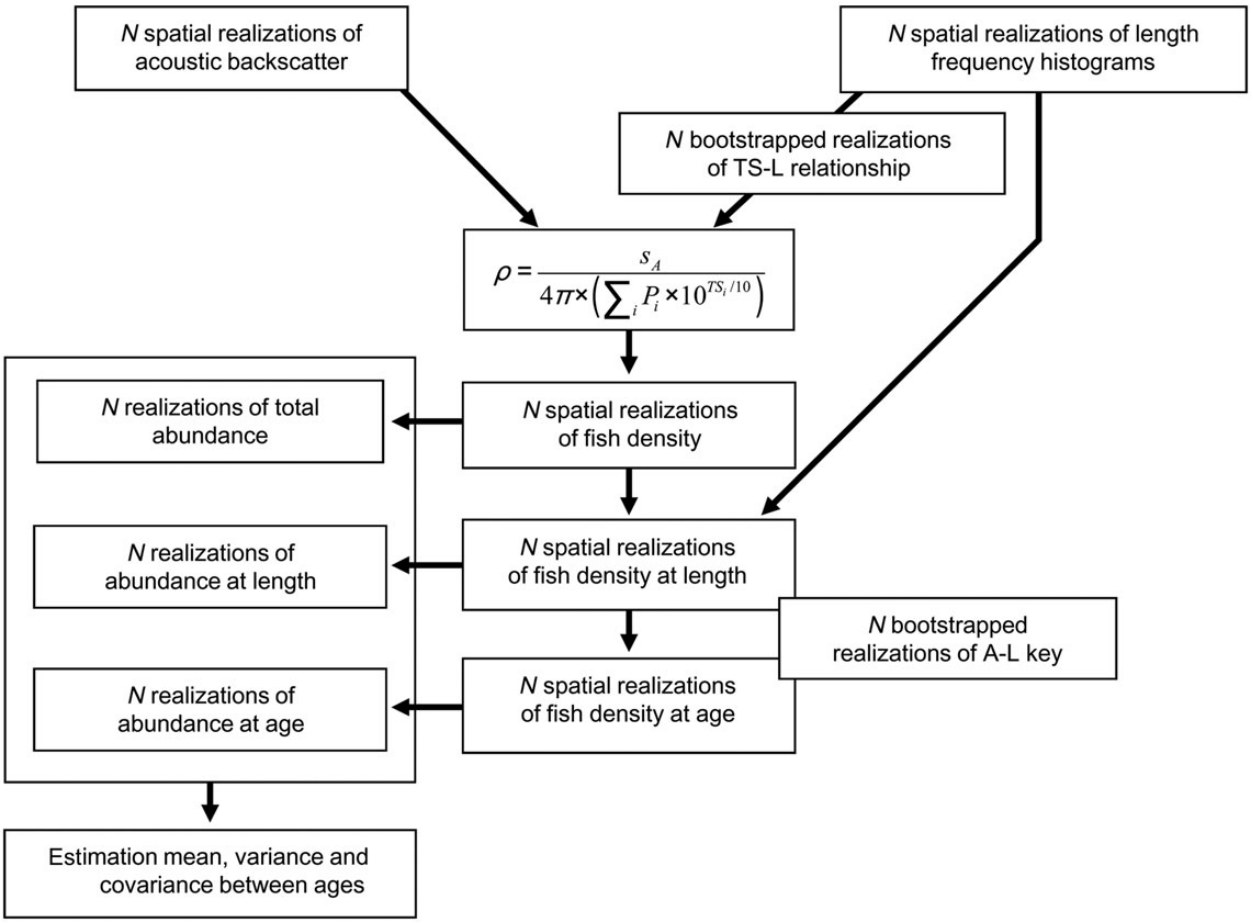

Figure 2. Flowchart showing how uncertainties are combined and propagated to final estimates.

region, conditional on the length frequency distributions observed at the sampled locations from the trawl hauls made during the survey (Figure 2).

The simulation approach considered the length data as discrete spatial random field variables. Quantiles from length frequency distributions were considered as co-regionalized variables monitored at some sampling locations (those of the trawl hauls). The spatial uncertainty associated with the values of the co-regionalization (the fact that values of the quantiles vary in space in between sample locations) can be assessed over the survey area using geostatistical co-simulation (Lantuéjoul, 2002; Emery, 2008), which provides realizations that reproduce the spatial variability of each variable as well as the relationships among the variables. The Gaussian random fields model, whose finite-dimensional distributions are multivariate normal, is especially well suited for conditional co-simulation (Chilès and Lantuéjoul, 2005), i.e. geostatistical co-simulation that honours multivariate data values known at data locations. Indeed, conditioning a simulation by adding a spatially-independent simulated kriging error to the kriging estimator holds in the multivariate case, except that co-kriging must be used instead of kriging (Emery, 2008).

However, most of the time, the Gaussian assumption is only acceptable after the original variables are transformed (cf. Gaussian anamorphosis). Therefore, conditional co-simulation of the length quantiles is made on transformed Gaussian variables, which are then back-transformed to obtain co-simulation of the original variables.

A structural model needs to be inferred from the multivariate data to perform the co-simulation. A common structural tool for multivariate analysis is the cross-variogram:

$$\gamma_{ij}(h) = \frac{1}{2} E\{[Y_i(x+h) - Y_i(x)][Y_j(x+h) - Y_j(x)]\} \quad (3)$$

which describes the spatial interaction for a given distance h between two transformed length quantiles i and j. Here a linear model of co-regionalization was assumed (Wackernagel, 1995; Chilès and Delfiner, 1999), meaning that all regionalized variables are thought to be generated by a set of biological processes acting additively at various spatial scales. Consequently, all simple- and cross-variograms $\gamma_{ij}(h)$ are linear combinations of the same basic structural components $\gamma^k(h)$ of unit variance:

$$\gamma_{ij}(h) = \sum_k b_{ij}^k \gamma^k(h), \quad (4)$$

where b_{ij}^k is the i–j component of the variance–covariance matrix. Each matrix b^k must be positive definite, implying the following inequality:

$$b_{ij}^k \leq \sqrt{b_{ii}^k b_{jj}^k}. \quad (5)$$

Thus, the basic structures comprising the cross-variograms are those already in the simple variograms. The basic structures are first identified and subsequently used to model the cross-variograms. Fitting of the (cross-) variogram models is conducted using the procedure developed by Goulard and Voltz (1992) that allows the parameters of the matrix model of co-regionalization b^k to be estimated using an iterative least-squares like technique.

At each trawl location, the empirical quantile function, i.e. the inverse cumulative distribution function, of the sampled length frequency distribution is modelled by a piece-wise linear interpolation to estimate the quantiles of the length for the probabilities $\{0.025, 0.05, 0.1, 0.15, 0.2, 0.3, \ldots, 0.9\}$. The quantiles of the length for probabilities 0 and 1 are set, respectively, to the minimum and the maximum length values recorded at each trawl station. The length quantiles are transformed by Gaussian anamorphosis as they are far from Gaussian. A linear model of co-regionalization is used to fit all simple- and cross-variograms of the transformed length quantiles then the inferred model is used in a multivariate Gaussian conditional simulation. Realizations of transformed length quantiles are produced and back transformed to realizations of raw length quantiles using appropriate anamorphosis models. At each grid node, simulated length distributions are built from the simulated length quantiles, and the order among length quantiles honoured. The reconstruction has two steps: first, the order among simulated length quantiles is checked, then the length classes are predicted using a model (i.e. a piece-wise linear interpolation function) from the simulated length quantiles. The check is performed first in the ascending order of simulated length quantiles; if the simulated length quantile of the higher order is below the current simulated length quantile, then the simulated length quantile of the higher order is set to the current simulated length quantile. Then, the check is performed in the descending order. What is kept for the prediction step of the length frequency distribution reconstruction, is the average value between the two checking process.

Fitting the linear model of co-regionalization and (co-)simulating the acoustic backscatter and the length quantiles were performed using the R library RGeostats (Renard *et al.*, 2010).

Bootstrapping target strength-at-length and weight-at-length relationships

Regressions relating target strength-at-length and weight-at-length are used when variables are combined to estimate abundance or biomass (Figure 2). The first relationship quantifies the acoustic reflectivity of a fish of length L, measured in terms of target strength (TS) as follows:

$$TS = b + m\log(L), \tag{6}$$

where b and m are species-specific coefficients, assumed to be known from experimental evidence. For walleye pollock b has been empirically estimated as -66 and m is set to 20 (Foote and Traynor, 1988; Traynor, 1996). The second relationship relates weight to length according to

$$W = cL^d \tag{7}$$

with c and d the regression model parameters.

The uncertainty arising from these two regressions is evaluated separately by bootstrapping respectively the data from Traynor (1996) and weight and length data collected in a given survey year. Resampling residuals are the approach considered here. First, a regression model is fitted to the data, say (x_i, y_i), $i = \{1, \ldots n\}$, where x_i is the ith independent variable, y_i is the ith the dependent variable, $\hat{y}_i$ is the ith fitted value and $\hat{e}_i = y_i - \hat{y}_i$ is the ith residual. A randomly resampled residual $\hat{e}_j$ is added to the response variable $\hat{y}_i$ for each pair, (x_i, y_i), i.e. a synthetic response variable $y_i^* = \hat{y}_i + \hat{e}_j$ is created where j is selected randomly from $(1, \ldots n)$ for every i. The regression model is then refitted using the new response variable y_i^*, and the quantities of interest retained, in our case, the parameter b for the *TS-L* relationship (m was set equal to 20) and the parameters c and d for the *W-L* relationship. For the latter, the process preserved the correlation between the pairs of parameters c and d. Finally, the steps of creating a synthetic response variable and of refitting is repeated many times to construct a probability density function for the quantities of interest.

Bootstrapping age–length keys

Age–length keys were bootstrapped to propagate the uncertainty of the relationship used to convert estimates of abundance-at-length or biomass-at-length into estimates of abundance-at-age or biomass-at-age (Figure 2). Following the current method from the Alaska Fisheries Science Center (AFSC), age–length keys were stratified east and west of 170° as pollock have been observed historically to grow at different rates and to have different age and length compositions in these areas (Honkalehto *et al.*, 2002), and the bootstrap was conditioned on this stratification. Thus, for each surveyed year and geographical stratum, a set of realizations of age–length keys was generated by bootstrapping based on the observed ages-at-length in the survey. The simulation design replicates the survey sampling design (i.e. sample hauls and observed ages (within length classes) within a haul). All sampling was conducted with replacement. Practically, the bootstrap samples were constructed by sampling N hauls with replacement from the N sampled tows within each stratum. Then, within each haul, age data were sampled with replacement (the sample size being equal to the dimension of the age data present in the bootstrapped sampled hauls). All sampled ages were pooled for each stratum and an age–length key computed. The age-frequency for length classes that are not sampled are set based on a long-term historical age–length key. However, the proportion of animals that come from the historical age–length key is very small. For year 2006, this represents only 0.4 and 0.9% in average for east and west of 170°W, respectively.

