Bayesian spatial multispecies modelling to assess pelagic fish stocks from acoustic- and trawl-survey data

Teppo Juntunen1*, Jarno Vanhatalo1, Heikki Peltonen2, and Samu Mäntyniemi1
1Fisheries and Environmental Management Group, Department of Environmental Sciences, University of Helsinki, PO Box 65, FI-00014 Helsinki, Finland
2Marine Research Centre, Finnish Environment Institute, PO Box 140, FI-00251 Helsinki, Finland

*Corresponding Author: tel: +358 50 4150608; fax: +358 9 19158754; e-mail: teppo.juntunen@helsinki.fi


Received 26 April 2011; accepted 21 October 2011.

A Bayesian spatial model was constructed to estimate the abundance of multiple fish species in a pelagic environment. Acoustic- and trawl-survey data were combined with environmental data to predict the spatial distribution of (i) the acoustic backscattering of fish, (ii) the relative proportion of each species, and (iii) their mean length in the Gulf of Finland in the northeastern Baltic Sea. By combining the three spatial model layers, the spatial distribution of the biomass of each species was estimated. The model consists of a linear predictor on environmental variables and a spatial random effect given by a Gaussian process. A Bayesian approach is a natural choice for the task because it provides a theoretically justified means of summarizing the uncertainties from various model layers. In the study area, three species dominate pelagic waters: sprat (Sprattus sprattus), herring (Clupea harengus), and three-spined stickleback (Gasterosteus aculeatus). Results are presented for each model layer and for estimated total biomass for each species in 2 × 2 km lattices. The posterior mean and central 95% credible intervals of total biomass were sprat 45.7 kt (27.7–71.6), herring 24.6 kt (9.7–41.3), and three-spined stickleback 19 kt (0.9–32).

Keywords: acoustic survey, Bayesian spatial modelling, Gaussian process, multispecies, trawl survey.

Introduction

Acoustic surveys have a long history in fisheries assessment and management (Simmonds and MacLennan, 2005). The method became feasible after understanding the relationship between acoustic backscattering from an echosounder and fish biomass, and the dependence of backscattering on fish size (Foote, 1983, 1987). In acoustic surveys, concurrent trawls are usually used to estimate the relative proportion of species and their size distribution. Incorporated acoustic and trawl surveys are important especially to give fisheries-independent information on the dynamics of fish stocks. Time-series of such surveys are often applied in tuning stock-assessment models (Iversen et al., 1993; Gjøsaeter et al., 1998; Korsbrekke et al., 2001).

Incorporation of acoustic-survey data with spatial data on environmental variables allows the construction of sophisticated models of species’ distributions. The spatial prediction of a species distribution produced by linking survey data with environmental variables has applications not only in the assessment and management of fisheries and other aquatic resources (Maravelias and Reid, 1995), but a wide range of methods has been extensively used, for example, in ecology and conservation planning to predict the abundance of species in locations outside the observation points (Guisan and Zimmermann, 2000; Austin, 2002; Latimer et al., 2006). Most of these works assume that environmental factors determine the distribution of a species and that species have reached equilibrium with respect to the factors. However, there are often other processes inducing spatial changes in the distribution of a species that are not linked to environmental variations, but arise, for example, from animal behaviour (schooling, migration, dispersion), so it would be an exception for a species to be in equilibrium with respect to environmental factors. In predicting North Sea herring (Clupea harengus) abundance, Maravelias et al. (1996) used kriging to take into account both the small-scale effect of schooling and the impacts of larger-scale oceanographic features. Here, we assign a Gaussian process prior for the spatial field, which is a Bayesian analogue to the kriging approach.

Bayesian methods provide an attractive alternative in spatial fish-stock assessment, as in many other fish-stock assessment tasks, where they have been applied recently (Michielsens et al., 2008; Rivot et al., 2008; Punt et al., 2011). The virtues of Bayesian modelling include flexibility in model structure and in
the application of different data sources, and they provide straightforward solutions on how to incorporate uncertainty. Hammond and Swartzman (2001) presented a spatial Bayesian tool for estimating fish composition based on echogram analysis, depth, and temperature, and the model was tuned with trawl-survey data. Brierley et al. (2003) used a Bayesian maximum-entropy approach to estimate Antarctic krill (*Euphausia superba*) biomass based on acoustic data, but without explanatory environmental variables. Assessment of fish stocks with acoustic- and trawl-survey data can be divided into four subtasks: (i) assessment of the spatial distribution of acoustic backscattering over the focus research area, (ii) estimation of the spatial distribution of the relative proportion of each fish species, (iii) estimation of the spatial distributions of fish size and acoustic target strength of each fish species in the area, and (iv) incorporation of the information from the three other subtasks to provide estimates of the spatial distribution of biomass of each fish species.

