ICES STATUTORY MEETING 1993

C.M. 1993/D: 13
Ref. M
Statistics Committee
Ref. Anacat Committee

BELIEF NETWORKS IN FISH STOCK ASSESSMENT – THE BALTIC SALMON CASE

OLLI VARIS*, SAKARI KUIKKA** AND JUHANI KETTUNEN**

* *Helsinki University of Technology, Laboratory of Hydrology and Water Resources Management, FIN-02150 Espoo, Finland.*
** *Finnish Game and Fisheries Research Institute, P.O. Box 202, FIN-00151 Helsinki, Finland.*

ABSTRACT

Belief networks contain a set of interlinked nodes. Based on Bayesian calculus, the links transfer information between the nodes containing probability distributions. The aim of the paper is to present a brief theoretical basis for belief networks and to illustrate their capability to support fish stock assessment. The flexibility of problem formulation and efficiency in knowledge acquisition are emphasised. It is also shown, how various types of uncertain information can be handled and merged in belief networks. The technique is applied to Baltic salmon assessment problems. Regression models, the VPA, and expert judgement are used together to estimate the parameters for the terminal stock, to produce stock forecasts, and to assist in the total allowable catch decision, which is done on the basis of highly uncertain information.

1. INTRODUCTION

Management decisions of natural and environmental resources need often be made under high uncertainty, and expert judgement is thus in central role. There are two key reasons to this. First, it were often most irrational in practice to thrive at collecting a waterproof empirical data. This is due to economic constraints. Second, the potential changes in the system in comparison to the past are often so high that extrapolation of past development is vague. This is due to high variability in semi-natural systems caused by numerous uncontrolled and controlled issues. Moreover, the management targets often at changing the system substantially to a desired direction, and invalidates the use of historical records.

A typical example from fisheries management is the annual Baltic salmon quota decision. The stocking of reared salmon to the Baltic has enhanced the salmon fisheries, and the wild stocks are under severe risk of being extinct (e.g., Anon. 1993). The managers (the International Baltic Sea Fisheries Commission) have defined that the goal for management is *to safeguard wild salmon stocks.* Stock assessment is made to support this goal. The economic rationale to gather empirical information is far too low to provide enough data for purely empirical stock forecasts. Furthermore, the system impacted by the management policy, including ecological, social, economic, and political facets, is under practically unpredictable changes and transitions (cf., Kuikka & Varis 1992, Kuikka 1993).

For the purposes of fish stock assessment, the information and experience available allows the use of empirical, regression-type of models for certain relations between sub-stock volume data, growth parameters, water quality data, and so on. Also the VPA equations (Beverton & Holt 1957, Gulland 1983) have been found very useful, although they are not identifiable from data and the essential parameters (mortalities) are assessed by experts. When producing age-structured stock forecasts from this information which is of rather split character, the role of experts is important. Some experts prefer the use of selected empirical models, while some rather use the VPA equations. Clearly, any present assessment technique alone suffers from severe limitations, and all possible, relevant information and models should be taken into account.

The objective of this study is to produce a computerised environment that allows the inclusion of empirical models and the VPA equations in one context. The uncertain information from multiple sources can be merged together by one, or preferably, several experts. The system allows interactively the detection of controversies in information, arbitrary weighting of different models, tuning of the VPA equations, calculation of forecasts, and definition of the fisheries quota (Total Allowable Catch) decision. Methodologically, this has been realised by using a probabilistic, belief network in which the above mentioned models have been embedded.

2. BALTIC SALMON MANAGEMENT – THE PREDICTION PROBLEM

The present state of wild Baltic salmon stocks is poor. The share of the reared stock from the whole stock has increased remarkably during the last 10 - 20 years. In 1980, the recruitment of wild stock to fishery was about 20 % of the total recruitment, but in 1988 - 1989 it was only about 9% (Anon. 1992, Tables 7.2.5.2.1 & 7.2.6.2.1). About 35% of the wild recruitment comes from the northernmost rivers, even though the smolt production potential of these rivers is high and their water quality is good. Lack of spawners is an obvious reason for the poor state of the northernmost stocks (Anon. 1993).

To achieve the management goal – to safeguard wild salmon stocks – managers have decided to use the Total Allowable Catch (TAC) policy for years 1991 - 1993. The ICES produces the information on the state of the stocks, which is needed in the management decisions.

Salmon is a short living species; the fishing mortality of Baltic Salmon is around 2/year in the most important age groups. This implies that around 90% of the individuals of these age groups are being caught annually. Fisheries is based on two age groups, A1 and A2. Even though the total recruitment is almost totally based on reared salmon released annually (known quantity) mainly to the northern part of the Baltic Sea, the recruitment variation is large. For example, in 1987 the post-smolt survival of the reared salmon was about 10 % and in 1988 around 30% (Anon. 1992).

The most important age group (A2) in the spawning stock of the target year (TAC year) is released two years earlier, i.e. during the latest data year. However, this age group does not recruit to any main fisheries during the first year in the sea. Therefore, there is no such catch information for this age group, which could be used in a usual way in VPA based assessment.

Moreover, variation in the growth rate is also large (Kuikka 1991). Growth rate improved almost by 60 % in the end of 1980's. While the selectivity of the most important fishery (drift net fishery) is very high, the fishing mortality of age group A1+ has large variation and it is not in clear connection with the total effort. Therefore, the tuning of this fishing mortality value is a difficult task.

Owing to these variations, the uncertainties of the stock predictions and TAC based management are high. Some additional variables having predictive power have been used in the assessments to reduce the uncertainties (Anon. 1992). Available data suggest, that both the growth of the post smolts in the sea during the first half year and the temperature during the first months in the sea can be used in the predictions of the size of age group A0. Moreover, the CPUE (catch per unit effort) data of the driftnet fishery and the growth rate can be used in the assessment of the terminal F values of age group A1. Effective use of these predictive variables requires methodology, that allows the consideration of all relevant, information that could be used in salmon stock assessment.

3. THE BELIEF NETWORK APPROACH

One of the ways a human mind comprehends a habit is through defining objects, and postulating associations between them. When considering a certain object in this context, she sees it simultaneously as one unit, and as a detail in interaction with the rest of the context. Systems, in which uncertain information is available on a set of mutually dependent objects, would be approached in Bayesian calculus by assigning a prior probability distribution to each object. Thereafter, the strength and character of the dependency between each object pair would be inserted. With this information, posterior probability distributions are calculated for each object. This is actually the key idea in belief networks. In the belief network terminology, the objects are called as nodes, their associations as links, and the context as a network.