Combining uncertainties for abundance and biomass estimates

Estimates of total abundance, abundance-at-age, total biomass, and biomass-at-age and their associated uncertainty are obtained by combining spatial realizations of acoustic backscatter, spatial realizations of the length frequency distribution, and bootstrapped regression results on target strength-at-length (*TS-L*) and weight-at-length (*W-L*), and bootstrapped age–length keys (Figure 2). At each simulation grid node, simulated acoustic backscatter (nautical area scattering coefficient, s_A in m^2·nautical $mile^{-2}$) are converted to fish density (in nb·nautical $mile^{-2}$) using simulated spherical-weighted mean backscattering cross section (in m^2):

$$\rho = s_A(4\pi \times \bar{\sigma}_{bs})^{-1}, \tag{8}$$

where the weighted mean backscattering cross section $\bar{\sigma}_{bs}$ is the average backscattering cross section taken over the simulated size

distribution of the insonified fish:

$$\bar{\sigma}_{bs} = \sum_i P_i \times \sigma_{bsi}, \tag{9}$$

where P_i is the simulated proportion of fish of length L_i. The backscattering cross section relates to the target strength of one fish TS_i (in dB re 1 m^2), which can be expressed as a function of its length L_i (in cm):

$$\sigma_{bsi} = 10^{TS_i/10}, \tag{10}$$

where $TS_i = b + m\log(L_i)$, b is taken from bootstrapped samples and m is set to 20.

For each grid node the simulated abundance-at-length AL_i is the product of the simulated proportion at length P_i, with the simulated pollock density ρ and the cell area a (in nautical $mile^2$). The simulated biomass-at-length BL_i is obtained by multiplying the simulated abundance-at-length by the simulated weight-at-length, W_i from Eqn. 8.

The simulated abundance-at-age AA_j and the simulated biomass-at-age BA_j are then derived from the simulated abundance-at-length and the simulated biomass-at-length using $Q_{i,j}$ the bootstrapped proportion of fish in a given length class i which are of age j (i.e. an age–length key):

$$AA_j = \sum_i Q_{i,j} \times AL_i, \tag{11}$$

$$BA_j = \sum_i Q_{i,j} \times BL_i, \tag{12}$$

where the bootstrapped age–length key depends on the stratum in which the grid is located. Simulated total abundance A and simulated total biomass B are obtained by summing over ages:

$$A = \sum_j AA_j, \tag{13}$$

$$B = \sum_j BA_j, \tag{14}$$

Simulated estimates over grid cells are summed over the entire survey domain providing estimates of total abundance and biomass. Statistics (sample mean and standard deviation) can be derived from simulation estimates, as well as associated upper and lower 95% confidence intervals.

Decomposing the total uncertainty

The aim is to assess the contribution to uncertainty from the five sources (the acoustic data, the length frequency distributions, the age–length keys, and the weight-at-length and the target strength-at-length relationships) to the whole. Because the individual uncertainties do not add up to the total uncertainty (they combine non-linearly, and they can actually contribute to each other in a hierarchical sense as well as being correlated with each other), the contribution from the kth source was defined as:

$$\sigma_k^2 = \sigma_{\text{total}}^2 - \sigma_{\text{total}-k}^2, \tag{15}$$

where $\sigma_{\text{total}-k}^2$ is the total uncertainty evaluated when the kth source is held fixed (Løland *et al.*, 2007). Negative values can thus appear if $\sigma_{\text{total}-k}^2 > \sigma_{\text{total}}^2$. The relative contribution from the kth source is then given by

$$\sigma_k^2 \left(\sum_{k=1}^{5} \sigma_k^2 \right)^{-1}. \tag{16}$$

Results

Geostatistical simulation of the acoustic backscatter

Over the period 2006–2010, the acoustic backscatter was characterized by a high proportion of zeros (between 25 and 51% for adults, and between 72 and 91% for juveniles) and a skewed histogram with extreme values (maximum between 2787 and 15 737 m^2·nautical $mile^{-2}$ for adults, and between 6322 and 26 697 m^2·nautical $mile^{-2}$ for the juveniles; Table 1). For the example year 2006, the experimental variogram of the lower-cut Gaussian variable Y^+ (Figure 3) was fitted indirectly using a nested structure with a nugget of 0.10, an exponential component of 0.60 with a range of 17.5 nautical miles, and an anisotropic spherical component of 0.30 with a range of 300 nautical miles, 135° for the direction of the main axis of the anisotropy ellipse, and an anisotropy ratio of 0.25. For the other years, the variogram models of the Gaussian Y presented either two (for juveniles), or three (for adults) structures (Table 2). The first structure was always a nugget. Juveniles always had a short-range structure, whereas adults had short- and long-range structures. The short-and long-range structures were modelled either by a spherical or an exponential component, which complicated comparing spatial ranges among years. The long-range structure for adults always exhibited anisotropy with the main direction 135° (along the EBS shelf) and an anisotropy ratio between 0.25 and 0.45 (Table 2).

Table 1. Descriptive statistics (mean, standard deviation (s.d.), skewness, maximum (max.), percentage of zeros (% zeros), and number of samples (N)) of the raw and simulated acoustic backscatter (m^2.nautical $mile^{-2}$) for the two walleye pollock aggregation types (adults or juveniles).

		Mean		s.d.		Skewness		Max.		% zeros		N	
Year	Aggregation type	Data	Simu	Data	Simu	Data	Simu	Data	Simu	Data	Simu	Data	Simu
2006	Adults	234	236	396	394	4	4	5017	6093	25	25	1653	11 753
2007	Adults	349	357	622	613	4	3	5637	5734	33	34	1314	9207
2007	Juveniles	150	151	1131	1112	17	18	23 094	33 846	91	91	524	3568
2008	Adults	150	146	259	238	4	3	2787	2836	39	40	1327	9422
2008	Juveniles	317	384	1673	2022	11	11	26 697	41 937	77	78	449	3113
2009	Adults	226	221	799	756	8	7	11 166	11 545	51	52	942	6828
2009	Juveniles	273	295	1017	1076	5	5	9096	12 566	80	80	354	2381
2010	Adults	389	395	1072	1004	7	6	15 737	17 005	48	49	1412	10 128
2010	Juveniles	175	198	536	628	6	7	6322	10 778	72	73	418	3019

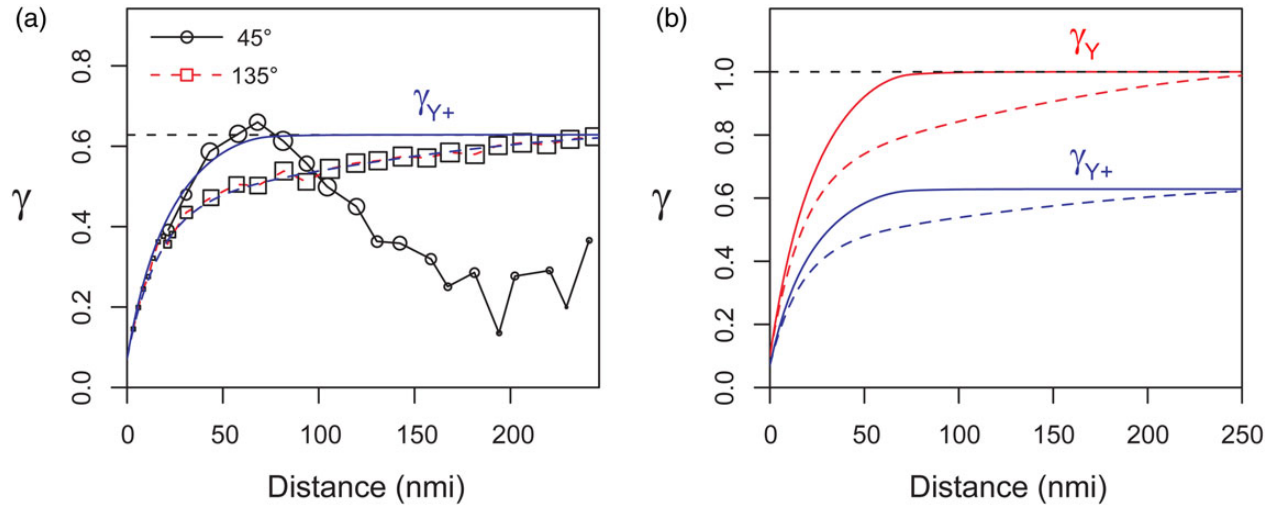

Figure 3. (a) Experimental variogram of the lower-cut Gaussian Y^+, with symbol size proportional to the number of data points, computed in two directions, 45° (circles, across shelf) and 135° (squares, alongshelf), and (b) the anisotropic variogram model of the Gaussian Y (solid, 45°; dotted, 135°) and the corresponding variogram for the lower-cut Gaussian Y^+. This figure is available in black and white in print and in colour at *ICES Journal of Marine Science* online.

Table 2. Models fit to experimental variograms of normal score transformed acoustic backscatter from EBS surveys 2006 to 2010.

			Structure 2			Structure 3				
Year	Aggregation type	Structure 1 Nugget	Type	Partial sill	Range	Type	Partial sill	Range	Direction	Ratio
2006	Adults	0.10	exp	0.60	17.5	sph	0.300	300	135	0.25
2007	Adults	0.10	exp	0.47	10.0	sph	0.430	280	135	0.35
	Juveniles	0.10	sph	0.90	30.0	–	–	–	–	–
2008	Adults	0.10	exp	0.40	8.0	exp	0.600	80	135	0.35
	Juveniles	0.20	sph	1.10	32.0	–	–	–	–	–
2009	Adults	0.10	sph	0.50	30.0	sph	0.750	80	135	0.45
	Juveniles	0.20	exp	1.10	14.0	–	–	–	–	–
2010	Adults	0.05	exp	0.55	12.0	sph	0.325	175	135	0.35
	Juveniles	0.10	exp	1.20	15.0	–	–	–	–	–

One hundred realizations of acoustic backscatter were produced for each year over the period 2006–2010. For year 2006, realizations with low, medium, and high mean levels over the domain illustrated the variability of the simulated spatial distribution of acoustic backscatter (Figure 4a–c). High values of simulated acoustic backscatter were found mostly in the northwest of the survey domain in 2006, while intermediate values were found over the whole EBS shelf except in the southeast away from the Aleutian Islands (Figure 4a–c). Descriptive statistics from simulated and raw acoustic backscatter were computed (Table 1). The conditional simulations of acoustic backscatter were able to reproduce the summary statistics of the raw data, except for the maxima (Table 1). Maxima were deliberately unbounded in the modelling of the Gaussian anamorphosis, resulting in simulated values higher than in the raw data.

Geostatistical co-simulation of the length frequency distribution

The quantile function of each sampled length frequency distribution (i.e. each trawl) was modelled using piece-wise linear interpolation and 14 length quantiles were estimated for the chosen probabilities (Figure 5). The length quantiles were then transformed into normally distributed variables. The experimental (simple and cross) variograms for the normal score length quantiles were computed at lags between 25 and 35 nautical mile, depending on the sampling lag of the year (e.g. 35 nautical mile for 2006). No anisotropy was detected, so the experimental variograms were omnidirectional. A linear model of co-regionalization was fitted for each year. For instance, in 2006, a nugget effect and two nested isotropic structures, the first one spherical with a range of 50 nautical mile and the second one linear, were fitted for the linear model of co-regionalization of the transformed length quantiles using the algorithm of Goulard and Voltz (1992)(Table 3). The resulting fit is illustrated in Figure 6. A third model component was needed to fit the trend present in the transformed length data. Since the trawl locations had good coverage over the surveyed region and especially over the fish distribution, there was no unwanted extrapolation of the model. For the other years, the fitted models were very similar, showing some consistency over time in the co-regionalization of the length quantiles (Table 3).