Here, we constructed a Bayesian model consisting of three geographic layers resembling GIS (geographic information system) applications to estimate spatial distributions and uncertainties in the abundance of three pelagic fish species. The model combines a linear predictor of environmental variables and a spatial random effect that accounts for factors, such as schooling, not explained by the environmental variables. Sophisticated computational methods were applied, allowing computations in a feasible time frame for high spatial resolution.

**Material and methods**

The study was conducted on the Gulf of Finland (GoF) in the northeastern Baltic Sea. The maximum length of the GoF is \( \sim 400 \) km, the surface area is \( \sim 30,000 \) km\(^2\), and the average depth is \( \sim 37 \) m. The GoF has strong estuarine characteristics mainly because of the substantial flow of freshwater into the easternmost part. Herring and sprat (*Sprattus sprattus*) dominate the pelagic fish assemblage and dominate fish catches in the GoF. Additionally, three-spined stickleback (*Gasterosteus aculeatus*) are found at high density in many locations, and although other marine and freshwater species are also in the GoF, they tend to be at low density in the pelagic zone.

**Sampling of fish stocks with trawl and hydroacoustics**

Material was collected on board RV “Muikku” during a research cruise from 17 to 28 July 2004 that focused on both ecosystem function and the characteristics of pelagic fish stocks in the research project “Trophic Interactions in the Baltic Sea: How are Zooplankton Communities and Commercially Important Fish Stocks Regulated?” As the survey was not designed solely to support fish-stock assessment, it did not cover parallel transects, as do many fisheries acoustic surveys, but followed a more adaptive strategy. Trawl stations were intended to cover as much of the survey area as possible, but some compromises were necessary because the acoustic and trawl surveys had to be integrated with the ecosystem survey. The cruise was conducted in territorial waters and Exclusive Economic Zones of Finland and Estonia, but did not cover the Russian zone in the eastern part of the GoF. Regardless, the whole GoF was included in the predictive spatial modelling conducted in the current study (Figure 1).

The research material consisted of fish samples obtained with a pelagic trawl and hydroacoustic data on pelagic fish. A set of 20 pelagic trawl hauls was made to sample the fish species composition and length distribution of each species. The vertical and horizontal openings of the trawl were ca. 15 and 35 m, respectively, and the codend mesh was 5-mm bar. Species composition of each catch was determined, and total fish length was recorded at intervals of 0.5 cm (for detail, see Peltonen et al., 2007).

The hydroacoustic data were collected with a Simrad EY500 scientific echosounder (Simrad, 1995) equipped with a split-beam Simrad ES38-12 transducer. The echosounder had a transmit frequency of 38 kHz and a transmit power of 250 W. The beam width of the transducer was 12° (to \(-3\) dB level from the acoustic axis). Pulse duration was set to 1 ms and pulse rate to 2 Hz. The hydroacoustic equipment was calibrated carefully with a standard copper sphere on 14 June 2004 using LOBE software (Simrad,

![Figure 1. The research area, showing the grid, depth, trawl sample sites, and EDSU points, along the acoustic-survey track.](https://academic.oup.com/icesjms/article-abstract/69/1/95/673665)
1995). During the cruise, the calibration was occasionally checked with the copper sphere. Estimates of the nautical area scattering coefficient (NASC) \( s_a \) (\( m^2 \) nautical mile\(^{-2} \)) \cite{MacLennan2002} were made with Sonar5-pro fisheries acoustic post-processing software \cite{Balk2004a, Balk2004b}. The data were analysed in sections (EDSUs, elementary distance sampling units) of 0.1 nautical miles (185.2 m). The dataset contained a total of 5821 EDSUs.

**Geographic material**

The study covered a geographic grid in which each cell covered an area of \( 2 \times 2 \) km, resulting in a total of 8515 cells in the focus area. The grid roughly covered ICES Subdivision 32, i.e. the GoF (Figure 1).

