NRC Research Press

ARTICLE

Incorporating stakeholders' knowledge to stock assessment: Central Baltic herring

Samu Mäntyniemi, Päivi Haapasaari, Sakari Kuikka, Raimo Parmanne, Maiju Lehtiniemi, and Joni Kaitaranta

Abstract: We present a method by which the knowledge of stakeholders can be taken into account in stock assessment. The approach consists of a structured interview process followed by quantitative modelling of the answers. The outcome is a set of probability models, each describing the views of different stakeholders. Individual models are then merged to a large model by applying the techniques of Bayesian model averaging, and this model is conditioned on stock assessment data. As a result, the views of interviewed stakeholders have been taken into account and weighed based on how well their views are supported by the observed data. We applied this method to the Baltic Sea herring (*Clupea harengus*) stock assessment by interviewing six stakeholders and conditioning the resulting models on stock assessment data provided by the International Council for the Exploration of the Sea.

Résumé : Nous présentons une méthode permettant de tenir compte, dans l'évaluation des stocks, des connaissances de différents intervenants. L'approche consiste en un processus d'entrevue structurée suivi de la modélisation quantitative des réponses. Le résultat est un ensemble de modèles probabilistes qui décrivent chacun les opinions de différents intervenants. Ces modèles sont ensuite fusionnés individuellement à un modèle plus large en appliquant les techniques de combinaison bayésienne de modèles, et ce modèle plus large est conditionné sur les données d'évaluation des stocks. Ainsi, la prise en compte des opinions des intervenants interviewés est pondérée selon leur degré de coïncidence avec les données d'observation. Nous avons appliqué cette méthode à l'évaluation des stocks de harengs atlantique (*Clupea harengus*) de la mer Baltique en interviewant six intervenants et en conditionnant les modèles en découlant sur les données d'évaluation des stocks fournies par le Conseil International pour l'Exploration de la Mer. [Traduit par la Rédaction]

Introduction

It is widely recognized that the management of fisheries is mostly about management of people involved rather than directly affecting the fish population and its habitat (e.g., Jentoft and McCay 1995; Degnbol et al. 2006; Mackinson et al. 2011). Various studies have suggested that involving stakeholders in the process of decision making would improve the trust between stakeholders and managers and therefore improve the commitment to management, thus reducing the implementation uncertainty of management actions (Nielsen and Vedsmand 1999; Haapasaari et al. 2007; Newig et al. 2005; Lynam et al. 2007). In some cases this can also allow for a bit higher exploitation rates, because reduction of the implementation uncertainty improves the overall predictability of the fishery system (Haapasaari et al. 2007; Haapasaari and Karjalainen 2010; Röckmann et al. 2012).

The importance of involving interested parties in evaluating science related to complex problems and in forming the knowledge base for decision making has been widely emphasized (Ludwig 2001; Voinov and Bousquet 2010). Participation is believed to enrich models and their transparency and to enhance understanding and acceptance of resulting models among stakeholders. It can also facilitate the urgently needed review of the models by stimulating dialogue between all parties involved (Voinov and Bousquet 2010; Schnute and Richards 2001; Prell et al. 2007; Henriksen et al. 2007).

Our focus in this paper is on the method that could be used to explicitly involve the stakeholders in the process of gathering the knowledge about the fishery system. "Stakeholder" usually refers to fishermen, native groups, conservation organizations, and the public. However, in this paper we also include scientists and managers into this set; we use their knowledge in the same way as the knowledge of traditional stakeholders.

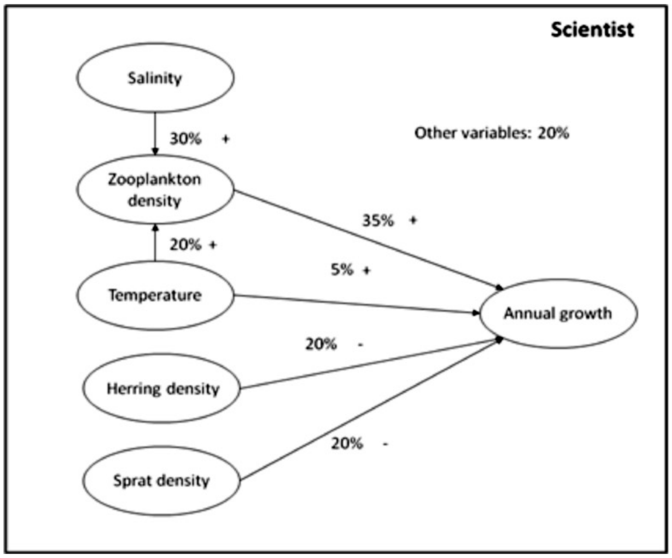
Our approach is simply to view stakeholders as experts and then use methods developed for pooling of expert opinion (Hammond and O'Brien 2001; O'Hagan et al. 2006; Uusitalo et al. 2005). Consequently, the quantification and pooling of the stakeholder views is conducted within the Bayesian approach to scientific reasoning (de Finetti 1975). The core idea of the Bayesian approach is to use the concept of probability as a measure of knowledge (Ramsey 1926; Savage 1954). The relative beliefs on different values of the quantity of interest are expressed as a probability distribution, where probability is interpreted as degree of belief (de Finetti 1975).

Within the Bayesian approach, the views of the individual stakeholders are first elicited and formulated as probability statements (O'Hagan et al. 2006). Then the probability models of stakeholders are viewed as alternative hypotheses about the problem, and these alternatives are initially given equal weights to equally respect the view of each stakeholder. Thus, the resulting metamodel includes the views of all stakeholders. This approach is

Received 14 July 2012. Accepted 6 January 2013.

Paper handled by Associate Editor Ray Hilborn.

S. Mäntyniemi, P. Haapasaari, S. Kuikka, and J. Kaitaranta.* Fisheries and Environmental Management Group (FEM), Department of Environmental Sciences, Viikinkaari 2, P.O. Box 65, FI-00014 University of Helsinki, Finland.
R. Parmanne. Finnish Game and Fisheries Research Institute, P.O. Box 2, FI-00791 Helsinki, Finland.
M. Lehtiniemi. Finnish Environment Institute (SYKE), Marine Research Centre – State of the Marine Environment Unit, Erik Palménin aukio 1, Helsinki, P.O. Box 140, FI-00251 Helsinki, Finland.
Corresponding author: Samu Mäntyniemi (e-mail: samu.mantyniemi@helsinki.fi).
*Present address: Helsinki Commission, Baltic Marine Environment Protection Commission, Katajanokanlaituri 6 B, FI-00160 Helsinki, Finland.



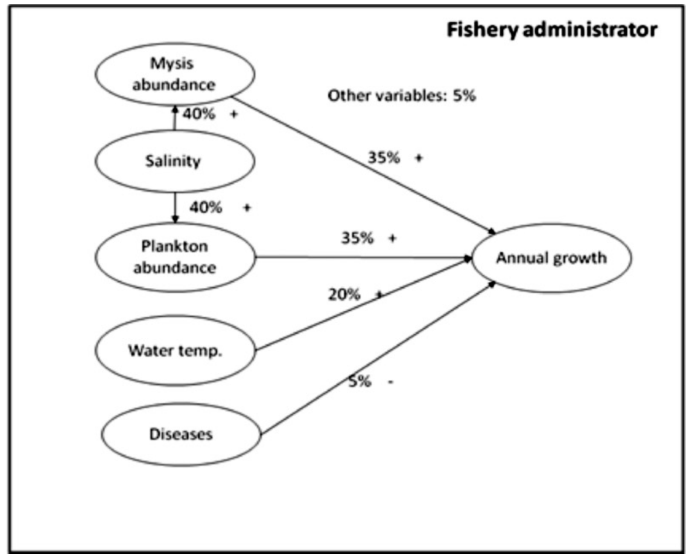


Fig. 1. Directed acyclic graphs (DAGs) representing the views of two stakeholders about the factors affecting the annual growth of Baltic herring. Ovals denote the variables, arrows indicate the direction of the causal relationship, percentages indicate the reduction of variance in dependent variable given that the cause is known, and plus and minus signs denote the direction of the relationship. All models provided by the stakeholders can be found in the online Supplementary Appendix B[1].

explained and illustrated with a simple example by Hammond and O'Brien (2001).