Belief networks have emerged from the tradition of Bayesian statistics in the 1980s (see Shafer & Pearl 1990). The cornerstones were laid by Pearl (1986, 1988). The key idea is that any new information introduced in the net can be propagated to any direction, not only to one direction (Fig. 1). This feature has been realised using bi-directional information flow in the links. The nodes are able to merge the information from these systems and to update it. Pearl (1988) presented a sequence of algorithms starting from a chain, and proceeding through trees and polytrees to networks. The basic problem in network algorithms is to cope with circular references. The algorithms presented consist of approximate methods such as simulation.

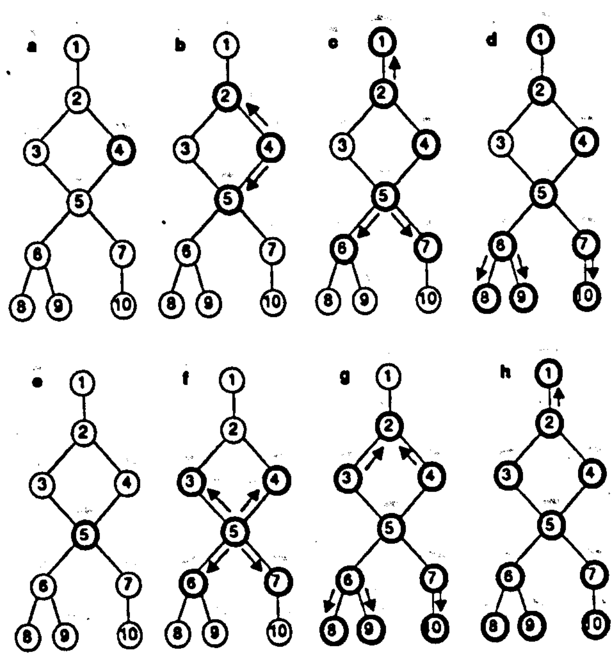


Fig. 1. Example on bi-directional propagation using the approach by Varis (1993). In *a*, a new message updates the node 4. The message is propagated through the net by two belief trees in opposite directions. The trees are directed by the sequence of nodes. If the message to node 4 were to update also the node 3, then there should be a link in between, or the node sequence should be different. Again, in *e*, a new message updates the node 5. The message is propagated and posterior distributions are being calculated.

As a summary, a network consists of n nodes that can be arbitrarily linked to one another. The prior probabilities assigned to the outcomes are updated with the information linked from other parts of the net, yielding the posterior probability distribution. A network is constructed and modified interactively during the modelling procedure. Essential is to find (1) most relevant variables to a specific problem, and (2) define links between them in the best possible way. The methodology presented in short below (Varis 1993) is deeply rooted to the work by Pearl (1986, 1988), it has adsorbed certain features from the influence diagram methodology by Shachter (1986), and a number of extensions have been made.

Nodes

Each node i in the network contains:

- A vector of possible (discrete) *outcomes* $\mathbf{y}_i$. They can be defined as inputs or they can depend on outcome values of other nodes.
- A *prior probability distribution*, expressed with probabilities $e_1...e_k$ assigned to k outcomes given, summing up to unity. These constitute a k dimensional vector $\mathbf{e}_i$, also known as evidence vector. If no prior belief exists, a non-informative prior, e.g., a unit vector, is used.
- A *sign* indicating the direction of change. The sign may either positive (implying growth, increase, addition, enlargement, etc.) or negative (decline, decrease, reduction, lessening, etc.).
- A *posterior probability distribution* $\mathbf{Bel}_i$.

In general, the nodes are probabilistic (uncertain). They can, however, have one outcome with the value 1 and the others with 0, and be thus certain. If an outcome in a prior distribution gets the value zero, then also the posterior distribution will also have a zero value in the respective outcome. Only uncertain issues accept updating from other parts of the model.

In decision analysis or optimization, some nodes may be understood as controllable, decision nodes. One or several nodes can act as criteria or constraints to decision making, and constitute one or more objective functions. Removing uncertainty from a node, i.e., selecting one of its outcomes to have a probability 1, may be used to simulate a decision or other action that has been or will possibly be made. Its implications are propagated through the network, and they can be observed at each of the successor nodes. Some of them are usually more critical than the others, presenting objectives or constraints, and the adjustment of the control policy simulated can be based on observed changes in those nodes.

Links

A link transmits information from a node to another node. When defining the concept of link, Pearl (1988) lists the following four primitives, for which we have created examples: (1) *likelihood*: fish growth is more likely to be increased than decreased; (2) *conditioning*: if temperature increases, then fish will grow faster; (3) *relevance*: whether the fish will grow faster depends on whether there will be increased temperature; and (4) *causation*: increasing temperature will enhance fish growth. For more discussion and illustration, see Pearl (1988). Another classification based on the information source to the link is given by Varis (1993): (1) *deductive*: there is prior knowledge, theory, or belief concerning the interdependency of the two nodes; and (2) *inductive*: there is, correspondingly, empirical evidence or data.

Links are in two layers. An *uncertainty link* is defined as the link matrix $\mathbf{M}_{i|j}$ between two nodes i and j. An *outcome link* presents a relation between outcomes $\mathbf{y}_i$ and $\mathbf{y}_j$ of i and j. Because the approach requires that each outcome has one value, the propagation of outcome values is unidirectional, and a functional relationship exists, $\mathbf{y}_j = f(\mathbf{y}_i)$. This relation can be either *deterministic* (numerical), or *logical* (rule-based). Fig. 2 (see also Appendix 1) portrays the idea of two different layers of links between three nodes: probabilistic links propagated bi-directionally through link matrices $M_{i|j}$, and outcome links $y_j = f(y_j)$.

The links can be direction specific, i.e., $M_{i|j} \neq M_{j|i}$, or they can be negative, i.e., increase in node *i* implies decrease in node *j*. A link matrix can be either symmetric or non-symmetric. If the link represents pure correlation between two variables, then there is no reason to use direction specific or non-symmetric links, but negative links are very useful. In many cases, there is a clear causal dependency between two nodes, and direction specific links and non-symmetric link matrices are well applicable. Varis (1993) presents indices for both defining the information content of a link, expressed as a value ranging between -1 and 1 (0 is non-informative), and inverse use of these indices for generating a link matrix from a link strength index value.