In addition to fish-stock data, geographic data were compiled to support the modelling. Depth data for the grid area were generated by resampling from digital topography data for the Baltic Sea \cite{Seifert2001}, where the original resolution was 2 min longitude and 1 min latitude (corresponding approximately to 1.85 \( \times \) 1.85 km at the coordinates of the research area). In Seifert et al. (2001), depth was an average of all known depth values in each grid cell.

Distance to shore was incorporated as an explanatory factor in the model, because it was expected that fish species and size categories would differ in spatial distribution owing to their affinity to the coast or to the open sea \cite{Uusitalo2005}. Distance to shore was determined as the shortest distance to the shore from the centre of each cell. The shoreline also included large islands.

Finally, longitude was used as an explanatory geographic variable. Many hydrodynamic, chemical, and biological variables can influence the distribution of fish species in the area \cite{Peltonen2007}. Several variables exhibit trend-like changes from west to east in the GoF mainly because of its strong estuarine characteristics, with substantial inflow of freshwater in the east, and western areas influenced by saline water inflow from the Baltic Proper. For the purposes of fish-stock assessment, it would be demanding to map a wide range of environmental variables during each survey cruise. That would make the modelling approach more data-demanding, whereas one of the goals was to develop methodologies applicable also in data-poor fisheries and in limited-resource surveys. Therefore, the assumption was made that the dynamic environmental variables could be replaced satisfactorily with stable geographic covariates in modelling the distribution of fish in the GoF.

**Model structure**

To assess the spatial distribution of the biomass of each fish species in the research area from acoustic- and trawl-survey data, we needed to predict the spatial distribution of (i) the acoustic-survey results, i.e. the NASC, (ii) the proportion of each fish species, and (iii) the average length and acoustic target strength (TS) of each species. In the third phase, modelling the average size of each species instead of the size distribution was considered appropriate because of the small size range of each fish species. Information from the three models was then combined to predict the spatial distribution of biomass of each species in the whole GoF. The construction of the model is summarized below.

**Nautical area scattering coefficient**

The midpoints of the 5821 EDSUs from the acoustic-survey material constituted the datapoints for the analyses. As we predicted at a cell level, it was natural also to treat the data at a cell level so that there was no need to account for the change in support problem resulting from data being in different resolution from the prediction \cite{Gelfand2001}. Each cell, having multiple datapoints, was given a single value that was the mean of the measurements within the boundaries of that cell, resulting in a total of 530 cells with one or more EDSUs. The variance of the resulting mean observation depends on the number of measurements averaged. However, as the numbers of averaged measurements were close to each other, we did not consider the differences between cells to be important and assumed the same variance for each cell. In every cell \( i \), the logarithm of the mean of NASC, \( \log(s_a) \), was assumed to have a Gaussian distribution with mean \( \mu_i \) and variance \( \sigma^2 \) of \( \log(s_a) \sim \text{Normal}(\mu_i, \sigma^2) \). The mean was modelled as a sum of a linear function of environmental variables and a spatially structured random effect \( \mu_i = \beta_0 + x_1^T \beta + s_i \). Here, \( x_i \) is a vector of length \( m \) that collects the environmental covariates, \( \beta \) a vector of regression parameters, \( s_i \) the spatial random effect, and \( \beta_0 \) a constant term. The regression parameters \( \beta_0, \ldots, \beta_m \) are given independent zero-mean Gaussian priors, \( \beta_i \sim \text{Normal}(0, \sigma^2) \).

The spatially structured random effect was given a zero-mean Gaussian process (i.e. Gaussian random field) prior \cite{Rasmussen2006}, which implies that any subset of random effects \( s = (s_1, \ldots, s_n) \) has a multivariate Gaussian distribution \( s \sim \text{Normal}(0, \Sigma) \). The covariance matrix \( \Sigma \) incorporates the spatial association between different locations and can be represented as a function that describes the decay in correlation between pairs of points with distance \cite{Gelfand2010}. The function can take several forms, and here we used the so-called Matern covariance function with 3/2 degrees of freedom: \( \Sigma_{ij} = \sigma^2(l - \sqrt{3} \delta_{ij}) \exp(-\sqrt{3} \delta_{ij}/l) \). The parameter \( l \) is a length scale that defines how rapidly the correlation declines with distance \( \delta_{ij} \) and \( \sigma^2 \) is a scaling parameter that governs the overall variability of the spatial effect.