Methods

Bayesian model averaging (BMA) is a consistent method for taking structural uncertainty into account when representing and updating knowledge about a true state of nature using Bayesian statistical methods (Hoeting et al. 1999; Gibbons et al. 2008). Our approach is to cast the problem of pooling the stakeholder knowledge into the Bayesian framework, which has been suggested as a normative approach to environmental decision making (e.g., (Punt and Hilborn 1997; Dorazio and Johnson 2003; Mäntyniemi et al. 2009). As a consequence, competing stakeholder views should logically be weighted using the principles of BMA.

If new information (i.e., observed data) becomes available after formulation of the meta-model, the combined views of the stakeholders can be updated. The update takes place in two levels: (*i*) The views of each stakeholder will be updated according to the logic specified by each stakeholder. (*ii*) The weights of the stakeholders will be updated based on the consistency with the new information. In this way, the stakeholder showing the highest level of expertise about how the system works will gain dominance of the meta-model.

The stakeholder models can be visualized by using directed acyclic graphs (DAGs; Spiegelhalter et al. 1996; Haapasaari et al. 2012). A DAG shows all the included variables and their assumed dependencies and independencies. DAGs can be used for qualitative comparison of the beliefs of stakeholders. For example, Fig. 1 illustrates how two different stakeholders see the factors affecting the growth of Baltic herring (*Clupea harengus*).

Case study: Baltic herring fishery

Within the JAKFISH project, we have used the Central Baltic herring (CBH) fishery as a case study to study and develop the process of participatory modelling. The main idea is to assess the stock by using a stock assessment model specified by each stakeholder and using the International Council for the Exploration of the Sea (ICES) data sets and environmental observations.

ICES Assessments of CBH stocks have indicated that the spawning stock biomass of the stock was high in the 1970s but declined until 2001. The mean mass-at-age of individual herring has decreased considerably, by 15%–45%, since the 1990s, and has stabilized at a low level in recent years (ICES 2009*a*, 2011; CEC 2011). Reasons for the herring stock's poor state are largely unknown, and scientists do not have a single agreed-upon causal structure to describe its biological productivity. According to ICES (2011), the stock is harvested outside of safe biological limits.

In principle we could develop the entire assessment model for each stakeholder from scratch. However, we chose to simplify the task as much as possible for efficient development; potential problems are likely to arise with a simplified model. Our simplification was to construct a basic structure for the population dynamics model (e.g., the age structure), which was given to stakeholders as a fixed, unchangeable structure. Other parts such as the source of mortality were subject to definition based on different hypotheses by various stakeholders.

The population dynamics model is an age–structured, size-based stochastic model formulated as annual transition equations. This model is documented in detail in elsewhere (S. Mäntyniemi, L. Uusitalo, H. Peltonen, P. Haapasaari, and S. Kuikka, unpublished data).

The three main parameters of the annual transition are the growth rate of individuals, their natural mortality, and the survival of spawned eggs. This new integrated assessment framework was developed because the assessment method used by ICES does not treat any of these variables as uncertain and because existing assessment frameworks such as Stock Synthesis (Methot and Wetzel 2012) and CASAL (Bull et al. 2002) are not able to take into account the model uncertainty and uncertainty about the survival process variance (Mäntyniemi et al 2012).

The interview process

The views of stakeholders were used to specify a model about the key factors that affect these three parameters. For each of the three parameters, the stakeholders were asked to identify the five most important explanatory variables. In the second stage, stakeholders

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2012-0316.

stated their opinions about how strongly and to what direction each of the explanatory variables would affect the parameter of interest. In the third stage, the stakeholders were asked to specify how uncertain they were about the strengths of the effects.

In this case, the purpose of the interview was to extract the current, existing knowledge of the stakeholders. Therefore, the topic of the interview was only vaguely introduced to the stakeholder before the interview. This was to prevent them updating their knowledge about the particular questions before the interview. The structure of the interview was as follows:

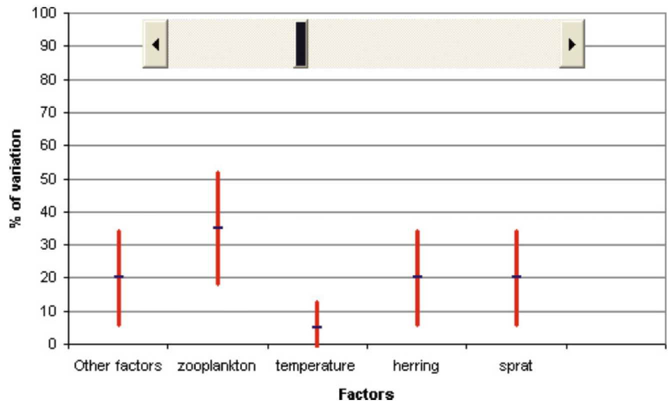
(1) Introduction. At this phase the purpose and the scope of the interview was presented. The structure of the interview was presented, and the basic principles of constructing a DAG were discussed with the help of a simple example.
(2) Definition of the parameter of interest. Here the structure of the population dynamics model was explained, and the parameter of interest was defined. The interview was repeated for each of the three parameters.
(3) 1st round: five most important predictors. The stakeholder was asked to identify up to five most important explanatory factors that would help to predict the parameter of interest.
(4) 2nd round: possibility to change the group of five. The purpose of this stage was to let the stakeholder reiterate the chosen factors. At this stage, the stakeholder was shown a list of other potential explanatory factors. The list was compiled beforehand by sending an email questionnaire to several herring experts, who were asked to list potential explanatory factors. The stakeholder was free to change some or all of originally indicated factors or to keep the list of variables unchanged.
(5) 3rd round: directions of causal effects and effect strengths. First, the arrows were drawn to indicate the causalities. Next, the strengths of the causalities were asked in terms of coefficient of determination (i.e., how much of the predictive variance of the variable of interest would be reduced if the explanatory factor was known, or put another way, how much of the variation is explained by each factor). Finally the stakeholder was asked to indicate whether the causal relationship is positive or negative, for example, whether an increase in cod density is expected to increase or decrease the natural mortality of herring. Stakeholders were also given possibility to state if they believe that the relationship is dome-shaped instead of monotonic (see online Supplement[1]).
(6) Uncertainty about effect strengths. Finally the stakeholder was asked to express her or his uncertainty about the strengths of the causal effects. This task was aided by interactive graphics; the uncertainty about strength of causal effect was presented as 95% probability intervals, which could be made longer or shorter by dragging a slider (Fig. 2). For simplicity and mathematical convenience, the strengths of all effects were assessed simultaneously by assessing the amount of relative variation (Appendix A).