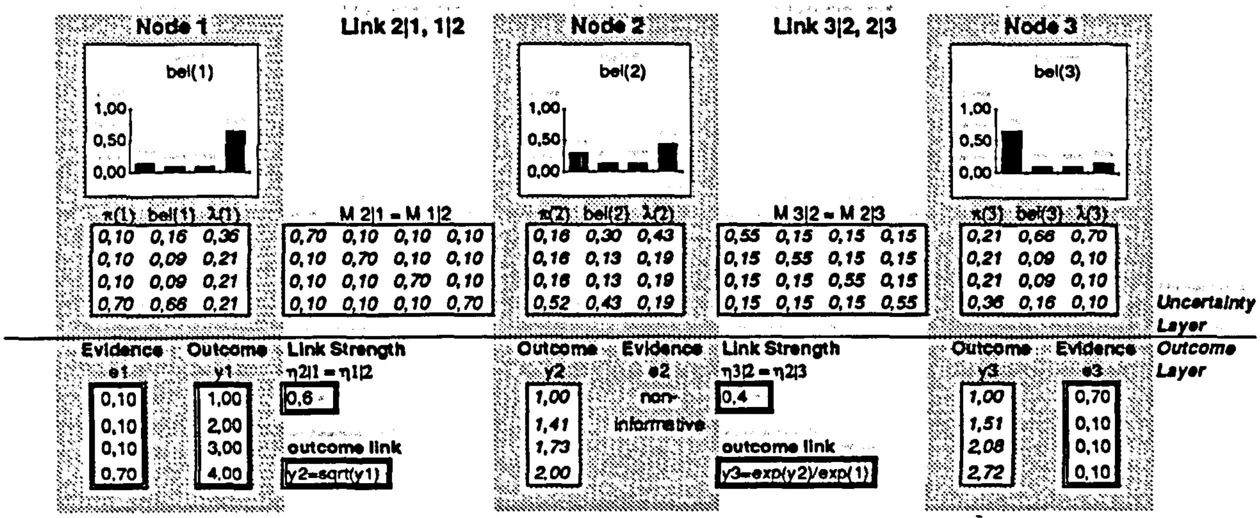


Fig. 2. An example three node belief network model with deterministic outcome links. Computed values are set in italics, and inputs are in double-line bordered cells. Uncertainty layer is above with figures, and outcome layer is below.

Network propagation

The question is how to calculate posterior belief distribution vectors $\mathbf{Bel}_j$ for the nodes, updating their prior information. The algorithm by Varis (1993) is based on Pearl's (1988) tree algorithm. Two independent tree messages (denoted usually as π and λ) are computed (cf., Fig. 1), The updated belief is obtained as their and the prior's convolution product. The nodes are linked with link matrices that can be chosen direction specific. Positive and negative dependencies between nodes are allowed. Computationally, all nodes are linked with each other, and a non-informative link implies no connection. In a non-informative link, each link matrix element has an equal value. All information on the probabilistic relations between nodes is expressed by unequal link matrix element values.

Prospects for applicability

Being a relatively novel approach, the number of applications is limited at this stage. Varis (1993) provides a prospective discussion and review of applicability of belief networks in the management of natural resources and the environment. He groups the potential modelling directions in the following five clusters, and discusses and illustrates them with examples: (1) belief and knowledge acquisition, (2) decision analytic use, (3) analytical, mechanistic and process modelling, (4) spatial and temporal correlations, and (5) learning and adaptive modelling. In practice, a belief network can include properties from each of these categories (Fig. 3), being a hybrid of several, conventionally distinct, computational modelling approaches. In this application, regression models and a deterministic model (VPA) are used together as described below.

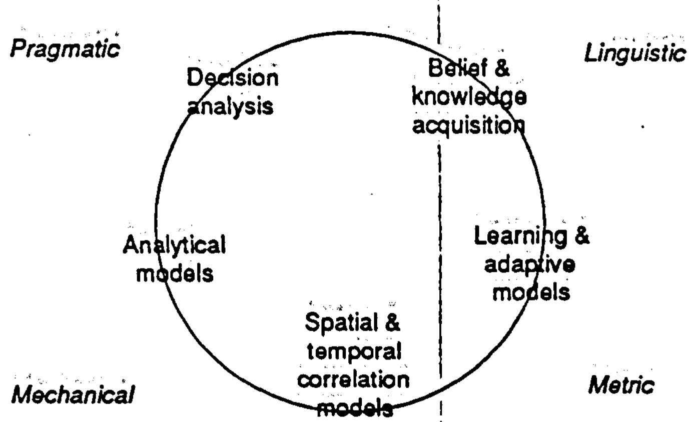


Fig. 3. The belief network approach facilitates the combined use of several, methodological and paradigmatic (in italics, see Beck 1991) facets that are often seen as being far from one another.

4. ASSESSMENT AND PREDICTION OF BALTIC SALMON STOCKS USING THE BELIEF NETWORK VPA

The assessment procedure consists of four steps. They are

- Regression models for predicting selected quantities of the salmon stock.
- Calibration of the VPA model and linking it and regression models with a belief network.
- Prediction of the stock for the present and the coming year using both the VPA and the belief network model.
- Definition of the total allowable catch using the predictions.

Fig. 4 illustrates the schematic structure of the assessment procedure showing the relations of different submodels used.