Conditional co-simulation of the length quantiles was performed on the same simulation grid as that used for the acoustic backscatter. One hundred realizations of transformed length quantiles were produced and back transformed to realizations of raw length quantiles using appropriate anamorphosis models. Each

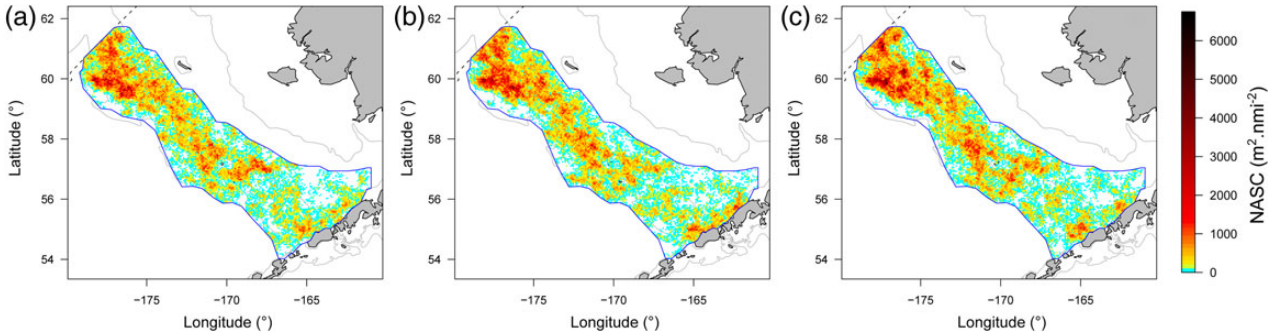

Figure 4. Realizations of the spatial distribution of acoustic backscatter (m^2.nautical $mile^{-2}$) with (a) low, (b) medium, and (c) and high mean levels over the domain. Only simulated values above zero have been coloured. This figure is available in black and white in print and in colour at *ICES Journal of Marine Science* online.

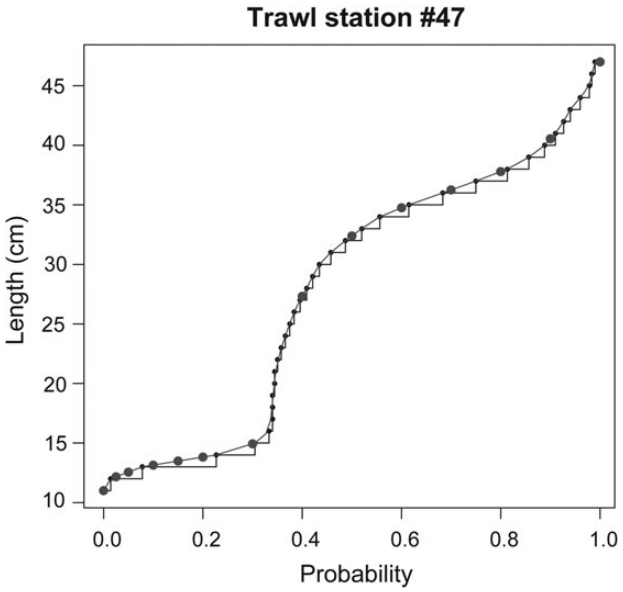

Figure 5. Inverse empirical cumulative distribution function of length for trawl station #47 in 2006. The quantile function was piece-wise interpolated (line) and length quantiles (dots) estimated for probabilities {0, 0.025, 0.05, 0.1, 0.15, 0.2, 0.3,. . ., 1}.

Table 3. Variogram models fitted for the linear model of co-regionalization of the transformed length quantiles using the algorithm of Goulard and Voltz (1992).

			Structure 2		
Year	Aggregation type	Structure 1	Type	Range	Structure 3 Type
2006	Adults	Nugget	Spherical	50	Linear
2007	Adults	Nugget	Spherical	50	Linear
	Juveniles	Nugget	Spherical	50	Linear
2008	Adults	Nugget	Spherical	50	Linear
	Juveniles	Nugget	Spherical	50	Linear
2009	Adults	Nugget	Spherical	25	Linear
	Juveniles	Nugget	Spherical	30	Linear
2010	Adults	Nugget	Spherical	50	Linear
	Juveniles	Nugget	Spherical	50	Linear

Gaussian anamorphosis was modelled using a piece-wise linear interpolation, similar to that for the acoustic backscatter, with bounds fixed to the maximum and the minimum values observed in the survey data of the length quantile under consideration. Finally, simulated length frequency distributions were rebuilt at each grid node from the simulated length quantile values, and multiple realizations of length frequency distributions were produced. Maps detailing the probability of finding small, medium or large pollock were produced for three different realizations to illustrate the spatially simulated length frequency distributions (Figure 7). Probabilities of finding small fish were higher in the northwest of the survey domain. Probability of finding medium fish were high inside and near the previous area, while probabilities of finding large fish were high in the rest of the survey area (Figure 7).

Bootstrapping

Residuals from the regression between target strength and length were resampled 100 times, providing 100 bootstrapped estimates for the parameter *b*. The distribution of the bootstrapped estimates was fairly symmetric and close to normal, despite the low number of points in the regression. The resulting relationship is illustrated with the 95% confidence intervals derived empirically using the bootstrap estimates (Figure 8).

Similarly, residuals from the regression between weight and length were resampled 100 times for each year, providing 100 bootstrapped estimates for the parameters *c* and *d*. The distributions of the bootstrapped estimates were fairly symmetric and close to normal. Bootstrapping allowed the uncertainty associated with the regression to be evaluated, as illustrated by the bootstrapped 95% confidence intervals in Figure 9.

One hundred bootstrapped age–length keys were produced for each stratum and year. The mean age–length key and its variance were computed for year 2006 (Figure 10). Variability was higher for older fish in both strata, for intermediate fish in the stratum east of 170°W, and for younger fish in stratum west of 170°W.

Combining uncertainties for abundance and biomass

One hundred realizations of the acoustic backscatter, length frequency distribution, target strength-at-length relationship, age–length key, and weight-at-length relationship were produced for years 2006–2010 either by geostatistical (co)-simulation or bootstrap. The realizations were combined to produce estimates of abundance- and biomass-at-length, abundance- and biomass-at-age, total abundance and biomass. Uncertainties associated with the abundance and biomass estimates are illustrated for year 2006 in Figure 11. Results from the other years were similar, but are not described in detail. These simulation results compared well with AFSC survey estimates. The AFSC survey estimates lie within the empirically derived 95%

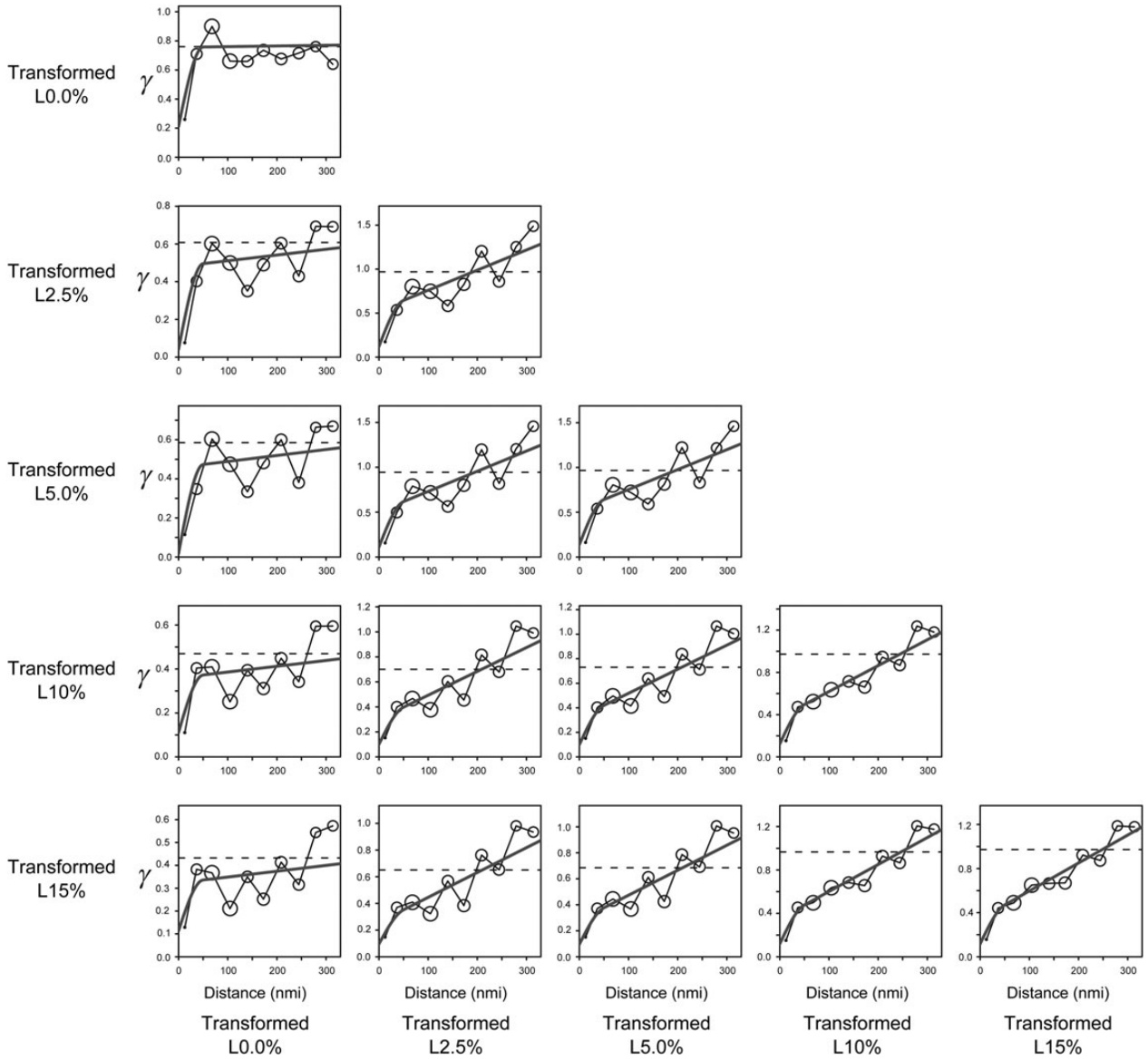

Figure 6. Experimental (black plain lines with circles sized in proportion to number of data points) and modelled (grey solid lines) simple and cross-variograms for normal scores of the first five length quantiles (L0%, L2.5%, L5%, L10%, and L15%) for year 2006.

confidence intervals of the simulated outcomes 100% of the times for total abundance or biomass, 100% of the times for abundance or biomass-at-age, 70% of the times for abundance-at-length and 58% of the times for biomass-at-length in 2006. Thus, some discrepancies exist for estimates at length. For instance, troughs (e.g. ∼18 cm) and peaks (e.g. ∼23 cm) in the survey abundance-at-length estimates are sometimes not well reproduced by the simulation estimates. In addition, simulation mean estimates were lower for small length classes (12–15 cm) and higher for medium length classes (43–51 cm). These discrepancies are explained by the methods used. While the AFSC stratum approach conserves the fluctuations seen in the sampled length frequencies, the geostatistical co-simulation tends to create more gradual transitions between sampled length frequencies at different locations, which results in dampened abundance-at-length variability over the survey domain where abrupt peaks and troughs occur in the survey data.

Otherwise, over the period 2006–2010, the coefficients of variation (CVs) of the abundances-at-age (the standard deviations of the simulated abundances divided by their means) ranged from 8 to 84%. High CVs were found for young ages (ages 1, 2, and to a lesser extent age 3). These young ages also had the highest variability over time. Otherwise, CVs were mostly between 12 and 18%, on average. The CVs of the total abundance were relatively high, ∼24% on average, but also quite variable (ranging from 8 to 51%). Young ages contribute substantially to uncertainty of total abundance; high CVs in total abundance are found for years where the CVs for the young ages are high (year 2007, 2009, and to a lesser extent 2008; Table 4).