Effect strengths and uncertainty

Turning the results of our interview into a complete probability model requires some additional steps that will depend on the choices of the analyst. The stakeholder has provided the strength and direction of the effect and uncertainty about the strength. It is then the responsibility of the analyst to decide about

(1) prior distribution for the marginal variance of the parameter of interest,
(2) prior distributions for the marginal variances of all explanatory factors,
(3) prior distribution for the mean of the parameter of interest,
(4) prior distributions for the means of all explanatory factors,
(5) the shape of the distribution of the annual variation of the parameter of interest,

Fig. 2. Interactive graph used in the interview process to elicit the uncertainty about the strength of effects. The bars show the 95% probability interval for the strength of each explanatory factor. The widths of intervals are proportional to the point estimates given in a DAG. Dragging the slider left and right increases and decreases the uncertainty, respectively. During the interview, the uncertainty was initially set to highest value and was then gradually adjusted until the stakeholder was happy about the uncertainty.



(6) the shape of the distribution of the annual variation of the explanatory variables, and
(7) the exact functional form of the dependency between the parameter of interest and the explanatory factors.

Our approach here is to use minimally informative priors and then apply suitable transformations to be able to present the probability model as a regression model with normally distributed responses and predictors. Both the responses and the predictors chosen by stakeholders were positive-valued. Therefore log-transformation was applied to all variables prior to modelling. Even though stakeholders specified dome-shaped relationship between some of the variables, these were not taken into account at this stage of the model development.

Under such linear regression, the strengths of the effects can be translated to regression coefficients. Thus, the uncertainty about the effect strengths defines the prior distribution of the regression coefficients together with the prior distributions assigned for the means and variances. See Appendix A for details.

By definition, the effect strengths and the proportion of unexplained variation need to sum up to 100%. Consequently, the prior distribution assigned to the proportions must obey this constraint. Our approach is to use the Dirichlet distribution as a prior for the proportions. The parameters of the prior distribution are directly obtained from the point estimates of the effect strengths and from the amount of uncertainty specified using the interactive graph (Appendix A). The slider in the graph (Fig. 2) is used to adjust the sum of the Dirichlet parameters.

Another way of looking at the model is to view it as a multivariate normal distribution conditional on the covariance matrix specified by the variances and the effect strengths. Details are presented in Appendix A.

The collection of data about the predictor variables was prioritized by calculating the sum of link strengths over all the stakeholder models shown below. Based on this criterion, eight most important variables were zooplankton density, cod abundance, water temperature, sprat abundance, herring abundance, harmful substances in water, bottom vegetation, and salinity. From these, herring abundance and harmful substances had to be left out of the analysis. Harmful substances was a too vague concept that would have needed more elaboration with the stakeholders. Herring abundance was originally represented by corresponding variable in the population dynamics model, but implementing it

Table 1. Explanatory variables and the posterior weights of the six stakeholders (SH) according to Bayesian model averaging using the International Council for the Exploration of the Sea (ICES) stock assessment data.

SH	Growth	Prob.	Nat. mortality	Prob.	Egg surv.	Prob.
1	Zoopl, Temp, Salinity	0.0011	Zoopl, Cod	0.0004	Temp, Secchi	0.0042
2	Zoopl, Salinity, Sprat	0.0086	Cod	0.0001	Temp, Cod, Sprat, Salinity	0.0236
3	Zoopl, Sprat	0.0202	Cod	0.0070	Temp	0.0350
4	Zoopl, Temp	0.0001	Cod	0.9890	Zoopl, Secchi	0.0014
5	Zoopl, Temp, Salinity	0.0170	Cod, Zoopl	0.0031	Cod	0.0018
6	Zoopl, Temp, Salinity, Sprat	0.9530	Cod, Zoopl	0.0004	Temp, Cod, Zoopl	0.9340

Note: Weights are estimated for each biological process separately, which enables a stakeholder to have a high weight in one process and lower weight in other processes.

as a covariate in the model resulted in very low speed of Markov chain Monte Carlo (MCMC) simulation. No direct data was available on bottom vegetation, but Secchi depth observations were used a proxy for the amount of bottom vegetation available.

Zooplankton monitoring was conducted by the Finnish Environment Institute (previously the Finnish Institute of Marine Research), starting from 1979 on several open sea stations as part of the HELCOM COMBINE monitoring program. Samples were collected annually in August by vertical tows of a 100 μm WP-2 net and analyzed to the species level if possible. For the data analyses, the biomasses of the dominant species of copepods and cladocerans in the Gulf of Finland and the northern Baltic Proper were taken into account.

Sprat (*Sprattus sprattus*) and Atlantic cod (*Gadus morhua*) biomass estimates were extracted from ICES (2009*b*), salinity values are from HELCOM monitoring data, and the sea surface temperature time series was provided by the Swedish Meteorological and Hydrological Institute.

Computational approach: how to obtain the model probabilities?

After completing the model structure as described above, the model can be fitted to stock assessment data. This can be done by directly coupling the stakeholder model specifications with the population dynamics model presented in Mäntyniemi et al. (2012).

In this study we used the method of Carlin and Chib (1995). The basic idea of the method is to express the meta-model as a finite mixture of alternative models by using one or more indicator variables to index the potentially correct model configurations and then use MCMC simulation (Gilks et al. 1996) to approximate the joint posterior distribution of the indicator variables. Four uncertain indicator variables were set up to allow (a priori) independent uncertainty about different parts of the population dynamics. These were (*i*) the shape of the stock–recruitment function (Beverton–Holt versus Ricker) and stakeholder beliefs about the explanatory factors affecting (*ii*) growth of herring, (*iii*) natural mortality, and (*iv*) survival of eggs to recruits. Thus, the number of potentially correct model configurations was $2 \times 6 \times 6 \times 6 = 432$. However, for the purposes of evaluating the weights of stakeholder knowledge, it is mostly of interest to look at the weights given to each stakeholder within each subject area. For example, we are looking at the probability that a fishery manager would be "right" about the factors affecting the growth of herring without assuming that any particular stakeholder would be "right" about recruitment and natural mortality.

Results

The resulting probabilities, or weights, are given in Tables 1 and 2. When stakeholders (SH1–SH6) are weighted separately for each of the biological processes, SH6 gets the highest weight both in growth and recruitment, but has a very low weight in natural mortality. In other words, SH6 was clearly the best in predicting the stock assessment data with his models about growth and recruitment, but his model for natural mortality was very unsuccessful.

Table 2. Explanatory variables and the posterior weights of the six stakeholders (SH) according to Bayesian model averaging using the International Council for the Exploration of the Sea (ICES) stock assessment data.

SH	Growth	Natural mortality	Egg survival	Probability
1	Zoopl, Temp, Salinity	Zoopl, Cod	Temp, Secchi	0.0017
2	Zoopl, Salinity, Sprat	Cod	Temp, Cod, Sprat, Salinity	0.0015
3	Zoopl, Sprat	Cod	Temp	0.9800
4	Zoopl, Temp	Cod	Zoopl, Secchi	0.0007
5	Zoopl, Temp, Salinity	Cod, Zoopl	Cod	0.0126
6	Zoopl, Temp, Salinity, Sprat	Cod, Zoopl	Temp, Cod, Zoopl	0.0035

Note: Weights are estimated by assuming that the models of the same stakeholder would be used for each of the biological processes (i.e., SH2 would be "right" about growth and natural mortality and recruitment).