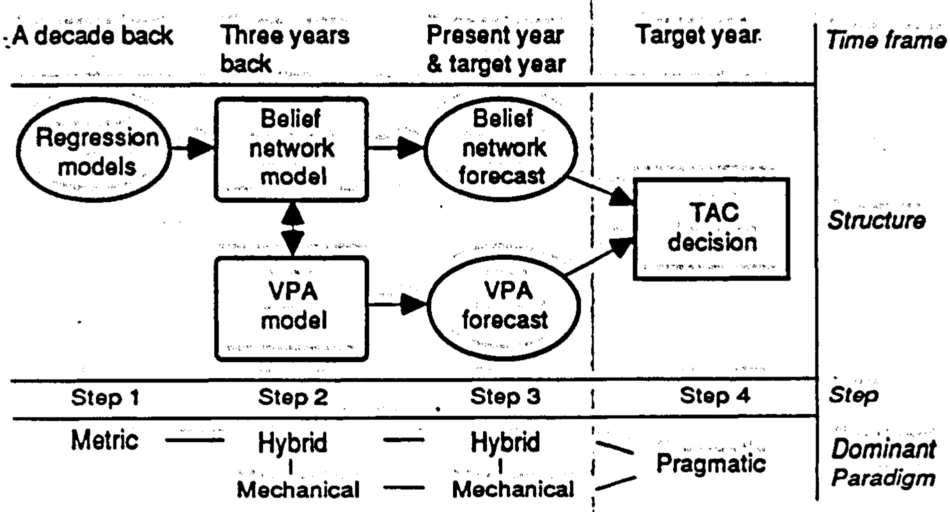


Fig. 4. Schematic diagram of the structure of the assessment procedure. The more angular a module is, the more important is the expert judgement component.

There are a number of inputs to the model during the procedure (Table 1). Their uncertainties are usually presented as the coefficient of variation value *cv*:

$$cv = \frac{\sigma}{\mu} \tag{1}$$

where σ = standard deviation and μ is mean. The normality assumption is used in the whole procedure. The belief network handles probability distributions discretised in three outcomes. The input distributions are discretised to have values $\mu - 0.97\sigma$, μ, and $\mu + 0.97\sigma$. This gives each outcome an equal prior probability, 1/3. This allows easy comparison of prior and posterior distributions.

Sample data from the situation from 1992 assessment work (data from 1991 and TAC advice for 1993) is used throughout this and next section. Catches of wild and reared salmon are summed up due to the large uncertainties in discrimination of wild/reared salmon by scales (Anon. 1993). The computer implementation of the procedure is described at the end of this Section. In addition, the implementation includes a collection of diagnostic plots, examples of which are given below.

Table 1. Inputs during the assessment procedure.

Step	Type of variables	Variable (see Table 3)	Parameter*	Substock
1	Information on regression models	(see text and Table 2)	(see Eq. 2)	each year, smolts, A1
2	VPA parameters	$m_{i,j}$ (natural mortality)	μ	each year & each year class
			cv	one value for all
		$F_{i,j}$ (fishing mortality)	μ, cv	terminal age groups
	Observations	$C_{i,j}$ (catch)	μ	each year & each year class
			cv	one value for all
		S_j (post-smolt survival)	d	each year
	Weights for (beliefs on) regression models		w	
	Belief network link strength parameters	Beliefs on VPA equations	α	
		year-to-year dependence	α	
		interannual dependence	α	
3	Decision from previous year	TAC for current year	d	
	Assumption on stationarity	Volume parameter β	d	
4	Decision variable	TAC for coming year	d	
	Assumption on stationarity	Volume parameter β	d	
	Properties of the stock	Percentage of wild stock	μ, cv	each year class
		Percentage of wild females	μ, cv	each year class

* μ = mean, cv = coefficient of variation, d = single value, α = link strengths of the belief network.

Step 1: Regression models

The application makes use of four linear regression models (Table 2). They are based on historical monitoring data, available from nine to ten years backwards. These regressions have been found useful, simple predictors in Baltic salmon stock assessment (cf, Anon. 1992).

Table 2. Regression models.

Independent (x)	Dependent (y)	Equation	r^2	n
Temperature at Seili [°C]	Post-smolt survival [%]	$y = -45.8 + 4.64x$	0.73	10
Growth of A0+ salmon [cm]	Post-smolt survival [%]	$y = 10.9 + 1{,}43x$	0.60	10
Catch per unit effort for A1+ [number of individuals]	Stock size of year class A1	$y = 355{,}5 + 33x$	0.23	10
Mean weight of A1+	Fishing mortality of year class A1	$y = -1.06 + 0.59x$	0.70	9

A regression model $y = a + bx$ is used in prediction in the following manner. Given the observation of the predictor x_m, the mean of the predicted value is $\hat{y} = a + bx_m$, and the standard deviation is

$$\hat{\sigma}_y = \sqrt{\sigma_y^2 (1 - r^2) \frac{n-1}{n-2}} \qquad (2)$$

where n is the number of observations.

Step 2: Calibration of the VPA model and linking it and regression models with a belief network

VPA equations by Pope (1972) were used, in the form shown in Table 3. The fishing mortality values for terminal stocks are assumed. Three year classes are included, and three years are calculated backwards, including the terminal year. Stocking data was used to calculate post-smolt survival rates for each of the years included. The proportion of the year class A3 is very low (Kuikka & Varis 1991). Therefore, it was excluded.

Table 3. VPA and post-smolt survival equations. $m_{i,j}$ is natural mortality of age group i in year j.

Variable	Year -2	Year -1	Year 0
Post-smolt survival (S_j)	$S_2 = N_{0,2}/s_2$	$S_1 = N_{0,1}/s_1$	$S_0 = N_{0,0}/s_0$
Fishing mortality for A0 ($F_{0,j}$)	$F_{0,2} = -m_{0,2} - \ln(N_{1,1}/N_{0,2})$	$F_{0,1} = -m_{0,1} - \ln(N_{1,0}/N_{0,1})$	$F_{0,0}$
Stock size of A0 ($N_{0,j}$)	$N_{0,2} = C_{0,2}e^{m_{0,2}/2} + N_{1,1}e^{m_{0,2}}$	$N_{0,1} = C_{0,1}e^{m_{0,1}/2} + N_{1,0}e^{m_{0,1}}$	$N_{0,0} = C_{0,0}\frac{m_{0,0} + F_{0,0}}{F_{0,0}\left(1 - e^{-m_{0,0} - F_{0,0}}\right)}$
Fishing mortality for A1 ($F_{1,j}$)	$F_{1,2} = -m_{1,2} - \ln(N_{2,1}/N_{1,2})$	$F_{1,1} = -m_{1,1} - \ln(N_{2,0}/N_{1,1})$	$F_{1,0}$
Stock size of A1 ($N_{1,j}$)	$N_{1,2} = C_{1,2}e^{m_{1,2}/2} + N_{2,1}e^{m_{1,2}}$	$N_{1,1} = C_{1,1}e^{m_{1,1}/2} + N_{2,0}e^{m_{1,1}}$	$N_{1,0} = C_{1,0}\frac{m_{1,0} + F_{1,0}}{F_{1,0}\left(1 - e^{-m_{1,0} - F_{1,0}}\right)}$
Fishing mortality for A2 ($F_{2,j}$)	-	$F_{2,1}$	$F_{2,0}$
Stock size of A2 ($N_{2,j}$)	-	$N_{2,1} = C_{2,1}\frac{m_{2,1} + F_{2,1}}{F_{2,1}\left(1 - e^{-m_{2,1} - F_{2,1}}\right)}$	$N_{2,0} = C_{2,0}\frac{m_{2,0} + F_{2,0}}{F_{2,0}\left(1 - e^{-m_{2,0} - F_{2,0}}\right)}$