Decomposing the total uncertainty

The relative contribution of each source of uncertainty was evaluated independently for total abundance, biomass, abundance-at-age, and

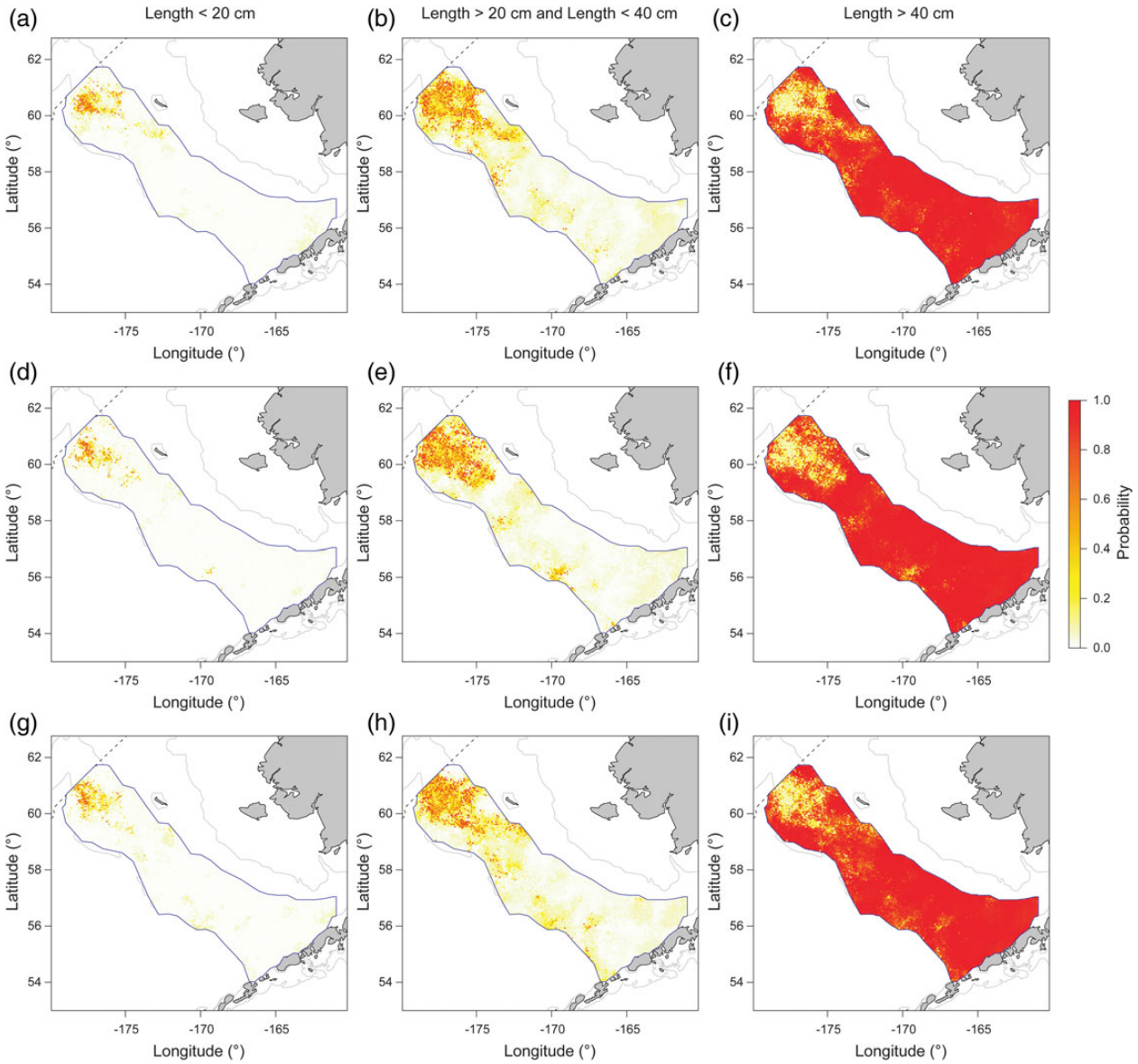

Figure 7. Three realizations of a spatial length frequency distribution with a low (a, b, and c), medium (d, e, and f), and high mean length (g, h, and i) mapped as the proportion of walleye pollock <20 cm (a, d, and g), >20 cm and <40 cm (b, e, and h), and >40 cm (c, f, and i). This figure is available in black and white in print and in colour at *ICES Journal of Marine Science* online.

biomass-at-age for each survey (Table 5). In 50 of 55 cases, the sources that contribute the most to the total uncertainty or uncertainty-at-age are the same for abundance and biomass. All sources except the weight-at-length relationship were the major source of uncertainty at least once. In contrast, the contribution of the weight-at-length relationship to total uncertainty was negligible (always close to zero). The contribution of age–length keys to total uncertainty was also negligible, but only for the total abundance or total biomass estimates. Some relative contributions appeared negative because of the variation overlaps between all the variables. For the example, in 2006, the length frequency distribution was the major source of uncertainty for ages 1–3, the target strength-at-length relationship was the largest contributor for ages 4–7, and the age–length key dominated for ages 8–10. For the total abundance, the target strength-at-length relationship contributed the most to uncertainty, while acoustic backscatter and length were comparable (24 and 30%, respectively). The target strength-at-length relationship contributed the most to uncertainty of total abundance and biomass. For each estimate of 2006, the major sources of uncertainty indicate retrospectively where more sampling effort should have been spent to reduce uncertainty. However, general recommendations are hard to draw as the major sources of uncertainty change over the time. Finally, the relative contribution to total uncertainty of the target strength-at-length relationship tended to be high (Table 5) when the overall uncertainty was low for years 2006 and 2010 (Table 4). This is because this uncertainty is fixed over years.

Structure of the estimation variance–covariance matrices

Variance–covariance matrices for abundance-at-age and biomass-at-age were very similar (not shown), and differed slightly from variance–covariance matrices of proportions-at-age, that were conveniently represented in terms of correlation (Figure 12).

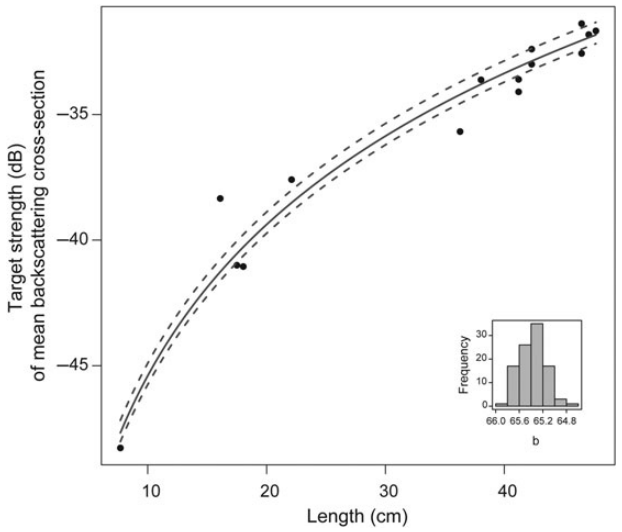

Figure 8. Bootstrapped 95% confidence intervals (dashed lines) generated by resampling residuals of the regression (solid line) relating walleye pollock target strength-at-length (data represented as points), with inset histogram of the bootstrapped estimates of the regression parameter (b in equation 6).

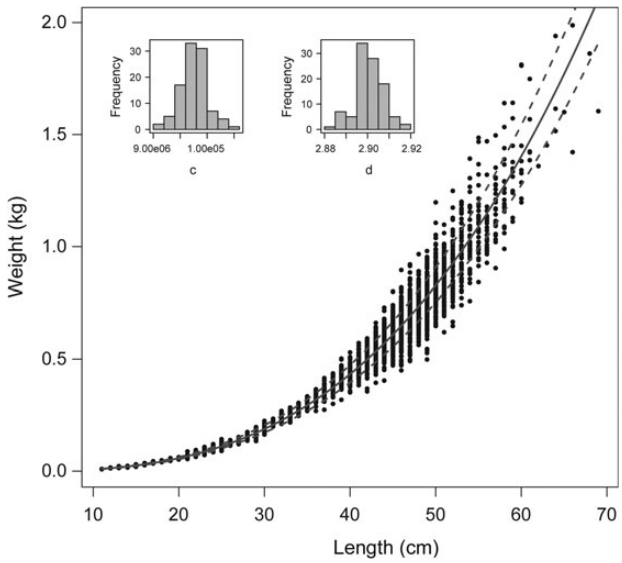

Figure 9. Bootstrapped 95% confidence intervals (dashed lines) generated by resampling residuals of the regression (solid line) relating walleye pollock weight-at-length (data represented as points), with inset histograms of the bootstrapped estimates of the regression parameters (c and d in equation 7).

Correlation between ages is generally positive, and the proportion-at-age for groups of ages is highly correlated with each other. For the example year 2006, group of ages were positively correlated together (ages 1–3, ages 3 and 4, and ages 5 and higher), but negatively correlated between groups. Otherwise, the years 2007 and 2009 differed greatly from other years. In these 2 years, there was very strong positive correlation between proportion at age estimates for all ages other than age-1. The reason was that age-1 year-class was dominant in these years, as a result the proportions of other ages were all low and correlated.

Discussion

The simulation framework developed here allowed the evaluation of the total uncertainty associated with abundance- and biomass-at-age estimates for EBS walleye pollock using geostatistical co-simulation and bootstrap techniques. The simulation approach used for the length frequency distribution, i.e. the co-simulation of the length quantiles, is original. Previous simulation approaches used mean length or a root mean square length (Gimona and Fernandes, 2003; Walline, 2007; Woillez *et al.*, 2009), or resampled length frequencies using nearest neighbour techniques (Zwolinski *et al.*, 2009). Other approaches used for estimation rather than simulation are co-kriging of polynomial coefficients that model the length frequency distributions (Petitgas *et al.*, 2011), and aggregation of length frequency distributions by strata based on similarity (Simmonds and MacLennan, 2005) or through expert opinion (Honkalehto *et al.*, 2012). Among these techniques, a co-simulation approach was tested, which was based on polynomial coefficients that were used to model the length-frequency distributions. However, the quantiles approach was preferred, because summary statistics over the simulated field were not well reproduced for the polynomial approach; major peaks were underestimated and the relative proportion of the largest fish was overestimated (not shown).

Discrepancies were observed between survey and simulation estimates of abundance and biomass (Figure 11). They might be explained by methodological differences. That is, the survey method grouped length frequency distribution data together by stratum and associated these data with corresponding acoustic backscatter, while the simulation method created smoother transitions in the spatial distribution of length frequency distributions. In addition, the length frequency distributions were modelled from a reduced number of quantiles and not from all the length classes, which might explain why some troughs and peaks were not well reproduced. For example, for the first length quantiles (e.g. 0.025 quantile), if small values at given sample location were surrounded by larger values from other sample locations, it is likely that the simulation would limit the influence of this small value by more than what occurs with the current AFSC method. An alternative explanation would come from the (relative) continuity of the transformed length quantiles (i.e. small nugget effect), meaning that the regionalized variables have a very smooth shared behaviour. Test results comparing the assessment model using the standard AFSC values with the simulation estimates were nearly identical, indicating that the differences were minor. This exercise simply used different point estimates for abundance-at-ages (using the two methods), but applied the same covariance matrix as derived from the simulation framework. Generally, the lower estimates of age-1 recruits in some years and higher CVs found with our simulation framework demonstrates how accounting for more aspects of the sampling process (and uncertainty) can affect the assessment model estimates. These results are consistent with the higher uncertainty found for age-1 pollock, which had a low or negative covariance structure with the others ages (see Table 4 and Figure 12).