The model definition was concluded by setting a hyperprior for the variance parameters (the scaling parameter in the covariance function, and the variance in the observation model) and length scale. We used half a Student’s \( t \)-distribution that works as a weakly informative prior \cite{Gelman2006, Vanhatalo2010}, assumed that the spatial correlations were \textit{a priori} short, and adjusted the range of our beliefs with the scale of the Student’s \( t \)-distribution. However, when using few degrees of freedom, e.g. 4, the posterior of the length scale was allowed to concentrate into larger values if the likelihood justified this. Similarly, the process variance that is governed mainly by variance parameters was assumed to be \textit{a priori} small, but using small degrees of freedom, was allowed to grow if necessary in light of the data. In general, the predictive behaviour of a Gaussian process with Matérn covariance function is not sensitive for a specific choice of the covariance function parameters \cite{Zhang2004}. Regardless of this, we assessed the sensitivity of the results to the priors of hyperparameters by conducting the analysis with different choices of variance for the Student’s \( t \)-distribution. The results were not sensitive to these different choices of prior.

The model construction above is valid for any number of grid points. Hence, it was possible to infer the posterior of the model parameters by including in the model only grid cells that...
Model for the relative proportion of species

The second model predicts the distribution of the proportions of each species (the main pelagic fish species, sprat, herring, and three-spined stickleback) based on species composition in trawl catches. Other species were found only sporadically and in very small numbers, so they were omitted from the modelling. Each trawl catch was considered as a representative sample of the fish community at the sample site. Each sample site with its known explanatory environmental factors gives rise to a multinomial response with a logistic link. For sample site $i$, the observed vector of counts of different species $Y_i = (y_{i1}, y_{i2}, y_{i3})$ has a distribution $Y_i \sim \text{Multinomial}(p_{i1}, p_{i2}, p_{i3})$, where $n_i = \sum_{i=1}^{d} Y_{ij}$, and $p_{id}$ is the relative proportion of species $d$ given as $p_{id} = \Phi_{id}/\sum_{d} \Phi_{id}$ and $\log \Phi_{id} = \beta_{d,0} + x_i^T \beta_{d} + s_{d,i}$. We set $\beta_{3,0} = 0$, $\beta_{1,0} = 0$, and $s_{1,1} = 0$ for identifiability of parameters. The spatial random effect $s_{d,i}$ for $d = 2, 3$ were given independent Gaussian process priors with the Matern covariance function described above. Also, the parameters were given the same priors as in the NASC model.

This model definition is much like the one for the NASC, but the observation model is not Gaussian. However, as the latent variable $\log \Phi_{id}$ followed the Gaussian model, we could again conduct the posterior inference for the parameters by treating only the grid cells from which we had observations. After this, we calculated the posterior of $\log \Phi_{id}$ in other cells and continued to calculate the posterior of the relative proportion $p_{id}$, which is the parameter of interest in this model.

Species-specific length model

Because it was the average length of each species that was of interest, this could be used to derive the average acoustic backscattering from individual fish in a grid cell, discussed below. The average length $L_d$ of species $d$ in area $i$ was modelled with a Gaussian distribution $L_d \sim \text{Normal}(\mu_d, \sigma_d^2)$, where each species has its own mean function, which is assumed to follow the similar linear plus spatial model described for the NASC. The model is identical with the NASC model, except for the observation model part, which is here normal.

Biomass model

Finally, we integrated our spatial models to estimate the distribution of biomass of each species. The integrated biomass model included a deterministic component in the relationship between fish length and acoustic-backscattering cross section and in the length–weight relationship.

Combining the acoustic signal strength, the species distribution, and mean-length models, we were able to calculate the total biomass of fish in the area. The acoustic signal strength, $s_{d,i}$, could be converted to the number of all fish $N_i$ in cell $i$ with relation $N_i = s_{d,i}/x_i^T \pi \sigma_i^2$. The average backscattering strength of one fish (not knowing its species) in cell $i$ was calculated according to $\bar{a}_i = \sum_{d} a_d \sigma_{bid} s_{d,i}$, where $\sigma_{bid}$ is the backscattering strength of one fish of species $d$ in cell $i$. The backscattering of a single fish $\sigma_{bid} = 10^{(\bar{a}_i-10)/10}$, where the target strength $TS_{id}$ of species $d$ in cell $i$ was estimated from $TS_{id} = a_d \log_{10} L_{id} + b_d$, and where $L_{id}$ is the mean length of species $d$ in cm in cell $i$, and $a_d$ and $b_d$ are the species-specific coefficients (Table 1). The mean weight $W_{id}$ in g of each species $d$ in every cell $i$ was estimated from $W_{id} = a_{id} L_{id}^{2.5}$, where $a_{id}$ and $b_{id}$ are species-specific coefficients (Table 1). Finally, the total biomass $B_d$ for each species $d$ in the whole research area was obtained by summing over all grid cells: $B_d = \sum_i N_i p_{id} W_{id}$.