The VPA model was used as such in further steps, and it was also used as an outcome layer in a belief network model. The architecture of the uncertainty layer is presented in Fig. 5. Each link has a symmetric link matrix $M_{i|j}$, which is not direction specific. The matrices are presented as single input parameters, using the link strength parameter approach (see Varis 1993).

The interpretation of the links is belief on the level of dependency, in the nature, of the respective variable pair. There are links that refer to interannual dependence of the year classes, links that stand for within-year dependence of age groups, links indicating the belief on the level of description of a mathematical relation (VPA equation) in the model, and combinations of these.

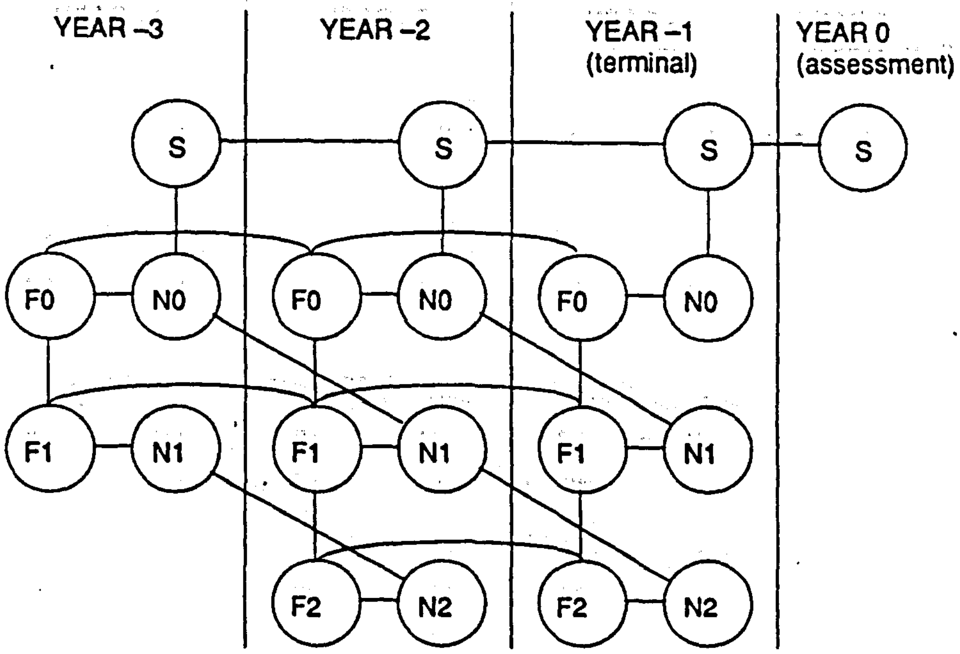


Fig. 5. The belief network architecture: uncertainty layer.

In addition to the VPA model information, the predictions from the regression models (Table 2) are merged in the belief network model. This is done in nodes PSM, F1, and N1 using the following equations, under the normality assumption:

$$\mu' = \frac{\mu_{vpa} + \sum \mu_k w_k}{1 + \sum w_k} \tag{3}$$

$$\sigma'^2 = \frac{1}{\frac{1}{\sigma_{vpa}^2} + \sum_k \frac{w_k^2}{\sigma_k^2}} \tag{4}$$

where μ' is joint mean and σ'^2 is joint standard deviation of the VPA and regression models k, and w_k is the weight assigned to the regression model k. For the above mentioned nodes, the normal distribution represented by these parameters is used as the prior distribution to the belief network. For the other nodes, the VPA outcome represented with μ_{vpa} and σ_{vpa}, is used as the prior distribution.

After each change, the belief network returns the updated posterior distributions for each node. These distributions are used in (1) iteration of terminal fishing mortality rates (both means and coefficients of variation), (2) other diagnostic purposes concerning the information available at this stage, and (3) as inputs to the predictive model. The first of these uses deserves more detailed explanation. The differences between the prior and posterior distributions indicate the additional information obtained from the belief network, using desired link strength parameter values. If desired, both the prior mean and the prior coefficient of variation can be iterated quite quickly to equal the corresponding posterior values, once there are not major controversies between different sources of information. These, possible controversies, together with relative weighting of information from different sources as merged using Eqs. 3 and 4, can be illustrated, e.g., as shown in Fig. 6.

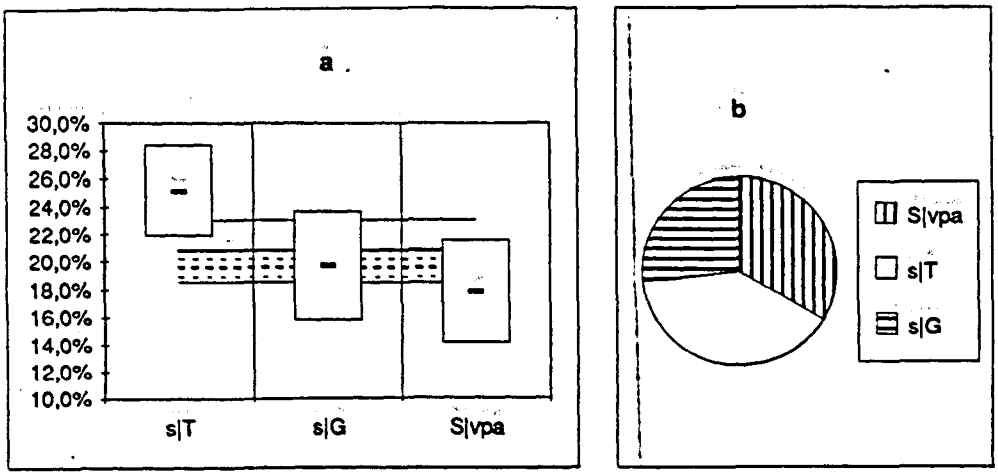


Fig. 6. Sample diagnostic plots for merging of VPA and regression information. (a) Post-smolt survival rate for year –1 from temperature regression (S I T), from the regression based on growth data of post smolts (S I G), and from VPA model (S I VPA). The prior and posterior distributions of the belief network are also shown. The figure shows the $\mu + \sigma$, μ, and $\mu - \sigma$ values for each distribution. Solid, horizontal lines are for prior and dotted for posterior distributions. (b) Proportion of explanation of the three different models to the prior distribution of the belief network model (values in the denominator of Eq. 4 are compared).