The simulation approach could be improved in several ways. For years when adult and juvenile types of walleye pollock backscatter were identified, co-simulations of acoustic backscatter and length frequency distributions were first performed independently then combined. Such an approach more closely simulates the method currently used at the AFSC. However, there is no biological justification for treating these data types separately, and there may be drawbacks. For example, the sampling intensity is lower for juvenile

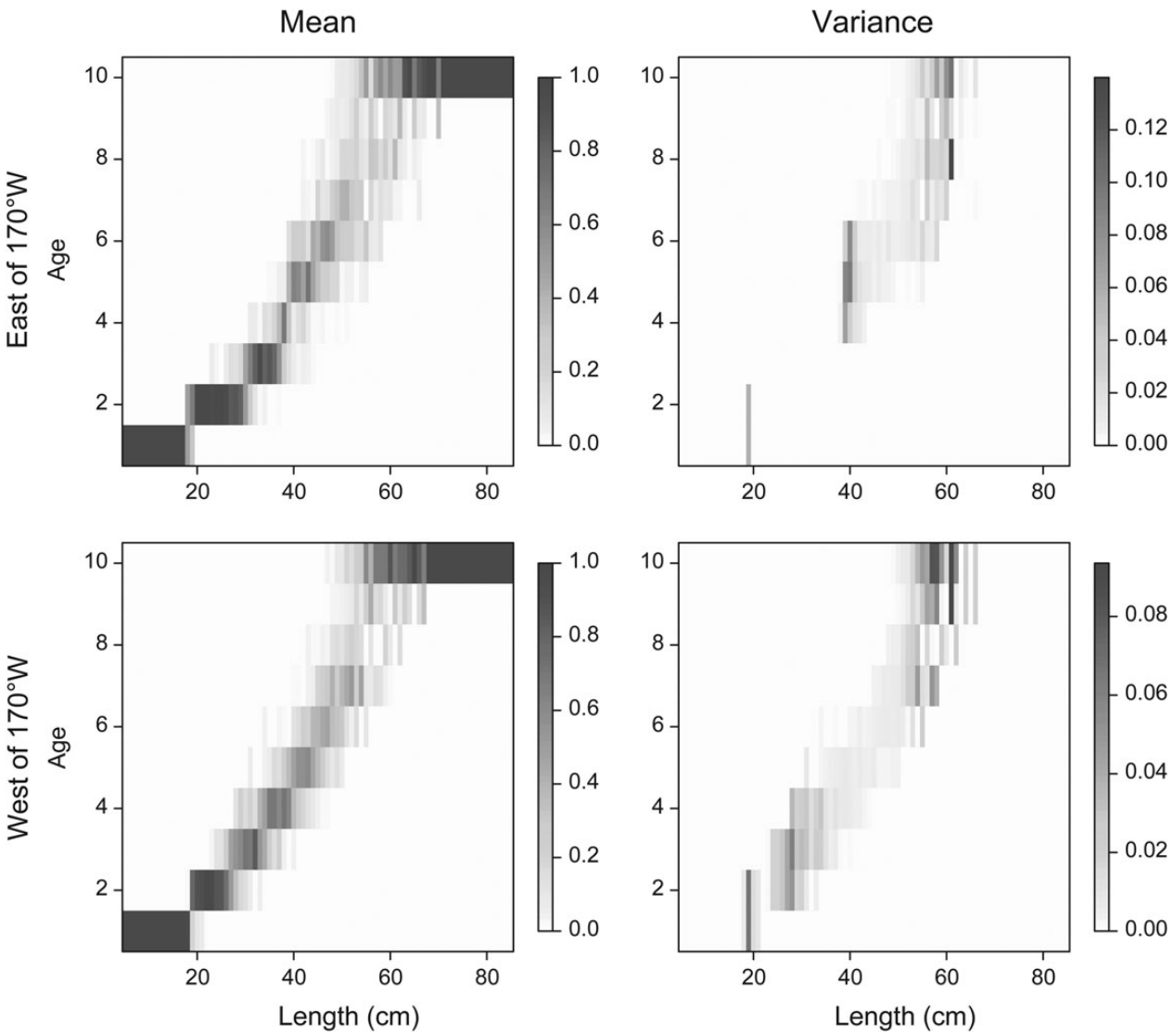

Figure 10. Mean (right) and variance (left) of 100 bootstrapped age–length keys for strata east (top) and west of 170°W (bottom).

pollock because fewer trawls targeted the associated layers. Consequently, fitting the variograms and covariograms is less reliable. The original choice was supported because of morphological and vertical positioning differences between aggregations (juvenile pollock aggregations are usually found higher in the water column in dense concentrations over a reduced area of the EBS shelf). This results in very different length frequency distributions in neighbouring trawl locations, or in trawl hauls made at different depths in the same location. This problem could be overcome in the modelling by including the vertical dimension. Another potential source of bias in the AT survey is trawl size selectivity where juvenile pollock under 25 cm are underrepresented in net samples (Williams *et al.*, 2011). This is particularly problematic when juveniles are present with older fish. Such bias could have been compensated in the quantiles simulation framework. Another aspect that could have been investigated and included in the simulation procedure is the potential co-variability between the acoustic backscatter and length data. If present, the simulation procedure should take it into account. This would mean that the simulation of one variable would depend on the other one, implying a probable slight reduction of the total uncertainty. Furthermore, a multiyear modelling approach could be investigated to reduce uncertainty and use the knowledge of prior years for a given year as it is likely that there is some stability over years in the co-regionalization of the length quantiles. In terms of simulation technique, the Gibbs sampler appeared to be an appropriate method to correctly specify the Gaussian variable Y at all data points, so that its histogram is totally Gaussian (without the spike corresponding to the zeros) and its spatial structure respects the inferred Gaussian model of Y. However, the convergence rate towards this equilibrium is not well known and is under study (Galli and Gao, 2001; Lantuéjoul, 2002). Recently, Lantuéjoul and Desassis (2012) proposed a propagative version of the Gibbs sampler that no longer has this limitation. It could be used instead of the Gibbs sampler in our simulation procedure.

The main source of uncertainty was not always the acoustic backscatter in contrast to most pelagic species surveyed using acoustic and trawl surveys (e.g. Zwolinski *et al.*, 2009). Decomposing the total uncertainty into the relative contribution of each source showed that length or age–length keys matter too, especially for abundance- and biomass-at-age estimates. This is in agreement

with Løland *et al.*(2007). Uncertainty coming from the target strength-at-length relationship is assumed to be invariant between years. As a result, its relative contribution is variable over years, being more important when total uncertainty was low, and less important when total uncertainty was high. It should be noted that only a few data points were used to estimate the *TS-L* relation, making it quite uncertain. Additional experimental points would help to reduce uncertainty due to the *TS-L* relationship. Results from Table 5 are very useful in indicating where more sampling effort should have been spent retrospectively to make most gains in terms of reduced uncertainty. The simulation analysis could be repeated considering alternative sample sizes for each of the five quantities on which the analysis is based, to try and gain an understanding on the issue of optimizing sampling efficiency. However, such analysis was not the aim of the paper. Otherwise, additional sources of uncertainty (e.g. species identification, vessel avoidance, acoustic shadowing, depth-dependent target strength, and potential ageing errors) could be incorporated into this framework to provide more comprehensive variance estimates of the abundance.

A frequentist approach was used here to evaluate the uncertainty of the abundance and biomass estimates. A Bayesian approach is an alternative that could have been considered. Several attempts have already been made to analyse AT data using Bayesian method, such as Juntunen *et al.*(2012), and Boyd *et al.* (2015). Juntunen *et al.*(2012) constructed a Bayesian spatial model to estimate the abundance of multiple pelagic fish species. In this work, the logarithm of the mean of NASC was the variable to be simulated, but it was unclear how the zero values in the acoustic backscatter data were treated. Boyd *et al.*(2015) made Bayesian posterior predictions based on a truncated GRF. This was effective at reproducing the patchiness of the observed spatial distribution of anchoveta off Peru.

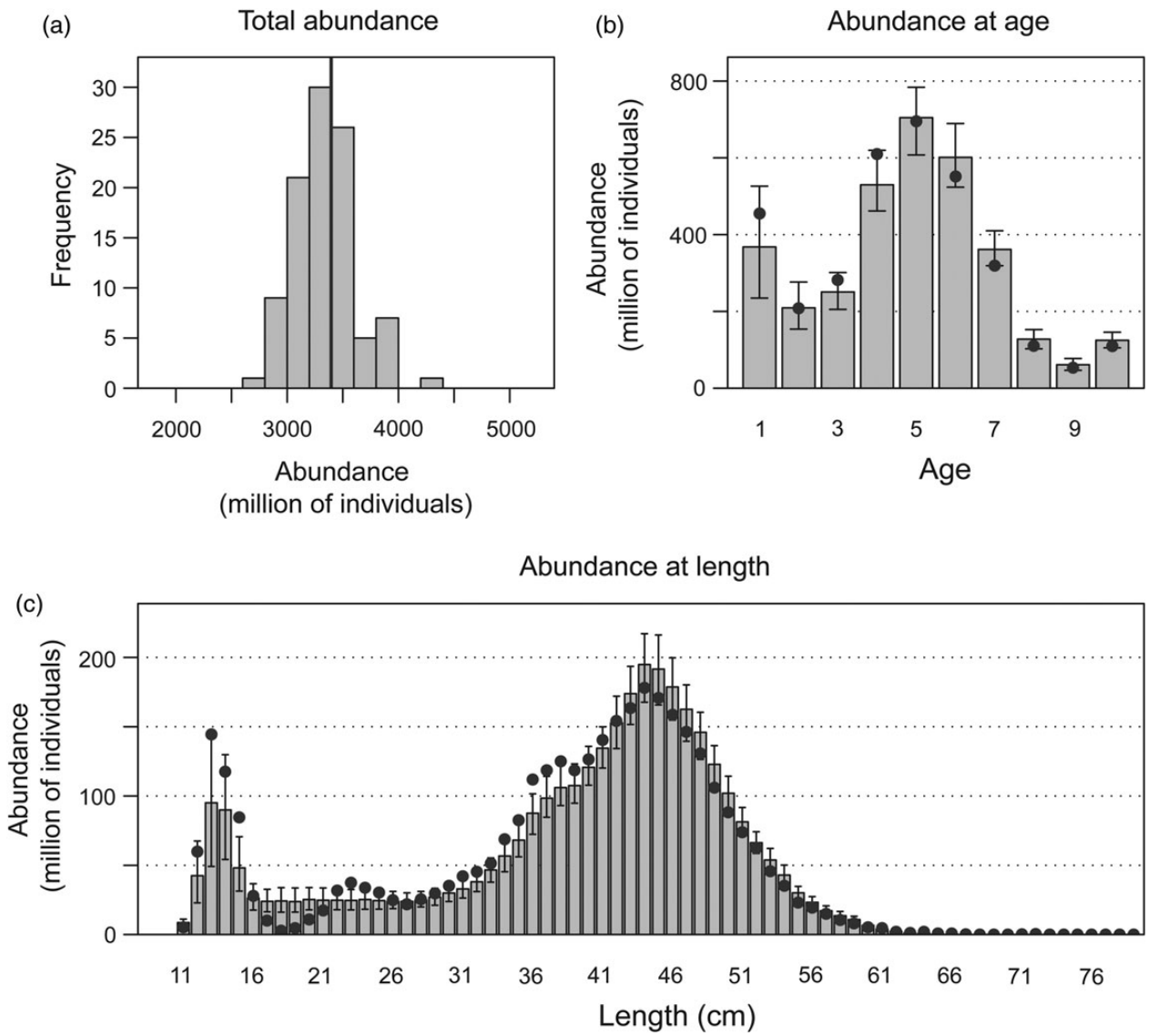

Figure 11. Distributions obtained from simulations for (a) total abundance, (b) abundance-at-age, (c) abundance-at-length, (d) total biomass, (e) biomass-at-age, and (f) biomass-at-length in 2006. Distributions of simulated total abundance and simulated total biomass are represented as histograms (a and d), while mean estimates for simulated abundance- and simulated biomass-at-age or -at-length are represented with bars where whiskers represent the associated 95% confidence intervals (b, c, e, and f). Grey lines or dots represent AFSC survey estimates in 2006.