Parameter estimation and spatial predictions

A fully Bayesian approach was applied to estimate the model parameters and predict the grid cells that did not contain data. The first step was to integrate over the regression parameters $\beta$ in the linear function $f_i = \beta_0 + x_i^T \beta$. As the prior for $\beta$ is Gaussian, we could integrate over them analytically to obtain $f_i \sim \text{Normal}(0, x_i^T \Sigma_\beta x_i)$, where $\Sigma_\beta$ is a diagonal matrix collecting the prior variances of $\beta_i$, and $x$ is augmented with 1 to take the constant term in the same formulation. As the covariance above is a function of a continuous variable $x_i$, this implies a Gaussian process prior for the function $f(x)$ with covariance function $x_i^T \Sigma_\beta x_j$ (see Rasmussen and Williams, 2006), so the mean $\mu = f(x) + s$ also has a zero-mean Gaussian process prior $\mu$ with covariance function $\Omega_{ij} = x_i^T \Sigma_\beta x_j + \Sigma_{ij}$.

For the log of NASC and the average-length models, we could also integrate analytically over the mean $\mu$, because the observation model is Gaussian. This treatment left only the hyperparameters in the covariance function $\Omega$ and the variance of the observation model $\sigma^2$ to be inferred numerically. We applied Markov chain Monte Carlo (MCMC) methods for them. The MCMC algorithm used was the Hamiltonian Monte Carlo (HMC; Duane et al., 1987; Neal, 1996), which was used to sample all parameters simultaneously. The HMC algorithm is a variant of the Metropolis–Hastings algorithm, which utilizes the gradient information of the posterior distribution to direct the sampling to regions with high posterior probability. We sampled 1500 iterations, with a thinning of 6, and discarded the first 500 samples as a burn-in. This left 1000 samples, for which we calculated the potential scale-reduction factor to assess the convergence (Gelman et al., 2004). We took four steps with a step-size adjustment of 0.04 between each consecutive proposal and used moment persistence with decay 0.6 (see Neal, 1996, for detail on the HMC parameters).

With the multinomial observation model in the species–composition model, we could not marginalize out the latent variable $\log \Phi_{id}$. The approach taken here was to use Gibbs sampling and to alternate the sampling of the latent variables, given hyperparameters, and the sampling of hyperparameters, given the latent variables. All latent variables were updated simultaneously using a scaled HMC described by Vanhatalo and Vehtari (2007). We took 20 steps with a step-size adjustment of 0.15 for the latent variables.

The hyperparameters were sampled with regular HMC, as with the Gaussian models, but here the number of steps was four, the step-

Table 1. Parameter values for $a$ and $b$ used in target strength (TJ Fässler and Gorska, 2009) equations, and the $\alpha$ and $\beta$ parameters of length–weight equations estimated from the trawl samples analysed in this study (for stickleback, the $T5$–length dependence was assumed the same as for herring).

<table>
<thead>
<tr>
<th>Species</th>
<th>$a$</th>
<th>$b$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sprat</td>
<td>20</td>
<td>-65.08</td>
<td>0.0073</td>
<td>2.9791</td>
</tr>
<tr>
<td>Herring</td>
<td>20</td>
<td>-63.88</td>
<td>0.0093</td>
<td>2.8802</td>
</tr>
<tr>
<td>Stickleback</td>
<td>20</td>
<td>-63.88</td>
<td>0.0099</td>
<td>3.0058</td>
</tr>
</tbody>
</table>
size adjustment was 0.1, and we did not use moment persistence. In all, 1700 samples were taken, thinning 90, and removing 700 as a burn-in. Such thinning was required because the correlation between latent variables and hyperparameters was strong, leading to high autocorrelation in the Markov chain (Vanhatalo and Vehtari, 2007).