Step 3: Prediction

The assessment procedure produces two predictions, using Eqs. in Table 3. They are:

- VPA prediction.
- Belief network model prediction, which aggregates the VPA information with regression models and the inserted weights and beliefs on these information, and uses this as input to the predictive equations of the VPA.

The most recent stocking information is given as input, together with a volume parameter β which allows the inclusion of an assumption of the level of stationarity of the structure of fishing between consecutive years. It is left as the task of the expert(s) to choose heuristically the recommendation for the TAC decision, given two uncertain predictions (Fig. 7). Besides the belief network prediction, the VPA prediction has been in such a wide use within ICES that its inclusion as a comparative element in the procedure has been considered important.

Step 4: Definition of the Total Allowable Catch

At this step, a set of additional monitored, but still rather uncertain parameters must be fixed. They are the proportions of wild salmon and females, both for all three age groups (A1, A2 and A3). For them all, also coefficients of variation are needed to allow handling of uncertainty. Also here, the parameter β can be used to describe the effect of the total effort change. Scale sample data on the proportion of wilds (including some error in the discriminate assessment by scales) and estimates on the wild and reared smolt productions are available (e.g., Anon. 1993). The uncertain forecasts of the total stock, the age structured stock, and the number of wild, spawning females are the basis for the TAC recommendation (Appendix 2c).

Sensitivity analysis of belief network links

In order to analyse and illustrate the roles of different links in the network, two sensitivity analyses were carried out. In the first one, the studied variables were the fishing mortalities for the terminal year, and in the second, the stock sizes for the terminal year (Table 4). The relative difference in the deviation between the prior and the posterior means was used as the indicator of sensitivity Δ:

$$\Delta_n = \frac{\frac{\mu'_n - \mu_n}{\mu_n} - \frac{\mu'_N - \mu_N}{\mu_N}}{\frac{\mu'_N - \mu_N}{\mu_N}} \tag{5}$$

where μ is prior and μ' is posterior mean. N refers to nominal, and n to perturbed link strength parameter values. The used values were 0.1, 0.5, and 0.9.

Only the links in which one or both of the interlinked variables were fishing mortalities, induced changes in the first sensitivity study (Fig. 7). Terminal fishing mortalities were most sensitive to the links between *F*s of two consequent age groups, within one year.

In the second study, each studied link group caused changes in stock sizes of the terminal year. The most sensitive link groups were those between post-smolt survivals and *N*(A0) (within one year), and those between post-smolt survivals and between *N*'s, both between two subsequent years.

Table 4. Classification of links for the two link sensitivity analyses, and indexing in Figs 7 and 8. The indices in bold indicate the most sensitive cases. + denotes next age group, * denotes next year.

Time domain	Link type	Terminal *F* (Fig. 7)	Terminal *N* (Fig. 8)
Within a year	Between post-smolt survivals and *N*: A0	*no influence*	S — N0
	Between *F* and *N* of an age group	F — N	F — N
	Between *F*s of two consequent age groups	F — F+	F — F+
Between two subsequent years	Between post-smolt survivals	*no influence*	S — S*
	Between *F*s	F — F*+	F — F*+
	Between *N*'s	*no influence*	N — N*

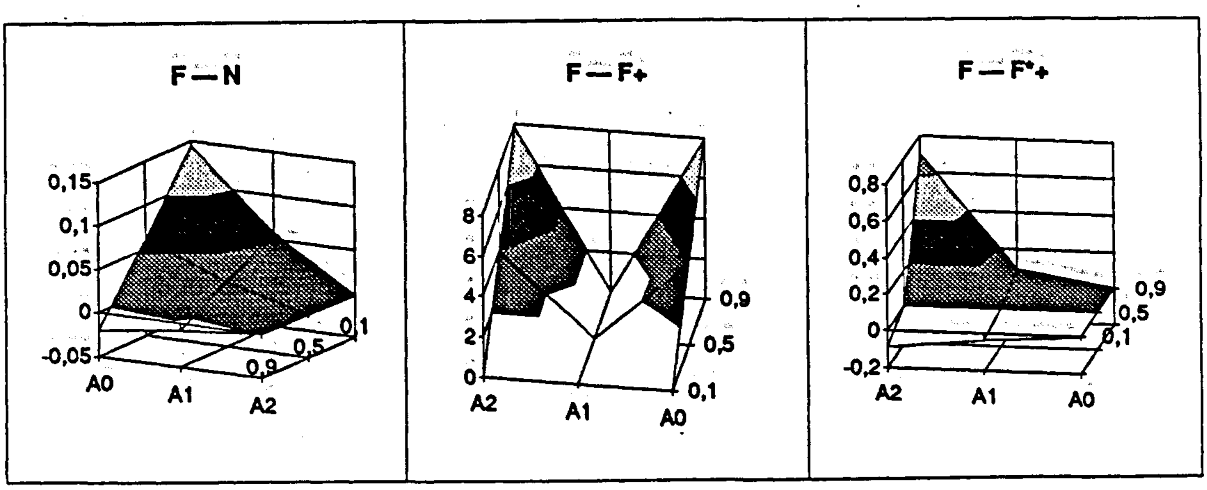


Fig. 7. Results of the sensitivity study for fishing mortalities of the terminal year, with respect to links as grouped in Table 4.