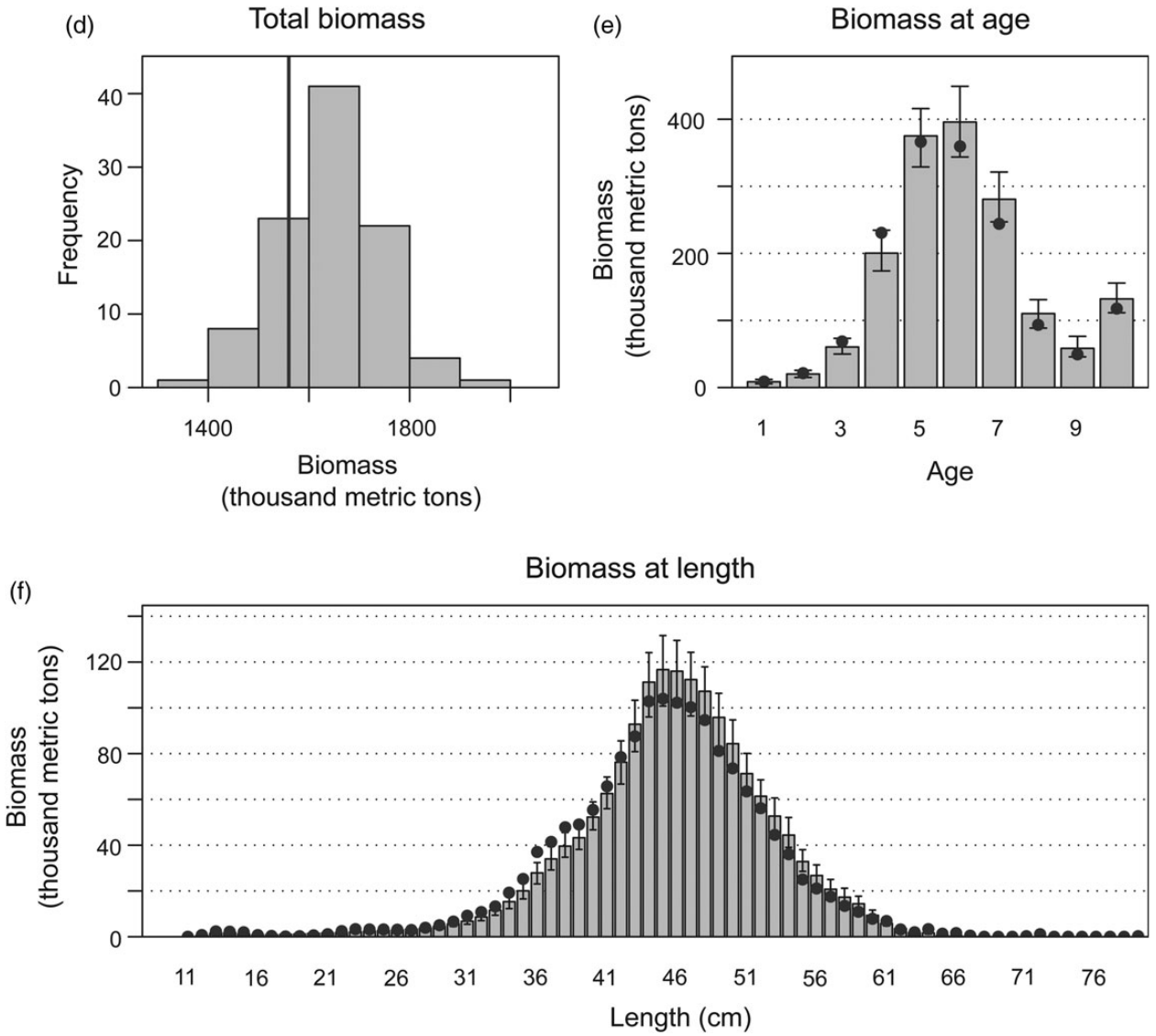

Figure 11 Continued

Table 4. CVs (%) of abundance-at-age, total abundance, biomass-at-age, and total biomass estimates obtained from simulations for years 2006 – 2010.

	Abundance							**Biomass**						
	CV							**CV**						
Age	**2006**	**2007**	**2008**	**2009**	**2010**	**Mean**	**s.d.**	**2006**	**2007**	**2008**	**2009**	**2010**	**Mean**	**s.d.**
1	28	76	38	84	13	48	31	26	56	44	68	12	41	22
2	18	56	23	37	10	29	18	18	51	23	29	10	26	16
3	12	21	21	44	10	22	14	12	17	19	42	9	20	13
4	10	11	15	23	10	14	6	9	9	14	19	10	12	4
5	8	8	11	23	11	12	6	8	8	11	21	11	12	5
6	8	8	9	19	16	12	5	8	8	9	18	17	12	5
7	8	9	8	18	28	14	9	8	9	8	17	27	14	8
8	11	9	10	17	27	15	8	11	9	11	17	27	15	7
9	16	14	10	23	19	16	5	16	14	11	22	21	17	5
10	10	14	19	22	23	18	5	10	15	20	21	24	18	6
Total	8	34	17	51	9	24	19	7	7	10	13	8	9	3

Mean CV and standard deviation (s.d.) CV are computed over the time-series.

Table 5. Relative contribution of each source of uncertainty in abundance and biomass estimates.

		Relative contribution for abundance				Relative contribution for biomass				
Year	**Age**	**A**	**L**	**ALK**	**TS-L**	**A**	**L**	**ALK**	**W-L**	**TS-L**
2006	1	11.6	81.9	−0.6	7.1	11.9	80.3	0.0	−0.4	8.2
	2	16.4	69.0	8.5	6.1	16.4	62.9	13.3	−0.1	7.4
	3	24.5	51.8	7.7	16.0	22.8	45.8	15.1	−0.2	16.6
	4	27.9	23.8	13.6	34.7	25.3	17.0	16.9	−0.4	41.2
	5	31.0	1.1	22.6	45.2	29.1	1.5	22.2	−0.5	47.7
	6	12.7	6.4	22.8	58.1	12.1	5.1	23.7	0.2	58.9
	7	16.7	−3.9	33.6	53.7	16.1	−1.9	35.5	−0.1	50.4
	8	6.4	−8.3	66.8	35.2	8.6	−6.3	65.8	−0.1	32.0
	9	1.4	−10.7	80.9	28.4	2.4	−7.5	78.4	0.1	26.6
	10	20.3	−8.1	57.6	30.2	23.6	3.6	42.8	0.7	29.3
	Total	23.9	29.7	0.0	46.3	25.6	−2.7	0.0	−0.1	77.3
2007	1	20.0	81.5	−0.7	−0.8	23.9	78.9	−2.4	−0.5	0.1
	2	21.9	73.2	3.1	1.8	18.2	77.8	2.5	0.4	1.1
	3	25.8	46.0	12.9	15.3	23.4	46.9	9.6	0.3	19.8
	4	13.8	39.6	17.4	29.3	15.8	31.5	17.4	0.1	35.2
	5	22.4	−0.8	21.9	56.5	21.9	−1.4	23.8	0.0	55.7
	6	34.0	−5.1	17.1	54.0	33.8	−8.5	17.6	0.4	56.7
	7	18.0	−8.7	37.4	53.3	17.0	−6.4	39.6	0.1	49.8
	8	19.0	−22.6	54.6	49.0	23.7	−32.7	53.1	−0.2	56.2
	9	9.4	3.1	67.4	20.1	9.8	5.1	63.1	0.1	21.8
	10	13.5	29.9	52.3	4.3	12.3	37.7	45.0	−0.1	5.1
	Total	36.0	61.9	0.0	2.1	32.5	−2.2	0.0	0.4	69.3
2008	1	35.3	36.6	26.2	1.8	28.5	33.8	36.6	0.2	0.9
	2	72.2	26.0	1.3	0.5	76.4	23.5	1.5	−0.7	−0.7
	3	89.5	−7.1	11.3	6.3	98.2	−12.9	9.0	−0.6	6.2
	4	53.9	−5.4	36.2	15.3	52.2	−1.8	33.8	−0.6	16.4
	5	36.8	7.2	38.0	18.0	30.6	7.4	40.8	0.1	21.1
	6	18.6	21.7	26.7	33.0	18.6	18.8	27.1	0.1	35.4
	7	32.0	−15.9	35.8	48.1	29.1	−13.6	32.2	0.6	51.7
	8	16.0	8.0	44.0	32.1	17.7	25.0	29.1	1.5	26.7
	9	14.9	4.1	48.8	32.2	17.5	24.4	31.4	0.8	25.8
	10	17.3	39.1	28.3	15.2	18.1	46.2	20.7	0.6	14.3
	Total	80.7	15.4	0.0	3.9	81.5	−8.5	0.0	−0.4	27.4
2009	1	14.3	80.6	0.0	5.0	14.1	80.1	0.1	0.2	5.5
	2	2.1	98.6	7.2	−7.8	3.8	93.1	8.6	0.8	−6.2
	3	12.0	86.4	0.1	1.5	12.2	85.6	0.7	−0.2	1.7
	4	18.0	56.2	17.2	8.7	19.5	50.2	20.7	−0.2	9.8
	5	36.2	0.8	69.1	−6.1	41.2	2.1	63.3	0.4	−7.0
	6	53.6	−41.1	83.3	4.2	60.9	−36.8	74.1	−1.4	3.3
	7	114.0	−68.5	65.0	−10.8	96.1	−32.4	45.7	0.9	−10.3
	8	70.3	−13.4	31.2	11.8	76.2	−8.3	19.1	−0.7	13.7
	9	29.7	−6.1	66.2	10.1	32.3	7.8	52.9	−0.5	7.4
	10	49.9	11.1	35.4	3.6	45.8	26.0	24.0	0.0	4.2
	Total	17.2	75.6	0.0	7.2	57.7	23.4	0.0	−0.1	19.0
2010	1	28.3	56.8	−0.8	15.7	27.3	57.1	0.1	0.0	15.4
	2	29.0	26.4	12.1	32.5	31.3	23.1	10.4	−0.2	35.4
	3	42.8	6.2	20.2	30.8	48.3	−1.8	13.0	−0.3	40.9
	4	53.7	−20.7	12.3	54.7	54.6	−19.3	9.0	0.0	55.8
	5	30.7	−12.9	39.8	42.4	30.0	−0.6	32.8	0.0	37.8
	6	30.3	−44.9	92.5	22.1	29.4	−27.0	76.7	0.4	20.5
	7	−22.2	−77.2	206.0	−6.9	−8.8	−28.1	141.0	0.5	−4.1
	8	47.2	20.2	41.1	−8.6	50.0	16.4	42.7	0.4	−9.5
	9	50.5	−1.4	56.6	−5.7	57.1	4.7	46.7	0.1	−8.6
	10	74.3	22.2	6.2	−2.7	70.7	30.2	0.6	0.5	−2.1
	Total	35.2	25.9	0.0	38.9	45.4	−9.4	0.0	−0.1	64.1

A, acoustic backscatter; L, length frequency distribution; ALK, age–length key; TS-L, target strength-at-length relationship; W-L, weight-at-length relationship.