The most credible model for the natural mortality is the one specified by SH4, which gets almost all of the weight. Interestingly, all stakeholders have agreed that the biomass of cod should be used to predict the natural mortality, but still SH4 has much higher weight than the others. This reflects the fact that SH4 had predicted that only a small fraction of the variation in natural mortality could be explained by the cod biomass, whereas the other stakeholders had their prior beliefs on stronger influence. The estimates of natural mortality compared with the cod biomass do not show a strong covariation, which is then reflected by the weights given to stakeholders.

As an extreme case, stakeholder weights were also estimated assuming that a single stakeholder would have the most plausible model about all of the biological processes. For example, models specified for growth, mortality, and recruitment by SH2 were assumed to be "true" and all other models were assumed "false". Under such an assumption, SH3 gets almost all of the weight (Table 2). This seems logical since SH3 is the second best in each biological process when they are assessed separately (Table 1).

The extreme case where a single stakeholder is assumed to have the correct model about all of the biological processes was also used to study the differences in the stakeholders' perception about the past development of the stock. As is obvious, each stakeholder remained uncertain about the past development. For example, SH1 and SH6 had slightly different views about how the natural and fishing mortalities have varied in the past (Fig. 3).

Stakeholders 1–5 had quite similar estimates about the past development of the spawning stock and recruitment. These estimates led to strong favouring of the Ricker stock–recruitment function (Fig. 4, left panel). However, SH6 estimated that during the periods of higher egg production, the expected recruitment might have been higher, which then leads to higher uncertainty about the shape of the stock–recruitment function (Fig. 4 right panel).

Fig. 3. Estimated past and future variations of fishing and natural mortality according to two stakeholders. Upper panels show the estimates of mean fishing mortality in age groups 3–7, and the lower panels show the development of natural mortality at age 5. Bold lines connect the annual posterior medians, and transparent lines depict both the uncertainty of the annual estimate and uncertainty about the annual variation of the mortalities. Transparent lines are random draws from the joint posterior distribution of annual mortalities. Blue denotes the past and red denotes the future time series predicted assuming constant effort at the level of 2007 and predicting the explanatory variables (cod biomass and zooplankton) using a mean reverting red noise (e.g., Ruokolainen et al. 2009) model.

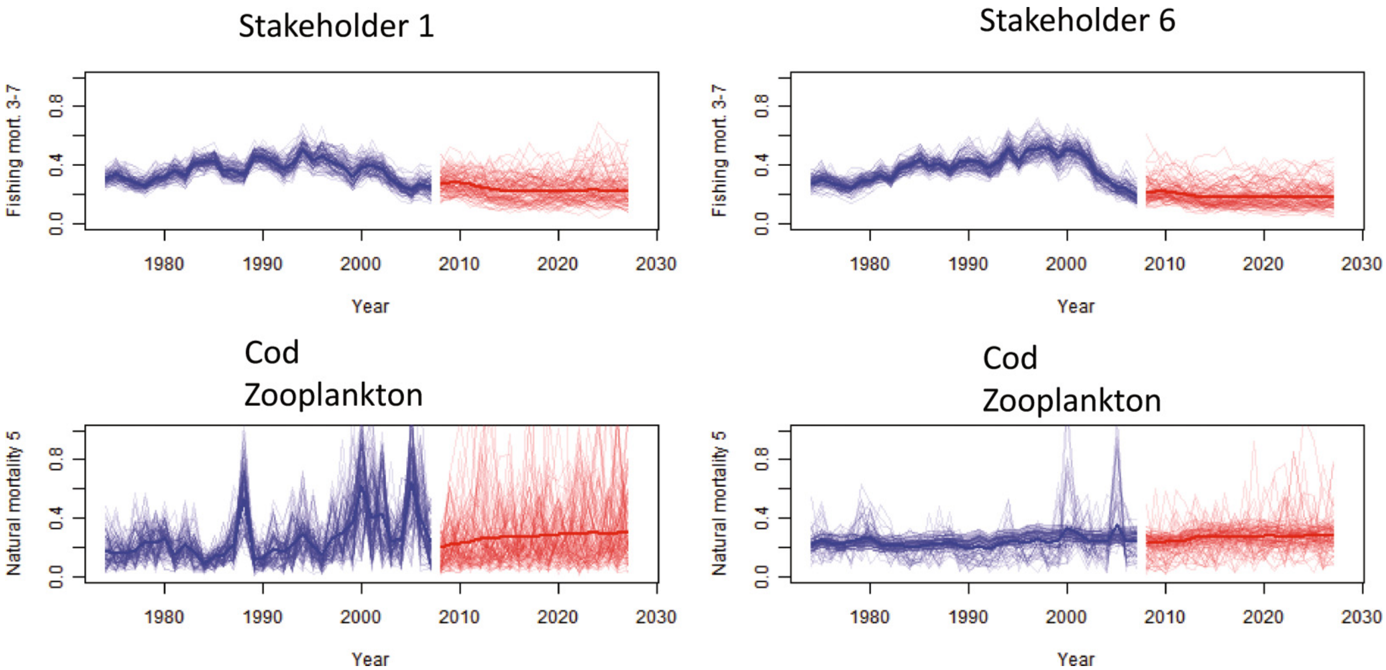
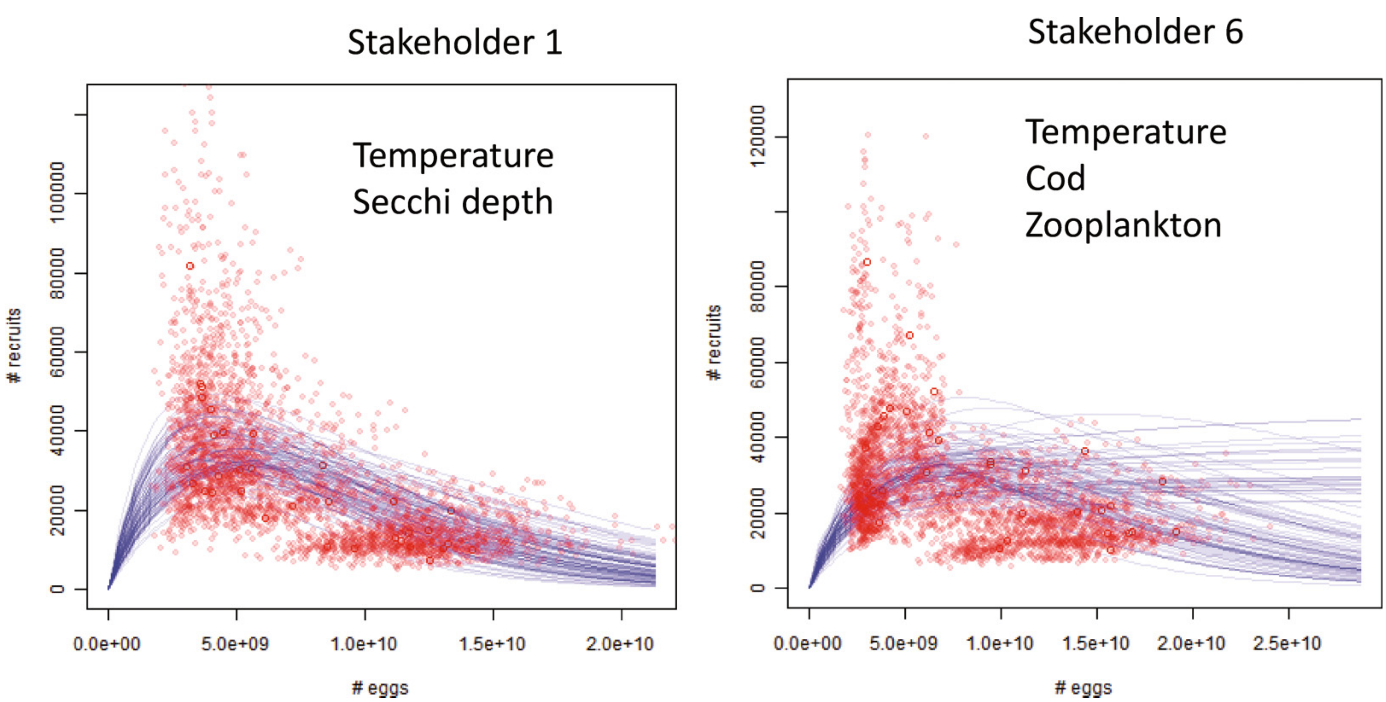
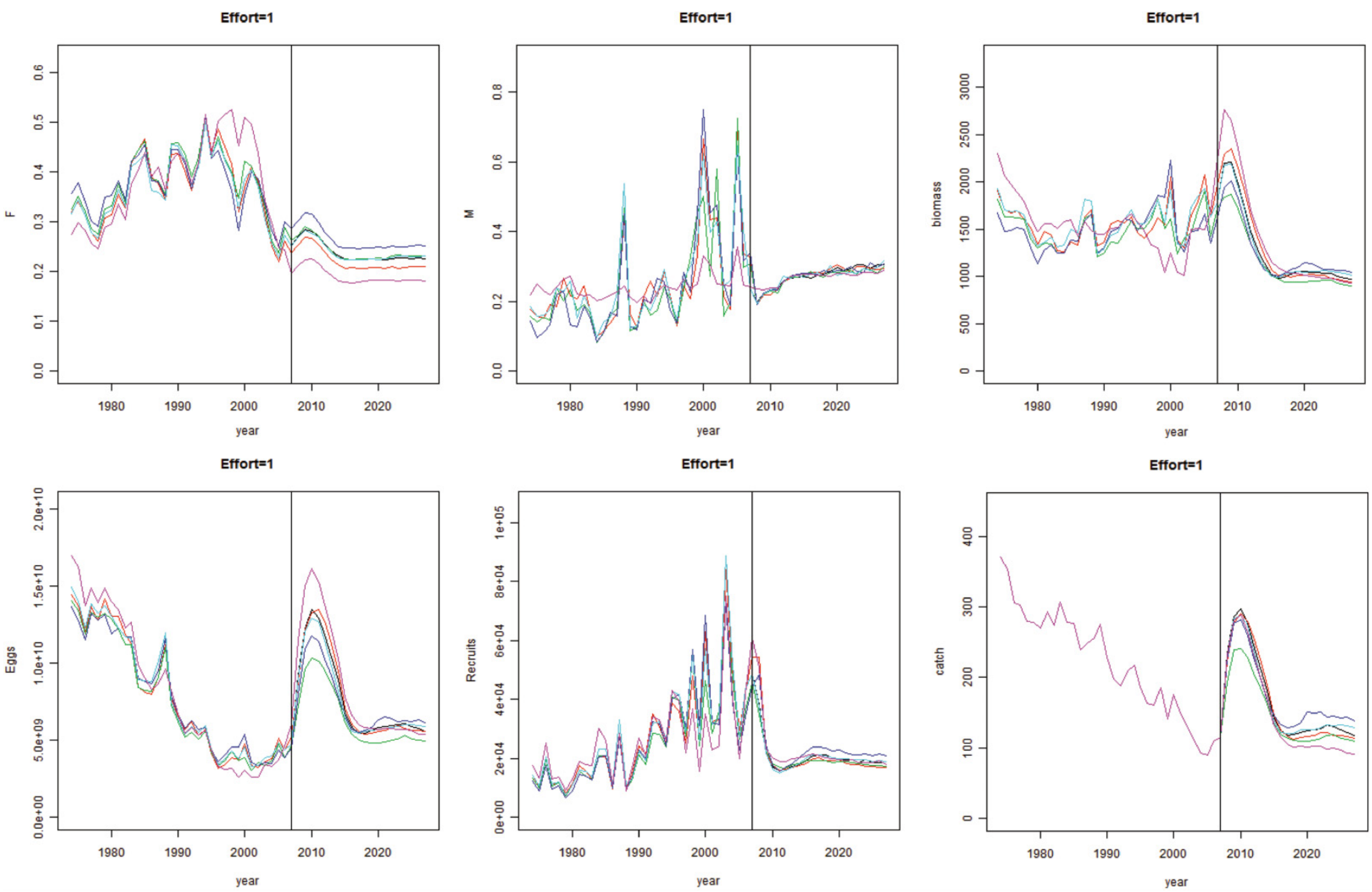