The prediction into grid cells without observations was straightforward, utilizing the properties of a Gaussian process (e.g. Vanhatalo, 2010), which allowed sampling of the parameters at prediction locations, given the posterior samples from the MCMC inference. However, the standard computationally efficient practice to sample only from the marginal distribution of the cell did not work here because we needed to integrate over all cells, requiring samples from a joint posterior.

We compared several models with a different spatial prior (i.e. a different covariance function from Matérn) and linear function with different subsets of explanatory covariates (longitude, distance to shore, and water depth). The model comparison was conducted using tenfold cross-validation for the log-predictive distribution (e.g. Vehtari and Lampinen, 2002). There was no significant difference between covariance functions, and the Matérn function was chosen because it is common in spatial modelling. In the models other than the average-length model, the best performance was obtained with a linear function utilizing all explanatory variables. In the average-length model, the explanatory variables were distance to shore and water depth.

With Gaussian observation models, the relevance of the linear predictors (longitude, water depth, and distance to shore) could be interpreted easily by examining their effect on $f$. The interpretation of the predictive effects of the environmental variables in the relative-proportion model was more cumbersome owing to the non-linear link function. However, using, for example, the average predictive comparison of Gelman and Pardoe (2007), we could examine the effect of individual components of the prediction on average and obtain a summary with interpretation similar to the weights of a linear model. All computations were conducted using the Matlab toolbox GPstuff, freely available from the Internet (http://www.lce.hut.fi/research/mm/gpstuff/).

Results

Nautical area scattering

The posterior of NASC is summarized in Figure 2. The model explained 90% of the total variation in the log of the NASC. The spatial component $s$ was more important than the linear model on environmental variables $f$ in the sense that $s$ contributed more to the total variation of log($sA$). The spatial random effect had a correlation distance of $\sim 10$ km, the distance between two locations where the spatial correlation had declined to 5% of its maximum value. Figure 3 presents the empirical semivariogram (Gelfand et al., 2010) and the posterior of the theoretical variogram calculated from the covariance function. The semivariogram fits well inside the 95% posterior credible interval. The semivariogram of the residual, $y - E[y|$data$]$, shows no spatial correlation in the residuals between model predictions and data. The spatial correlation is also clearly visible in the NASC prediction map, which reveals several high- and low-backscattering strength areas of size $\sim 10$ km radius along the survey route (Figure 2). There is a clear spatial pattern in the western GoF. As there were no observations from the eastern GoF, the predictions there were based only on the environmental part of the model, whereas the spatial component simply smoothed the predictive surface. The model produced increasing estimates for the posterior expectation of $sA$ towards the east, away from the shore, and towards deeper water (Figure 4). Also, the posterior uncertainty

Figure 3. The empirical semivariogram of the log NASC data together with the posterior mean and 95% credible interval of the theoretical variogram calculated from the covariance function of the spatial component in the model and the semivariogram of the residual of data and posterior predictive mean of the log NASC. The semivariogram was calculated with 1.5 km discretization. The empirical variogram fits well inside the 95% posterior credible interval of theoretical variogram.
increased towards the east, which is reasonable because there were no observations there.

Relative proportion of species
According to the model, sprat was the dominant species in the central GoF (Figure 5), and there was a clear increasing trend in the relative proportion of herring towards the east. The trend was opposite for stickleback, and there was no significant eastward effect for sprat. Herring, on the other hand, were more common near shore, and sticklebacks were predicted to be the most common in the western GoF. The uncertainties in the relative-proportion predictions were more inaccurate than the NASC predictions. The few trawl hauls relative to the hundreds of acoustic observations apparently influenced the uncertainty estimates.

The spatial-model component had practically no effect on the predictions for any of the species. The relative proportion of stickleback increased with an increase in depth, whereas the relative proportions of herring and sprat decreased towards deeper water. On the other hand, stickleback and herring were more common close to shore, and the relative proportion of sprat increased towards offshore. The relative proportions of sprat and herring increased and that of stickleback decreased towards the east (Figure 4).