Ianelli *et al.*(2011) treat the survey data as providing an index of total abundance, assuming a lognormal distribution and a separate multinomial component that relates to the age-composition data treated as proportions with a scaled sample size. The work presented here provides the ability to use, or at least, to take advantage of, the direct estimates from the sampling process and avoids having to

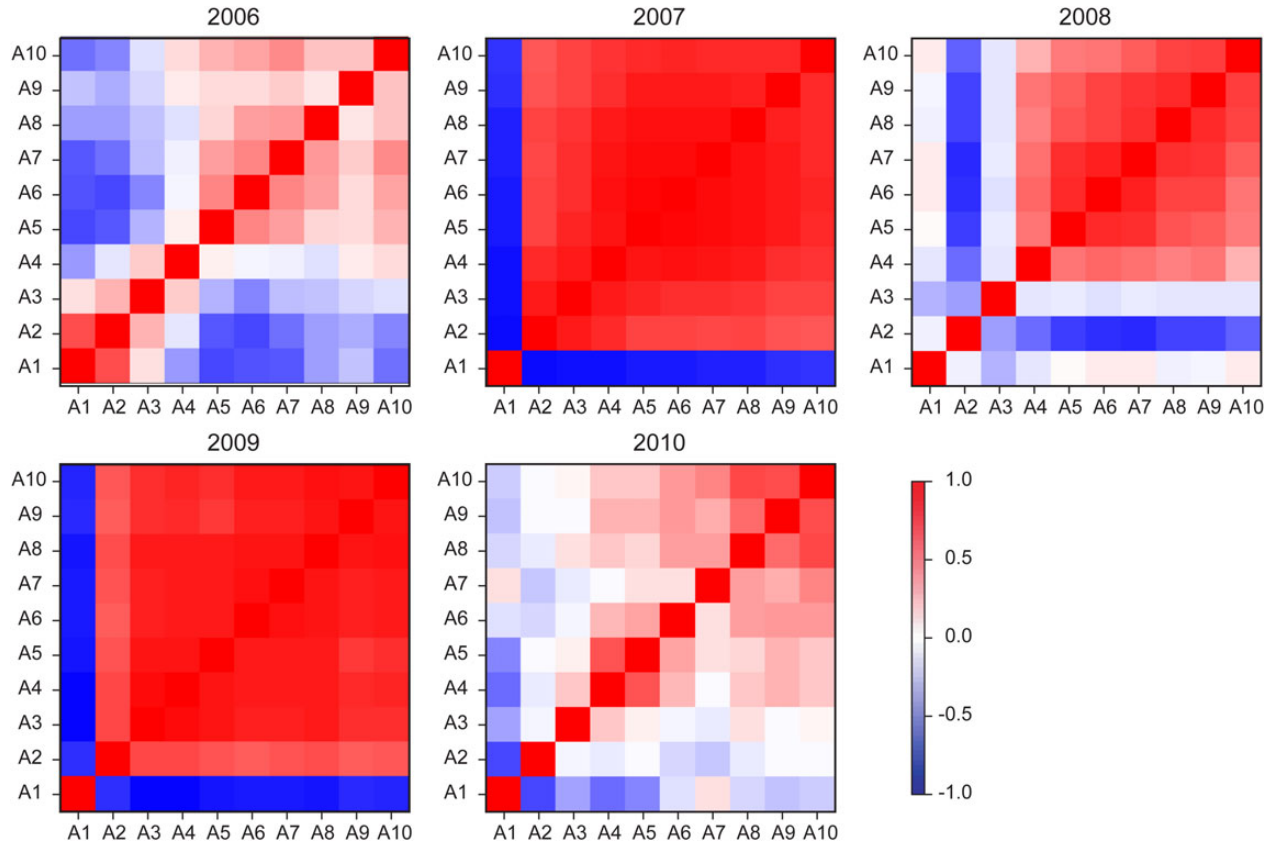

Figure 12. Correlation matrices of proportion-at-age estimates obtained by simulation for years 2006–2010.

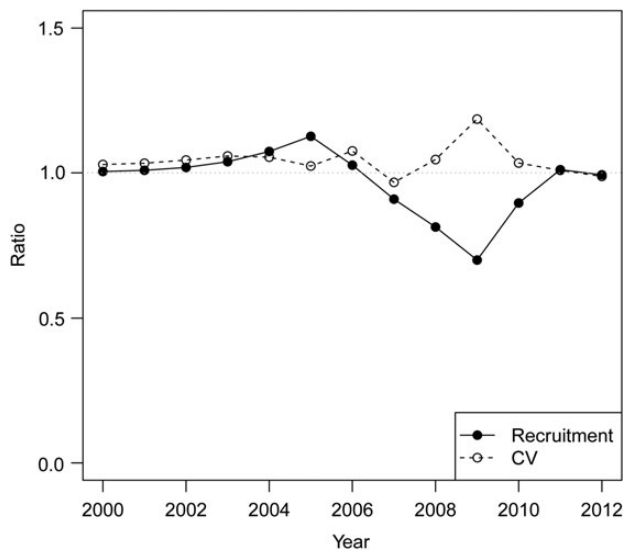

Figure 13. Impact of applying the estimated correlation matrices and associated abundance-at-age for years 2006–2010 within the stock assessment model. Solid line: ratio of assessment model estimates of age-1 recruits (new estimate on original one). Dashed line: analogous ratio of the CVs of the recruitment estimates.

assume a strict multinomial sampling process (in which all the correlations are by definition, negative). The correlation matrices produced by our simulation framework differ greatly from what the walleye pollock stock assessment assumes. Indeed, covariance terms for a multinomial distribution, whatever the chosen effective sample size, are always negative, which is clearly not the case in Figure 12.

Applying the revised estimates of numbers-at-age and their covariance estimates (from our simulation framework) for years 2006–2010 within the assessment model results in lowered recruitment estimates in some years and higher CVs, compared with the original assessment (Figure 13). Such changes in the point estimates have obvious consequences to subsequent biomass estimates used for management. Furthermore, accurate estimates of uncertainty within the assessment model are considered important. In this approach, this type of likelihood specification is an improvement over the traditional assumptions about the sampling process. Thus, the estimates of uncertainty shown here should be considered more reliable.

Finally, the methods presented here allowed us to assess the relative contribution of each of the major sources of uncertainty in the Bering Sea walleye pollock AT survey and to improve the way that survey data are included in stock assessment models. We suggest that simulated uni- and multivariate distributions of abundance-at-age estimates based on realized sampling characteristics are an improvement over estimates of covariance derived from assumed multinomial likelihood currently used in assessment models.

Acknowledgments

Data used were provided by the Resource Assessment and Conservation Engineering Division of the Alaska Fishery Science Center (NOAA). Funding for the work was provided by NOAA

Stock Assessment Analytical Methods project #0004. The findings and conclusions in the paper are those of the authors and do not necessarily represent the views of NOAA, National Marine Fisheries Service. We thank Pierre Petitgas for helpful discussions. The authors also thank the three anonymous reviewers and the editor, Richard O'Driscoll, for their constructive comments that helped improve the manuscript.

Annex1: the Gibbs sampler

The algorithm starts with the arbitrary assignment of values in the interval $[-\infty;\ y_c]$ to data points where Z equals zero. The algorithm proceeds iteratively. Each iteration entails selecting the set of samples to be modified at random. At iteration $(\nu+1)$ and for the current sample i, a new Gaussian value $Y_i^{\nu+1}$ is generated as a function of its conditional distribution to the other Gaussian values, both at the data points where the Gaussian value is determined based on the sampled data and at the data points where the Gaussian value is to be modified. Practically, this is equivalent to simulating the Gaussian value $Y_i^{\nu+1}$ such that

$$Y_i^{\nu+1} = Y_i^{KS} + \sigma_{KS}(i)R_i,$$

where Y_i^{KS} is the simple kriging estimate of Y_i derived from the other data (those for which the Gaussian value is determined from the sample data and those for which the Gaussian value is to be modified), $\sigma_{KS}(i)$ is the standard deviation of the simple kriging estimation error and R_i refers to a normalized Gaussian residual. This new Gaussian value is kept only if it falls within the interval $[-\infty;\ y_c]$; otherwise, another value is simulated. After a large number of iterations (in practice 1000), it can be shown that Y_i tends towards its conditional distribution defined from the starting values of Y (i.e. the determined Gaussian values and the Gaussian values that are undetermined but below y_c).

The Gibbs sampler thus allows correct specification of the variable Y at all data points, so that its histogram is Gaussian (without a spike) and its spatial structure respects the inferred Gaussian model of Y (Chilès and Delfiner, 1999; Lantuéjoul, 2002; Woillez, 2007). This new variable lends itself to geostatistical conditional simulations.

Annex2: Inference of the variogram of a Gaussian random field Y from a lower-cut Gaussian random field Y^+

Let $Y(x)$ be a regionalized Gaussian random field. The inference of its spatial model, i.e. its variogram, is direct through the structural analysis. However, when only a function of the Gaussian random field, for instance a lower-cut Gaussian random field (Gaussian histogram with a spike of equal values on the left part), is available, the inference of the variogram of the Gaussian is also possible. In this case, the inference is indirect and based on some developments in Hermite polynomials (Woillez, 2007).

Hermite polynomials

The Hermite polynomials are defined from the Rodrigues' formula as

$$H_n(y) = \frac{1}{\sqrt{n!}g(y)}\frac{d^n g(y)}{dy^n},$$

where $1/\sqrt{n!}$ is a normalizing factor and g is the standard Gaussian pdf. $H_n(y)$ is a polynomial of degree n defined by recurrence $(n>0)$:

$$\begin{aligned} H_0(y) &= 1 \\ H_1(y) &= -y \\ H_2(y) &= \frac{1}{\sqrt{2}}(y^2-1) \\ H_{n+1}(y) &= -\frac{1}{\sqrt{n+1}}yH_n(y) - \sqrt{\frac{n}{n+1}}H_{n-1}(y). \end{aligned}$$

These polynomials, for $n>0$, have a null mean ($Y(x)$ being a standard gaussian RF):

$$E[H_n[Y(x)]] = \int H_n(y)g(y)\mathrm{d}y.$$

A unit variance:

$$\mathrm{Var}[H_n[Y(x)]] = E[H_n[Y(x)]^2] = 1.$$

And check, for $n \neq p$, positive or null:

$$\mathrm{Cov}[H_p[Y(x)], H_n[Y(x)]] = E[H_p[Y(x)]H_n[Y(x)]] = 0.$$

Development of a lower-cut Gaussian in Hermite polynomials

All function $f[Y(x)]$ can be developed using Hermite polynomials:

$$f[Y(x)] = f_0 + f_1H_1[Y(x)] + f_2H_2[Y(x)] + \cdots = \sum_{n=0}^{\infty} f_nH_n[Y(x)].$$

As Hermite polynomials are orthonormal, one can write: $E[f[Y(x)]H_n[Y(x)]] = f_n$.