Fig. 4. Stock–recruitment relationship of Central Baltic herring according to two stakeholders. Blue lines depict the uncertainty about the expected recruitment from a given number of eggs, and red dots represent the uncertainty about the historical pairs of spawning stock and resulting recruitment. The predictors chosen by stakeholders are listed in each panel.



Large differences cannot be seen in the views about the future of the stock between stakeholders, when fishing effort is assumed to stay at the same level as 2007. Stakeholders expect that the fishing mortality would settle around 0.25, with some difference between stakeholders (Fig. 5). Approximately the same relative differences hold if the fishing effort was doubled or halved or if the fishery was closed. According to all stakeholders, the stock size has remained quite stable around 1.5 million tonnes, but the number of spawned eggs has been steadily going down to about one-third of the numbers compared with the beginning of the

Fig. 5. Posterior expected values of past and future estimates of the development of the fishery assuming that the fishing effort would remain in the level of 2007 for the next 20 years. Each colour represents one stakeholder.



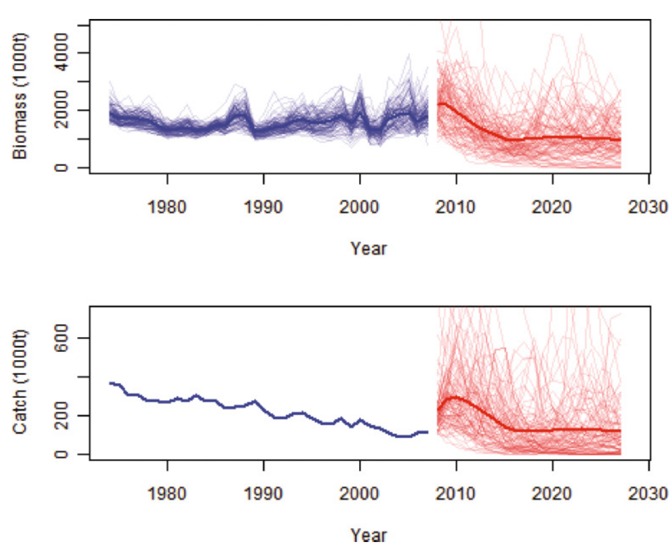
assessment period. This reflects the change in individual growth rates and consequently in the age structure and total fecundity of the population. However, an important aspect to be considered is the uncertainty about the future. Each stakeholder had a high uncertainty about the future, as exemplified in Fig. 6; uncertainty about the upcoming catch and biomass is very high. The realized catches for the two first prediction years show the increasing trend predicted by the stakeholder, but are located in the lower ends of the prediction intervals. Comparison of Figs. 5 and 6 shows that the predictive distributions of the stakeholders are overlapping a lot with high uncertainty in each.