Mean length
The predicted mean lengths in the research area for sprat were ca. 80–100 mm (Figure 6), and the mean lengths for herring were ~10 mm larger than for sprat. The largest fish were in the central GoF, and the smaller ones mainly in south coastal waters. For sprat and herring, there were no clear patterns in uncertainty, but for stickleback, uncertainty increased towards shore. As the data on mean length were small compared with the acoustic dataset, and the observation locations were, on average, farther apart, the spatial component of the model was presumably less important than with the nautical-area scattering data. With herring and sprat, the spatial correlation length was ~5 km, so there was no spatial correlation between most of the datapoints. The linear

Figure 4. The average predictive comparison of the environmental variables. In the linear models (NASC and mean-length models), the average predictive comparison represents the weights of the linear model. For the non-linear model (relative-proportion model), it approximates the average change in the outcome with respect to a unit difference in the environmental variable, which has the same interpretation as the weight of a linear model. Crosses show the posterior mean, and the lines show the posterior central 95% credible interval.

Figure 5. Posterior mean and CV for the relative proportions of species. In certain areas, the relative proportions of herring and stickleback were so close to zero that the CV became unstable; for the purposes of illustration, the CV in those areas is limited to 2.
component of the model had a significant effect on the predictions for herring and sprat. For stickleback, the spatial correlation length was \( \sim 20 \) km.

**Biomass**

The posterior mean of the biomass of herring clearly increased towards the east, the biomass of sprat increased in an offshore direction, and the biomass of stickleback was highest in the south-western GoF (Figure 7). The uncertainty in the biomass estimate was smaller as the mean estimate increased, which can be seen from the CV maps in Figure 7.

The posterior mean for total species biomass (Figure 8) was 72.2 kt, with a central 95% credible interval of (59.2–91.7). The respective individual species estimates were 45.7 kt (27.7–71.6) for sprat, 24.6 kt (9.7–41.3) for herring, and 1.9 kt (0.9–3.2) for stickleback.

**Discussion**

The model performed well, especially with the acoustic-survey data, which had a large number of observations. It indicated a clear spatial effect on fish clustering in medium–long ranges of up to 10 km. The effects of environmental covariates became important mainly over longer distances.

Although Bayesian models can be computationally demanding, the model could be inferred rapidly with the analytical tools applied. Gaussian priors in the linear model allowed the application of partly analytical inference of posterior distributions, which allowed the use of efficient sampling algorithms tailored in particular for spatial models. The model did not account for all major sources of uncertainty, however, because we treated the TS parameters as fixed to keep the computation time short. However, the model, like any other model, could be enhanced in many ways. It would be beneficial, for instance,
to incorporate the uncertainty in TS, particularly because there has been much controversy about the dependence of the TS of clupeid fish on fish size in the Baltic and in more-saline marine areas (Peltonen and Balk, 2005; Fässler et al., 2009). Obvious improvements in modelling the NASC could also include, for example, better analysis of echograms (Fablet et al., 2009). A possible error source is the avoidance behaviour of fish near the research vessel, which leads to underestimation of stock size, although evidence of this is contradictory (Fernandes et al., 2000; Hjellvik et al., 2008). It might be questioned too whether the inclusion of the eastern GoF in the analysis, with no observations there, was justified. However, we felt it should be included because it is standard practice to generalize the results of acoustic surveys to larger areas. As is important in such extrapolations, our model predicted an increase and accumulation of uncertainty towards areas without observations. From a decision-making perspective, the great uncertainty means that the ranking of alternative management actions may become very sensitive to the potential data that could be collected from the area where prediction uncertainty is great. If this is the case, the concept of the value of information (McDonald and Smith, 1997; Mäntyniemi et al., 2009) could be used to pinpoint geographic areas where further sampling would have the greatest value in decision-making.

Model estimates for the relative proportions of fish species and their mean lengths suffer from the small dataset, given the patchy presence of fish species in the sea. For example, the model predicted stickleback to be the dominant species in certain southwestern areas of the GoF. This was because one trawl in that area showed stickleback to dominate, although there were four samples without any stickleback at all. Sprat were caught in every trawl, and herring were absent from just one trawl. Most of the results from the current study match those published before (e.g. Uusitalo et al., 2005) in that the model, for example, showed sprat to be dominant in deeper water in the central GoF, and herring more dominant close to shore. The mean length of stickleback was strongly correlated with distance from shore, suggesting that larger stickleback feed farther from shore. Generally, the sprat and herring were quite small, which may be explained partly by the time of year at which the study was undertaken, because larger herring and sprat tend to migrate seasonally between the GoF and the northern Baltic Proper (Parmanne et al., 1994).