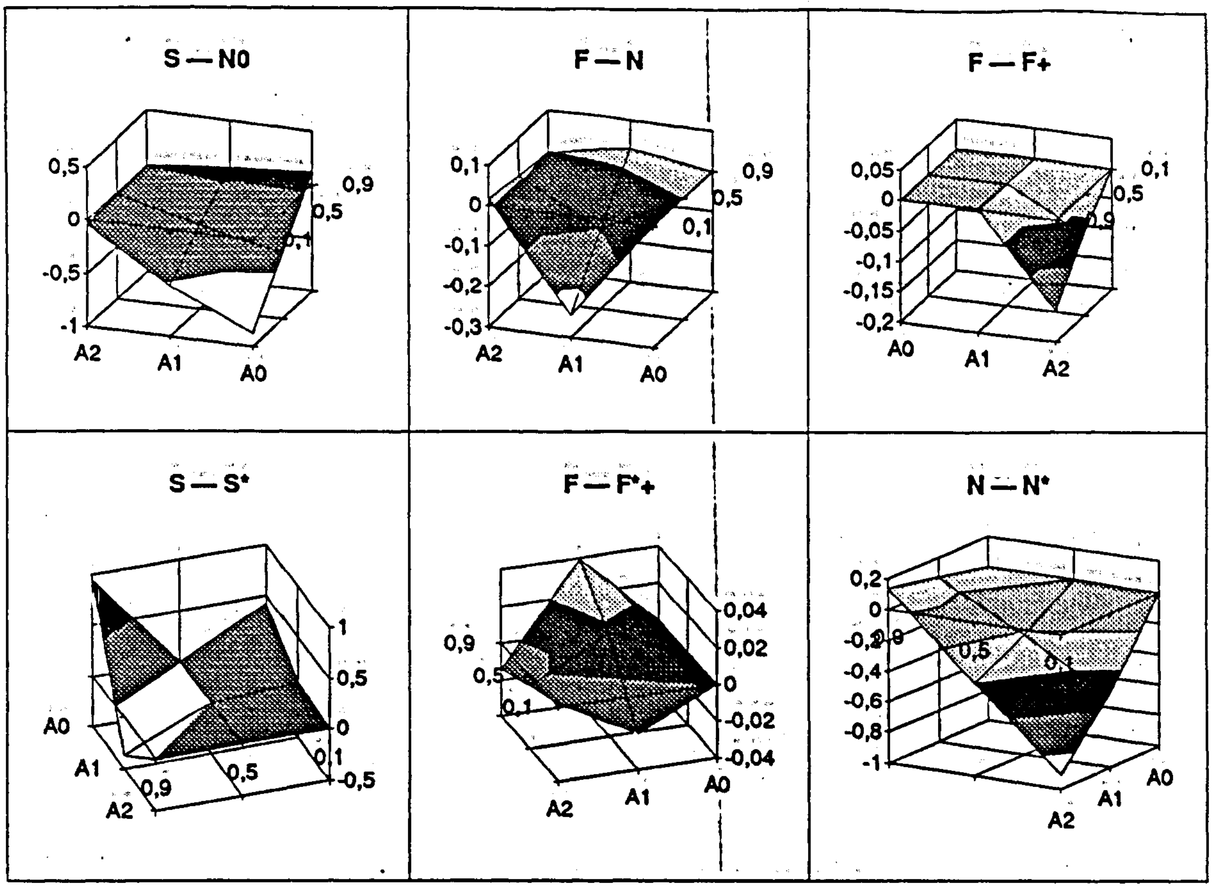


Fig. 8. Results of the sensitivity study for stock sizes of the terminal year, with respect to links as grouped in Table 4.

As a third sensitivity study, a corresponding procedure was performed to the weights of the two regression models predicting post-smolt survival (Fig. 9), with respect to stock sizes for the terminal year. The temperature regression appeared to have he greatest impact on A1, while the growth regression appeared the more influental the younger the fish group is.

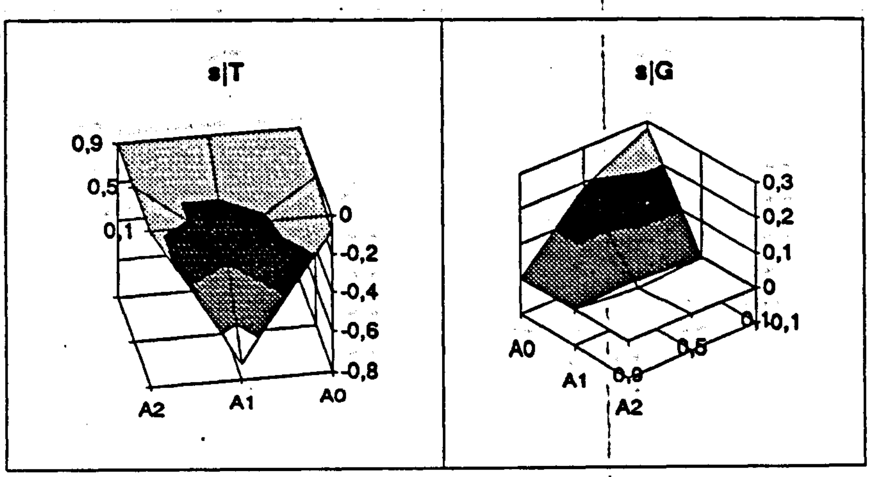


Fig. 9. Results of the sensitivity study for stock sizes of the terminal year, with respect to two regressions used to predict post-smolt survival.

Computer implementation

Microsoft® Excel 4.0 spreadsheet was used for coding the system described above. It allows the portability within Windows™ and Apple® Macintosh® environments. For the belief network, the FC BeNe (Varis 1992) spreadsheet toolkit was used. Excel contains a number of statistical, mathematical, and matrix functions that were very practical in the realization of the system. For instance, functions such as NORMDIST, NORMINV, and NORMSDIST were used frequently when dealing with normal probability distributions. As another example, Appendix 1 shows the source code of the belief network (actually a chain) in Fig. 2, that uses the MMULT function to perform belief updating in the net.

The user interface worksheet consists of six areas: one for each assessment step (Appendix 2) described above, a selection of (readily extendable and relocateable) diagnostic plots, and a variety of auxiliary, computational routines.