Discussion

We have shown how BMA can be used to account for the knowledge of stakeholders in fisheries stock assessment. We used a fairly generic state-space model (Mäntyniemi et al. 2012) to describe the core dynamics of the herring stock and then extended the model to include multiple competing hypotheses about the external factors that might affect the dynamics. These hypotheses were specified by volunteer stakeholders with the help of a modelling specialist.

The results of stakeholder pooling showed that while one of the stakeholders was in this case estimated to have a high weight, either as a single source of knowledge or when analysed separately for each biological process, the resulting beliefs about the future development of the stock based on estimates from different stakeholders were not dramatically different. There are a couple of points contributing to this result.

Fig. 6. Predicted and past stock biomass and catches according to stakeholder 1. Bold lines connect the annual posterior medians, and transparent lines depict both the uncertainty of the annual estimate and uncertainty about the annual variation of the mortalities. Transparent lines are random draws from the joint posterior distribution of annual mortalities. Blue denotes the past and red denotes the future time series.



Perhaps the most obvious one is that by the design of our study, the stakeholders were bound to interpret the observed data in the same way. This comes from the fact that the population dynamics model and associated observation models were assumed to be the same for each stakeholder, and differences were only allowed in external factors affecting the dynamics. Thus, once shown the data, the Bayes' rule updates the prior distributions provided by each stakeholder with the same likelihood function. This makes the posterior distributions of parameters specified by each stakeholder to be closer to each other than their prior distributions were. This is a generic phenomenon, expected to some extent to occur in all cases where BMA is going to be used to pool expert–stakeholder knowledge by using the same observation models for each participant. It follows directly from the fact that the posterior distribution is proportional to the prior distribution and the likelihood function.

In this case study, some stakeholders questioned the validity of the data used. This is an indication that the observation models should really be made stakeholder-specific. An interesting line of further research would then be the development of interview techniques by which the observation models could be elicited from stakeholders.

Having different observation models for each stakeholder would make it possible that after seeing the data, the beliefs of the stakeholders would divert even more than in the beginning. Thus, BMA could also be used to explicitly model the process of disagreement that might arise when new data accumulate.

Another point explaining the small differences in future projections between stakeholders is that all of the stakeholders were quite uncertain about the strengths of the effects that the external factors might have. As always in the Bayesian analysis, this has the consequence that even relatively weak signals coming from the data will be taken into account by the stakeholders so that their beliefs would be easily updated to be close to each other. Highly confident prior opinions would have stayed further apart.

Even though the posterior opinions of the stakeholders were quite similar, there was always a single stakeholder that happened to collect almost all the weight in BMA. This indicates that the one obtaining the highest weight was already consistent with the data before “seeing” it, while the others having smaller weight were much less capable of predicting the observations before seeing them. However, because of high prior uncertainty, the stakeholders with small posterior weight still learnt quickly from the data to agree more with the one who was already predicting the data well. Such a dominance of one stakeholder means that the observed data was interpreted to be informative enough to reduce the structural uncertainty that arose from the inclusion of multiple competing hypotheses. Some of our stakeholders saw this process as a democratic way of accounting for their knowledge in stock assessment, but those stakeholders that questioned interpretation of the data set (i.e., the likelihood function) were more skeptical about the validity of the approach (Haapasaari et al. 2013). This again points towards the need to use the stakeholder knowledge in the formulation of the observation models if the modelling approach is expected to increase the acceptance of the results by the stakeholders.

Formulating the stakeholder views as a mixture of multivariate normal distributions simplifies the modelling task and increases the possibility to take the stakeholder views into account in practice. However, such a simplification naturally reduces the chance to account for relationships that are difficult to linearize by using simple transformations. Here we assumed linear relationships between log-transformed variables and did not account for the possibility of nonmonotonicity suggested by some of the stakeholders. A future line of research could explore the use of nonparametric Gaussian process priors (e.g., Munch et al. 2005) in the expert elicitation of function shapes. Another more flexible alternative in terms of distributional assumptions would be to discretize the variables and then use the interview to ask conditional probability tables to describe the stakeholder beliefs about the causal relationships of the variables (e.g., Uusitalo et al. 2005; Uusitalo 2007). However, this would again considerably increase the time needed for the interview.

It is also worth noting that the approach used here results in a mixture of stakeholder views and the views of the analyst. The variables to be used and statements about their relationships come from the stakeholders, but the rest of the structure is dependant on the analyst. This balance could be changed by increasing the time to be used for interviewing the stakeholders. The interviews for the three parameters of interest lasted from 2 to 4 hours in total. In some cases it was evident that the interviewee got tired of thinking, especially about the uncertainty in the effect strength, towards the end of the interview. This suggests that if priors for means and variances were to be asked from the stakeholders, the interview should be divided to multiple sessions.

Elicitation of expert knowledge about uncertain quantities is not free of problems and is a wide field of research on its own. The structure of the interview, the way the task is described, and the questions asked can all affect the outcome of the elicitation (O'Hagan et al. 2006). Evaluating the success of the elicitation is particularly difficult, because the true value of the probability asked from the expert does not exist in a physically quantifiable form (Nau 2001) but is a description of the expert's thinking (Ramsey 1926) and often becomes formed when asked for (O'Hagan et al. 2006). These problems can be somewhat mitigated by providing feedback to the expert during the elicitation process and also by trying to ask the same things in different ways (O'Hagan et al. 2006). The effect of potential elicitation errors can be studied by conducting sensitivity analysis by repeating the analysis with small changes in prior distributions (O'Hagan et al. 2006).

The meta-model (updated or not) could be used to derive management advice once the alternative management actions and the objectives of the management have been explicitly coupled with the model. Such an influence diagram (Pearl 1988) can then be used for decision analysis (i.e., to find management actions that would produce highest expected utility), such as highest expected catch or profit or some other measure of success. The same optimization can be performed for each stakeholder model separately. This would provide material for assessing the sensitivity of management actions to the differences in stakeholder views.

The idea of using stakeholder knowledge in decision analysis has naturally raised doubtful voices. For example, Dennis (1996) suspected that such a participation would provide the stakeholders an opportunity to “play the game” for their own benefit by providing suitable misinformation. Ideally, the decision-making process would be arranged so that it would provide incentives to reveal true objectives and honest knowledge, but this topic goes outside the scope of this paper. However, using the BMA approach to weight the models of individual stakeholders can provide a buffer against potential attempts to provide misinformation. Models that are not in agreement with observed data will get less weight in the model averaging process (Hammond and O'Brien 2001; Hoeting et al. 1999).

It should be noted that while weighting of the stakeholder knowledge could be justified by using their ability to fit to the observed data, the same may not be true for the weighting of the objectives of the stakeholders in the decision analysis. Instead, the amount of weight given to objectives of each stakeholder is a political question for which the logic of scientific learning probably does not apply.

A potential problem in all statistical stock assessments is the violation of the likelihood principle (Berger and Wolpert 1988). The likelihood principle entails that all the interpretation of observed data should enter the inference only through the likelihood function (Berger and Wolpert 1988). Typical violations of the principle arise when the analyst chooses the structure of the likelihood function itself based on exploratory analysis of the same data set or based on the anticipated number of parameters that can be identified based on the data set. In Bayesian stock assessment, the likelihood principle can also get violated if the prior distribution assigned to model parameters is based on the same data that will be used in the likelihood function. In cases where the analysis is started after the data set has been collected, this possibility cannot be completely eliminated. However, stakeholders can be instructed that they should consciously try to avoid using their interpretation of the data set in their probability statements while encouraging the use of their experience and data from other similar cases and general knowledge about biology.