Indeed, we have:

$$\begin{aligned} E[f[Y(x)]H_n[Y(x)]] &= E\left[\left[\sum_{p=0}^{\infty} f_pH_p[Y(x)]\right]H_n[Y(x)]\right] \\ &= \sum_{p=0}^{\infty} f_pE[H_p[Y(x)]H_n[Y(x)]] \\ &= f_n. \end{aligned}$$

Given all the f's, computation of the coefficients is possible based on:

$$f_n = E[f[Y(x)]H_n[Y(x)]] = \int f(y)H_n(y)g(y)\mathrm{d}y$$

with $f_0 = E[f(Y(x))]$.

For the development of a lower-cut Gaussian, let's consider the function $f[Y(x)]$ of a variable equal to:

$$\begin{aligned} &y_c \text{ if } Z(x)=0, \text{ such as } G(y_c) = P(Y<y_c) = P(Z=0) \\ &Y(x) \text{ if } Z(x)>0. \end{aligned}$$

The considered variable is a lower-cut Gaussian that will be named $Y^+(x)$. This variable $Y^+(x)$ is a continuous function of the Gaussian variable $Y(x)$ such as:

$$Y^+(x) = y_cI_{Y(x)\leq y_c} + Y(x)I_{Y(x)>y_c} = y_cI_{Z(x)=0} + Y(x)I_{Z(x)>0}.$$

This function can be developed using Hermite polynomials and the analytical expression of its coefficients is:

$$f_n = \int (y_c I_{Y(x)\le y_c} + Y(x) I_{Y(x)>y_c}) H_n(y) g(y) \mathrm{d}y,$$
$$\text{for } n = 0,\ f_0 = g(y_c) + y_c G(y_c)$$
$$\text{for } n = 1,\ f_1 = G(y_c) - 1$$
$$\text{for } n \ge 2,\ f_n = \frac{1}{\sqrt{n(n-1)}} H_{n-2}(y_c) g(y_c) + \frac{y_c}{\sqrt{n}} H_{n-1}(y_c) g(y_c).$$

Thus, its development:

$$Y^+(x) = y_c G(y_c) + g(y_c) - y(G(y_c) - 1) + \sum_{n\ge 2} \left(\frac{y_c}{\sqrt{n}} H_{n-1}(y_c) g(y_c) + \frac{1}{\sqrt{n(n-1)}} H_{n-2}(y_c) g(y_c) \right).$$

Link between the covariance of a function of the Gaussian variable and the covariance of the Gaussian variable itself

The spatial covariance of a function of a Gaussian variable can be developed thanks to orthogonality of Hermite polynomials and the assumed bigaussian nature of the bivariate distributions $(Y(x), Y(x+h))$. We have then:

$$\mathrm{Cov}[f(Y(x+h)), f(Y(x))] = E[(f(Y(x+h)) - f_0)(f(Y(x)) - f_0)]$$
$$= E\left[\left(\sum_{n=1}^{\infty} f_n H_n[Y(x+h)] \right) \left(\sum_{p=1}^{\infty} f_p H_p[Y(x)] \right) \right]$$
$$= \sum_{n=1}^{\infty} (f_n)^2 E[H_n[Y(x+h)] H_n[Y(x)]]$$
$$= \sum_{n=1}^{\infty} (f_n)^2 [\mathrm{Cov}[Y(x+h), Y(x)] / \mathrm{Var}[Y(x)]]^n.$$

This expression allows the inference of the spatial model of the Gaussian by indirectly fitting the covariance of a function of the Gaussian variable, in our case the lower-cut Gaussian, to its known experimental variogram.

References

Bez, N. 2002. Global fish abundance estimation from regular sampling: the geostatistical transitive method. Canadian Journal of Fisheries and Aquatic Sciences, 59: 1921–1931.

Boyd, C., Woillez, M., Bertrand, S., Castillo, R., Bertrand, A., and Punt, A. E. 2015. Bayesian posterior prediction of the patchy spatial distributions of small pelagic fish in regions of suitable habitat. Canadian Journal of Fisheries and Aquatic Sciences, 72:290–303.

Chilès, J-P., and Delfiner, P. 1999. Geostatistics: Modelling Spatial Uncertainty. John Wiley, New York.

Chilès, J-P., and Lantuéjoul, C. 2005. Prediction by conditional simulation: models and algorithms. In Space, Structure and Randomness, pp. 39–68. Ed. by M. Bilodeau, F. Meyer, and M. Schmitt. Springer, New York.

Demer, D. A. 2004. An estimate of error for the CCAMLR 2000 survey estimate of krill biomass. Deep Sea Research Part II: Topical Studies in Oceanography, 51: 1237–1251.

Emery, X. 2008. A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. Computers & Geosciences, 34: 1850–1862.

Foote, K. G., and Traynor, J. J. 1988. Comparison of walleye Pollock target strength estimates determined from in situ measurements and calculations based on swimbladder form. Journal of the Acoustical Society of America, 88: 9–17.

Galli, A., and Gao, H. 2001. Rate of convergence of the Gibbs sampler in the Gaussian case. Mathematical Geology, 33: 653–677.

Gimona, A., and Fernandes, P. G. 2003. A conditional simulation of acoustic survey data: advantages and potential pitfalls. Aquatic Living Resources, 16: 123–129.

Goulard, M., and Voltz, M. 1992. Linear coregionalization model: tools for estimating and choice of multivariate variograms. Mathematical Geology, 24: 269–286.

Hanselman, D. 2009. A general method to adjust catch limits/targets with survey uncertainty. Draft document submitted to the 21–22 May 2009 ACL workshop.17 pp.

Honkalehto, T., Patton, W., de Blois, S., and Williamson, N. 2002. Echo integration-trawl survey results for walleye pollock (*Theragra chalcogramma*) on the Bering Sea shelf and slope during summer 2000. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-126. 66 pp.

Honkalehto, T., McCarthy, A., Ressler, P., Williams, K., and Jones, D. 2012. Results of the acoustic-trawl survey of walleye pollock (*Theragra chalcogramma*) on the U.S. and Russian Bering Sea shelf in June-August 2010 (DY1006). AFSC Processed Rep. 2012-01, 57 pp. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Ianelli, J. N., Honkalehto, T., Barbeaux, S., Kotwicki, S., Aydin, K., and Williamson, N. 2011. Eastern Bering Sea walleye pollock. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions. North Pac. Fish. Mgmt. Council, 605 W. 4th Ave., STE 306, Anchorage, AK 99501-2252, 1: 51–168.

Juntunen, T., Vanhatalo, J., Peltonen, H., and Mäntyniemi, S. 2012. Bayesian spatial multispecies modelling to assess pelagic fish stocks from acoustic- and trawl-survey data. ICES Journal of Marine Science, 69: 95–104.

Lantuéjoul, C. 2002. Geostatistical simulation: models and algorithms. Springer, Berlin. 256 pp.

Lantuéjoul, C., and Desassis, N. 2012. Simulation of a Gaussian random vector: a propagative version of the Gibbs sampler. The 9th International Geostatistics Congress, Oslo, Jun 2012, Oslo, Norway. 8 pp.

Løland, A., Aldrin, M., Ona, E., Hjellvik, V., and Holst, J. C. 2007. Estimating and decomposing total uncertainty for survey-based abundance estimates of Norwegian spring-spawning herring. ICES Journal of Marine Science, 64: 1302–1312.

MacLennan, D. N., Fernandes, P. G., and Dalen, J. 2002. A consistent approach to definitions and symbols in fisheries acoustics. ICES Journal of Marine Science, 59: 365–369.

O'Driscoll, R. L. 2004. Estimating uncertainty associated with acoustic surveys of spawning Hoki (*Macruronus Novaezelandiae*) in Cook Strait, New Zealand. ICES Journal of Marine Science, 61: 84–97.

Petitgas, P. 1993. Geostatistics for fish stock assessments: a review and an acoustic application. ICES Journal of Marine Science, 50: 285–298.

Petitgas, P., Doray, M., Massé, J., and Grellier, P. 2011. Spatially explicit estimation of fish length histograms, with application to anchovy habitats in the Bay of Biscay. ICES Journal of Marine Science, 68: 2086–2095.

Prager, M. H., Porch, C. E., Shertzer, K.W., and Caddy, J. F. 2003. Targets and limits for management of fisheries: a simple probability-based approach. North American Journal of Fisheries Management, 23: 349–361.

Renard, D., Bez, N., Desassis, N., Beucher, H., Ors, F., and Laporte, F. 2010. RGeostats: The Geostatistical Package 5.3.8. MINES ParisTech. http://cg.ensmp.fr/rgeostats.

Rivoirard, J., Simmonds, J., Foote, K. G., Fernandes, P., and Bez, N. 2000. Geostatistics for Estimating Fish Abundance. Blackwell Science, Oxford. 206 pp.

Rose, G., Gauthier, S., and Lawson, G. 2000. Acoustic surveys in the full monte: simulating uncertainty. Aquatic Living Resources, 13: 367–372.

Shertzer, C. E., Prager, M. H., and Williams, E. H. 2008. A probability-based approach to setting annual catch levels. Fishery Bulletin, 106: 225–232.

Simmonds, E. J., Gutiérrez, M., Chipollini, A., Gerlotto, F., Woillez, M., and Bertrand, A. 2009. Optimizing the design of acoustic surveys of Peruvian anchoveta. ICES Journal of Marine Science, 66: 1341–1348.

Simmonds, E. J., and MacLennan, D. 2005. Fisheries Acoustics: Theory and Practice. Blackwell Science, Oxford. 437 pp.

Traynor, J. J. 1996. Target strength measurements of walleye pollock (*Theragra chalcogramma*) and Pacific whiting (*Merluccius productus*). ICES Journal of Marine Science, 53: 253–258.

Wackernagel, H. 1995. Multivariate Geostatistics: An Introduction with Applications. Springer, Berlin.

Walline, P. D. 2007. Geostatistical simulations of eastern Bering Sea walleye pollock spatial distributions, to estimate sampling precision. ICES Journal of Marine Science, 64: 559–569.

Williams, K., Punt, A. E., Wilson, C. D., and Horne, J. K. 2011. Length-selective retention of walleye pollock, *Theragra chalcogramma*, by midwater trawls. ICES Journal of Marine Science, 68: 119–129.

Williamson, N. J., and Traynor, J. J. 1996. Application of a one-dimensional geostatistical procedure to fisheries acoustic surveys of alaskan pollock. ICES Journal of Marine Science, 53: 423–428.

Woillez, M. 2007. Contributions géostatistiques à la biologie halieutique. Thèse de Docteur en Géostatistique, Ecole Nationale Supérieure des Mines de Paris, France. 175 pp.

Woillez, M., Rivoirard, J., and Fernandes, P. G. 2009. Evaluating the uncertainty of abundance estimates from acoustic surveys using geostatistical simulations. ICES Journal of Marine Science, 66: 1377–1383.

Zwolinski, J., Fernandes, P. G., Marques, V., and Stratoudakis, Y. 2009. Estimating fish abundance from acoustic surveys: calculating variance due to acoustic backscatter and length distribution error. Canadian Journal of Fisheries and Aquatic Sciences, 66: 2081–2095.

Handling editor: Richard O'Driscoll