Figure 8. Total biomass and the biomass of each species, as a probability distribution. Distribution is calculated using kernel smoothing on MCMC samples. The grey area is the posterior 95% credible interval.

Biomass estimates here were generally lower than documented elsewhere. The ICES biomass estimate for herring in Subdivisions 25–29 and 32 (Baltic Proper and GoF, not including the Gulf of Riga) for the same year (2004) was 654 732 t. If that biomass were to be spread evenly over the whole assessment area, this would mean that the GoF held ca. 88 000 t, more than threefold higher than the posterior mean estimate here and double the 97.5% quantile. However, it is not known how evenly the biomass is distributed over the assessment area, and, as discussed earlier, seasonal migrations increase the uncertainty in this type of comparison. It is even more difficult to compare stickleback estimates with those of other assessments. However, the estimated stickleback biomass for the Gulf of Bothnia (Jurvelius et al., 1996) was ca. 25 000 t in an area roughly three times the posterior mean estimate of ca. 2000 t in the GoF.

It might have been enlightening to compare the results of this model with those obtained analysing the same data with a different methodology. However, there are two reasons why such a comparison is difficult. First, a conceptually valid comparison could
Bayesian spatial multispecies modelling of pelagic fish stocks

have only been made with another model in a Bayesian framework because the formal results produced by Bayesian and frequency statistics are in response to very different questions (Dennis, 1996; Nielsen and Ley, 2002). Even if we could overlook the conceptual difficulty of the comparison, another more practical problem would arise: the survey design had impacts on the type of conventional methodology that could be applied to the data. Although bootstrapping has been applied before in analyses of hydroacoustic fish-stock assessment data, the method was not applied to the present dataset because, as stated by Simmons and MacLennan (2005), 'the bootstrap does not take into account spatial distribution, and each observation is treated as a sample from a stationary pdf. Thus, if there is spatial correlation and the sampling design is non-random, the estimated variance will be wrong.'

Another problem is that the biomass estimate was highly dependent on a few parameters not inferred from the data. However, because it is common practice to use acoustic-survey data to estimate biomass, they were included in the model simply to show how uncertain the results were even without the uncertainty around all the variables. The distributions of posterior estimates were wide, as was evident in the results. Biomass estimates can be influenced by parameter choices in deterministic equations. Literature estimates of the depth in the common equation $BS = 20 \log L - b$ for Baltic herring vary between $-63.88$ and $-71.9 \text{ dB}$ (Foote, 1987; Peltonen and Balk, 2005; Fässler and Gorska, 2009). The trend has been for each new investigation to produce a higher value than the previous one. Species, fish depth, and orientation are factors that influence backscattering strength. Pelagic species in the GoF migrate diurnally (Peltonen et al., 2004, 2007); the deeper the fish, the smaller the swimbladder. Hence, the model should have a temporal aspect and take into account the observed fish depth. Fässler et al. (2009) present a Bayesian model for estimating TS that takes into account many of the issues listed above.

Another obvious improvement that could be made (although it was found to provide marginal improvement in the current study) could be to calculate the abundance of each length class of each species rather than using mean lengths. Adding more spatial environmental variables in a GIS layer format, such as bottom gradient, salinity, and temperature, should result in better predictive power of the model. In future, we will aim to increase the spatial resolution of the model and further improve computation of the spatial element from a cell to simultaneous point- and area-level calculations.

Acknowledgements
The research leading to these results was supported by the project “Effective Use of Ecosystem and Biological Knowledge in Fisheries” (ECOKNOWS) and the project “Judgement and Knowledge in Fisheries Involving Stakeholders” (JAKFISH), both funded by the European Community’s 7th Framework Programme (FP7/2007-2013) under grant agreements 244706 and 212969, respectively. Additionally, part of the work was supported by the project “Integrated Bayesian risk analysis of ecosystem management in the Gulf of Finland” (IBAM), funded by the European Community’s Seventh Framework Programme under grant agreement 217246 made with the joint Baltic Sea Research and Development Programme BONUS and the Academy of Finland. However, the paper does not necessarily reflect the views of the EC and in no way anticipates the Commission’s future policy. We thank the scientists and crew who contributed to data collection on the research vessel, and the Academy of Finland and the Employment and Economic Development Centre for Southeastern Finland for financial support for data acquisition.

References

Downloaded from https://academic.oup.com/icesjms/article-abstract/69/1/95/673665 by guest on 16 February 2019