Acknowledgements

We are grateful to the six volunteer stakeholders who participated in our research. We also thank the three anonymous referees whose comments greatly helped in improving the quality of this manuscript. This study was carried out with financial support from the Commission of the European Communities, "JAKFISH: Judgement and knowledge in fisheries involving stakeholders" (grant agreement No. 212969). This paper does not necessarily reflect views of the Commission and in no way anticipates the Commission's future policy in the area.

References

Berger, J.O., and Wolpert, R.L. 1988. The likelihood principle. The Institute of Mathematical Statistics, Haywood, Calif.

Bull, B., Francis, R., Dunn, A., and Gilbert, D.J. 2002. CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v1.02.2002/10/21.NIWA Technical Report 117, Wellington.

Carlin, B.P., and Chib, S. 1995. Bayesian model choice via Markov-chain Monte-Carlo methods. J. R. Stat. Soc. **57**(3): 473–484.

CEC. 2011. Commission non-paper: On the establishment of a multiannual plan for pelagic stocks in the Baltic Sea and the fisheries exploiting those stocks. Commission of the European Communities, Brussels, Belgium.

de Finetti, B. 1975. Theory of probability. John Wiley & Sons, Bristol.

Degnbol, P., Gislason, H., Hanna, S., Jentoft, S., Raakjær Nielsen, J., Sverdrup-Jensen, S., and Clyde Wilson, D. 2006. Painting the floor with a hammer: technical fixes in fisheries management. Mar. Pol. **30**(5): 534–543. doi:10.1016/j.marpol.2005.07.002.

Dennis, B. 1996. Discussion: Should ecologists become Bayesians? Ecol. Appl. **6**(4): 1095–1103. doi:10.2307/2269594.

Dorazio, R.M., and Johnson, F.A. 2003. Bayesian inference and decision theory — a framework for decision making in natural resource management. Ecol. Appl. **13**: 556–563. doi:10.1890/1051-0761(2003)013[0556:BIADTA]2.0.CO;2.

Gibbons, J.M., Cox, G.M., Wood, A.T.A., Craigon, J., Ramsden, S.J., Tarsitano, D., and Crout, N.M.J. 2008. Applying Bayesian model averaging to mechanistic models: an example and comparison of methods. Environmental Modelling and Software, **23**(8): 973–985. doi:10.1016/j.envsoft.2007.11.008.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. 1996. Markov chain Monte Carlo in practice. Chapman & Hall, London.

Haapasaari, P., and Karjalainen, T.P. 2010. Formalizing expert knowledge to compare alternative management plans: sociological perspective to the future management of Baltic salmon stocks. Mar. Pol. **34**(3): 477–486. doi:10.1016/j.marpol.2009.10.002.

Haapasaari, P., Michielsens, C.G.J., Karjalainen, T.P., Reinikainen, K., and Kuikka, S. 2007. Management measures and fishers' commitment to sustainable exploitation: a case study of Atlantic salmon fisheries in the Baltic Sea. ICES J. Mar. Sci. **64**: 825–833. doi:10.1093/icesjms/fsm002.

Haapasaari, P., Mäntyniemi, S., and Kuikka, S. 2012. Baltic herring fisheries management: stakeholder views to frame the problem. Ecol. Soc. **17**(3). doi:10.5751/ES-04907-170336.

Haapasaari, P., Mäntyniemi, S., and Kuikka, S. 2013. Involving stakeholders in build integrated fisheries models using Bayesian methods. Environmental Management. [In press.]

Hammond, T.R., and O'Brien, C.M. 2001. An application of the Bayesian approach to stock assessment model uncertainty. ICES J. Mar. Sci. **58**(3): 648–656. doi:10.1006/jmsc.2001.1051.

Henriksen, H.J., Rasmusssen, P., Brandt, G., Bulow, D.v., and Jensen, F.V. 2007. Bayesian networks as a participatory modelling tool for groundwater protection. *In* Topics on system analysis and integrated water resources management. *Edited by* A. Castelletti and R. Soncini-Sessa. Elsevier, Oxford. pp. 49–72.

Hoeting, J., Madigan, D., Raftery, A., and Volinsky, C. 1999. Bayesian model averaging: a tutorial. Stat. Sci. **14**: 382–417. doi:10.1214/ss/1009212519.

ICES. 2009*a*. 8.3.3.1 multi-annual management of pelagic fish stocks in the Baltic. *In* Advice 2009, book 8. *Edited by* ICES, Copenhagen, Denmark.

ICES. 2009*b*. Report of the Baltic Fisheries Assessment Working Group (WGBFAS). ICES CM 2009\ACOM:07. ICES, ICES HQ, Copenhagen, Denmark.

ICES. 2011. 8.4.4 ecoregion stock: Baltic sea: herring in subdivisions 25–29 and 32 (excluding gulf of riga herring). *In* Advice 2011, book 8. *Edited by* ICES, Copenhagen, Denmark.

Jentoft, S., and McCay, B. 1995. User participation in fisheries management: lessons drawn from international experiences. Mar. Pol. **19**(3): 227–246. doi:10.1016/0308-597X(94)00010-P.

Ludwig, D. 2001. The era of management is over. Ecosystems, **4**(8): 758–764. doi:10.1007/s10021-001-0044-x.

Lynam, T., De Jong, W., Sheil, D., Kusumanto, T., and Evans, K. 2007. A review of tools for incorporating community knowledge, preferences, and values into decision making in natural resources management. Ecol. Soc. **12**(1).

Mackinson, S., Wilson, D.C., Galiay, P., and Deas, B. 2011. Engaging stakeholders in fisheries and marine research. Mar. Pol. **35**(1): 7–7. doi:10.1016/j.marpol.2010.07.003.

Mäntyniemi, S., Kuikka, S., Rahikainen, M., Kell, L.T., and Kaitala, V. 2009. The value of information in fisheries management: North Sea herring as an example. ICES J. Mar. Sci. **66**(10): 2278–2283. doi:10.1093/icesjms/fsp206.

Methot, R.D., Jr., and Wetzel, C.R. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. [In press.] doi:10.1016/j.fishres.2012.10.012.

Munch, S.B., Kottas, A., and Mangel, M. 2005. Bayesian nonparametric analysis of stock–recruitment relationships. Can. J. Fish. Aquat. Sci. **62**(8): 1808–1821. doi:10.1139/f05-073.

Nau, R.F. 2001. De Finetti was right: probability does not exist. Theory and Decision, **51**: 89–124. doi:10.1023/A:1015525808214.

Newig, J., Pahl-Wostl, C., and Sigel, K. 2005. The role of public participation in managing uncertainty in the implementation of the water framework directive. Eur. Environ. **15**(6): 333–343. doi:10.1002/eet.398.

Nielsen, J.R., and Vedsmand, T. 1999. User participation and institutional change in fisheries management: a viable alternative to the failures of 'top-down' driven control? Ocean Coast. Manage. **42**(1): 19–37. doi:10.1016/S0964-5691(98)00085-4.

O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., and Rakow, T. 2006. Uncertain judgements: eliciting experts' probabilities. John Wiley & Sons, West Sussex.

Pearl, J. 1988. Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers, Inc., San Mateo, Calif.

Prell, C., Hubacek, K., Reed, M., Quinn, C., Jin, N., Holden, J., Burt, T., Kirby, M., and Sendzimir, J. 2007. If you have a hammer everything looks like a nail: traditional versus participatory model building. Interdiscip. Sci. Rev. **32**(3): 263–282. doi:10.1179/030801807X211720.

Punt, A.E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: The bayesian approach. Rev. Fish Biol. Fish. **7**: 35–63. doi:10.1023/A:1018419207494.

Ramsey, F.P. 1926. Truth and probability. *In* The foundations of mathematics and other logical essays, by Frank Plumpton Ramsey. *Edited by* R.B. Braithwaite. Hartcourt, Brace and Co., New York. pp. 156–198.

Röckmann, C., Ulrich, C., Dreyer, M., Bell, E., Borodzicz, E., Haapasaari, P., Hauge, K.H., Howell, D., Mäntyniemi, S., Miller, D., Tserpes, G., and Pastoors, M. 2012. The added value of participatory modelling in fisheries management — what has been learnt? Mar. Pol. **36**(5): 1072–1085. doi:10.1016/j.marpol.2012.02.027.

Ruokolainen, L., Lindén, A., Kaitala, V., and Fowler, M.S. 2009. Ecological and evolutionary dynamics under coloured environmental variation. Elsevier Science Publishers.

Savage, L.J. 1954. Foundations of statistics. John Wiley & Sons, New York.

Schnute, J.T., and Richards, L.J. 2001. Use and abuse of fishery models. Can. J. Fish. Aquat. Sci. **58**(1): 10–17. doi:10.1139/f00-150.

Spiegelhalter, D.J., Thomas, A., and Best, N.G. 1996. Computation on bayesian graphical models. *In* Bayesian statistics. *Edited by* J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith. Oxford University Press, New York.

Uusitalo, L. 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecol. Model. **203**: 312–318. doi:10.1016/j.ecolmodel.2006.11.033.

Uusitalo, L., Kuikka, S., and Romakkaniemi, A. 2005. Estimation of Atlantic salmon smolt carrying capacity of rivers using expert knowledge. ICES J. Mar. Sci. **62**: 708–722. doi:10.1016/j.icesjms.2005.02.005.

Voinov, A., and Bousquet, F. 2010. Modelling with stakeholders. Environ. Model. Softw. **25**(11): 1268–1281. doi:10.1016/j.envsoft.2010.03.007.

Appendix A

Converting effect strengths to priors for regression coefficients

This section uses a simple example to illustrate how the strengths of effects specified in a stakeholder interview can be used to derive prior distributions for regression coefficients. Assume that a stakeholder has specified how strongly variables B and C affect the variable of interest, A. This can be presented as a DAG in the form shown in Fig. A1. Thus, the stakeholder has specified the effect of B on A ($R^2_{B,A} = 0.3$) and the effect of C on

A ($R^2_{C,A} = 0.2$). The stakeholder has also specified the directions of the effects, which are denoted here as $i_{A,B} = -1$ and $i_{A,C} = +1$. Then assume that the analyst uses information from literature (or from another interview) to come up with prior distributions for the means and variances of the variables A, B, and C. In the following, the mean and the variance of A will be denoted as μ_A and σ^2_A, respectively. Similar notation is used for variables B and C.

The linear regression model used here takes the form

$$A = \alpha + \beta B + \gamma C + \epsilon$$

where parameter $\boldsymbol{\epsilon}$ represents random noise and the effects of variables that are not included in the model. This parameter is assumed to have zero mean and variance σ^2_ϵ. Since the stakeholder has assessed that B and C are a priori independent, the variance of A can be presented as

$$\sigma^2_A = \beta^2\sigma^2_B + \gamma^2\sigma^2_C + \sigma^2_\epsilon$$

The strengths of effects, seen as the coefficient of determination, are given by

$$R^2_{A,B} = \frac{\beta^2\sigma^2_B}{\sigma^2_A}$$

$$R^2_{A,C} = \frac{\gamma^2\sigma^2_C}{\sigma^2_A}$$

$$R^2_{A,\epsilon} = \frac{\epsilon}{\sigma^2_A}$$

Based on the above equations, the regression coefficients and the residual variance can be solved:

$$\beta = i_{A,B}\sqrt{\frac{\sigma^2_A}{\sigma^2_B}R^2_{A,B}}$$

$$\gamma = i_{A,C}\sqrt{\frac{\sigma^2_A}{\sigma^2_C}R^2_{A,C}}$$

$$\alpha = \mu_A - \beta\mu_B - \gamma\mu_C$$

$$\sigma^2_\epsilon = \left(1 - R^2_{A,\epsilon}\right)\sigma^2_A$$

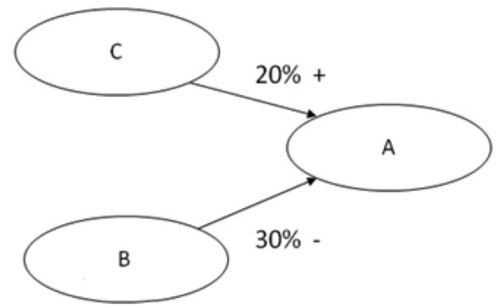
Converting the effect strengths to a covariance matrix

It may be useful for computational purposes to see the regression model as a joint probability distribution of the three variables A, B, and C. Then the distribution can be characterized with a vector of means and a covariance matrix. The mean vector is simply

$$\boldsymbol{\Omega} = (\mu_A, \mu_B, \mu_C)$$

The covariance matrix can be calculated based on the knowledge that the degree of determination depends on the covariance and on the marginal variances:

Fig. A1. Example directed acyclic graph.



$$\sqrt{R^2_{A,B}} = \frac{\mathrm{Cov}(A,B)}{\sigma_A\sigma_B}$$

Thus, the covariance matrix is given by

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2_A & i_{A,B}\sqrt{R^2_{A,B}\sigma^2_A\sigma^2_B} & i_{A,C}\sqrt{R^2_{A,C}\sigma^2_A\sigma^2_C} \\ i_{A,B}\sqrt{R^2_{A,B}\sigma^2_A\sigma^2_B} & \sigma^2_B & 0 \\ i_{A,C}\sqrt{R^2_{A,C}\sigma^2_A\sigma^2_C} & 0 & \sigma^2_C \end{pmatrix}$$

Dirichlet prior distribution for the effect strengths

The effect strengths $R^2_{A,B}$, $R^2_{A,C}$ and the proportion of unexplained variation $R^2_{A,\epsilon}$ must sum up to one. Since there is uncertainty about the strengths, the joint prior distribution of the strengths must be constructed in such a way that this constraint holds. Perhaps the simplest way to achieve this is to use a Dirichlet distribution as a prior for the vector of proportions:

$$\left(R^2_{A,B}, R^2_{A,C}, R^2_{A,\epsilon}\right) \sim \mathrm{Dir}(\theta_1, \theta_2, \theta_3)$$

where parameters θ are defined based on the point estimates of the effect strengths and based on the amount of relative uncertainty (η) specified using the interactive graph:

$$\theta_1 = \widehat{R^2_{A,B}}\eta$$

$$\theta_2 = \widehat{R^2_{A,C}}\eta$$

$$\theta_3 = \widehat{R^2_{A,\epsilon}}\eta$$

The above parameterization leads to marginal prior variances of the form

$$\mathrm{Var}\left(R^2_{A,B}\right) = \frac{\widehat{R^2_{A,B}}(1 - \widehat{R^2_{A,B}})}{\eta + 1}$$

and results in approximate 95% probability interval of

$$\widehat{R^2_{A,B}} \pm 1.96\sqrt{\frac{\widehat{R^2_{A,B}(1 - R^2_{A,B})}}{\eta + 1}}$$