5. DISCUSSION AND CONCLUSIONS

In the management of natural and environmental resources, there is a need to produce tools – computerised systems – that provide aid to experts in combining the information from multiple sources, consisting typically of numerous, very uncertain entities. It is often useful to be able to combine various empirical information to structurally experienced, deterministic models such as the VPA in fisheries management. In addition to the support in tuning of the parameter values, and in the assessment of the associated uncertainty is very important. All this is because data is very often exceedingly sparse, the costs for data collection allowing purely empirical forecasts are far too high to be rational, and the importance of the analytical inclusion of the associative way of human judgement in complex problems (cf., Rowe & Boulgarides 1983, Rowe & Watkins 1992). Yet above all, we are dealing with important, real world problems that call for the best available methodology. Computational problem solving should target, in particular, at enhancing the learning about the problems (Shafer 1981).

The present application shows an example on a management problem where plenty of expert judgement is needed. Especially, the combined use of different information sources is a difficult and time consuming process, e.g., in the working groups of ICES. Moreover, the decision problem and the system (ecological, political, etc.) are subjected to continuous, substantial, almost unpredictable, changes. Due to the short time series, r^2 is a poor basis to judge the relevancy of different information sources. In this case, the historical data gives a good correlation for the mean weight of A1+ and fishing mortality, but due to the changes of the market prices, the effort of the off shore fishery has decreased and the *F* prediction is not relevant anymore.

With respect to the advice given to the managers, the belief network approach offers possibilities for constructing computerised environments that allow systematic group discussions on the role, reliability, and usability of information from various, different sources, and of varying character. Systematic use of link parameters gives a good overview of the role of different information. The belief network helps both in the diagnosis of the problem, including information available, and in the predictions. The more deviations there are between priors and posteriors, and between VPA forecasts and belief network forecasts, the more inconsistencies there are in the system.

Even though the role of subjective information is often understated and even denied at ICES, subjective evaluation is very often needed. It is very important that the role of expert judgement is made clear, and the assessment procedure is open for discussion, e.g., at workshop meetings. In addition, it is important that all the relevant information including computational models are set in a framework that allows their inclusion or exclusion, or merging and weighting – depending on what is seen reasonable – to produce the best available forecast for the given purpose. In the sample case, the belief network model produced a more accurate forecast than the VPA model alone. This is due to the ability to include empirical information.

REFERENCES

Anon. 1992. Report of the Baltic Salmon and Trout Assessment Working Group. *ICES* C.M. 1992/Assess: 10.

Anon. 1993. Report of the Baltic Salmon and Trout Assessment Working Group. *ICES* C.M. 1993/Assess: 14.

Beck, M.B., 1991. Forecasting environmental change. *J. Forecasting* 10: 3-21.

Beverton, R.J.H. & Holt, S.J. 1957. On the dynamics of exploited fish populations. *Fishery Investig. Series* 2, 19: 1-533.

Gulland, J.A. 1983. *Fish Stock Assessment – A Manual of Basic Methods.* J. Wiley & Sons, Chichester.

Kuikka, S. 1991. Effects of some external factors on the predictability and production capacity of Baltic salmon stocks. *ICES* C.M. 1991/M:29.

Kuikka, S. 1993. Salmon is a political animal (Lohi on poliittinen eläin, in Finnish). *Helsingin Sanomat*, 5 June.

Kuikka, S. & Varis, O. 1991. Probabilistic assessment of TAC based fisheries management of Baltic salmon stocks. *ICES* C.M. 1991/M:30.

Kuikka, S. & Varis, O. 1992. Use of Bayesian influence diagram in fisheries management – the Baltic salmon case. *ICES* C.M. 1992/D:5.

Pearl, J. 1986. Fusion, propagation, and structuring in belief networks. *Artificial Intellig.* 29: 241–288.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan-Kaufmann, San Mateo, CA.

Pope, J.G. 1972. An investigation of the accuracy of virtual population analyses using cohort analysis. *International Commission of the Northwest Atlantic Fisheries, Research Bulletin* 9: 65-74.

Rowe, A.J. & Boulgarides, J.D. 1983. Decision styles – a perspective. *Leadersh. Organiz. Develop. J.* 4(4): 3-9.

Rowe, A.J. & Watkins, P.R. 1992. Beyond expert systems – reasoning, judgment, and wisdom. *Expert Syst. with Applicat.* 4: 1-10.

Shafer, G. 1981. Constructive probability. *Synthese* 48: 1-60.

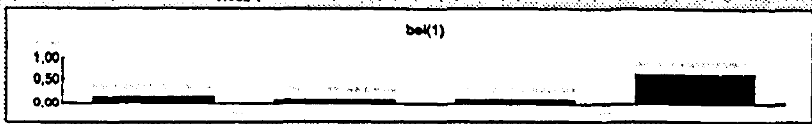
Shafer, G. & Pearl, J. (Eds.) 1990. *Uncertain Reasoning*. Morgan-Kaufmann, San Mateo, CA.

Varis, O. 1992. *F.C. BeNe (Fully Connected Belief Networks): A Spreadsheet Toolkit for Problems with Several, Highly Uncertain and Interrelated Variables. Release β User's Guide.* Helsinki University of Technology, Laboratory of Hydrology and Water Resources Management, Espoo.

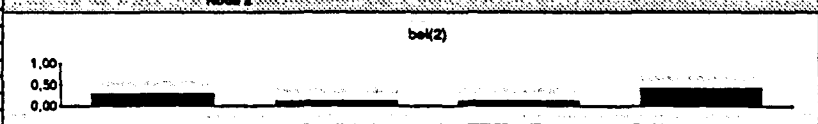
Varis, O. 1993. A belief network methodology for modelling environmental change. *Forthcoming.*

APPENDIX 1

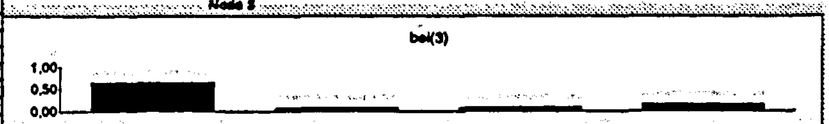
Source code (Microsoft® Excel 4.0) of the example belief network model in Fig. 2.



	C	D	E	F	G	H	I	J	K	L
3			Node 1				Link 2|1, 1|2			
4										
5										
6										
7										
8										
9										
10										
11		p(1)	bel(1)	l(1)			M 2|1 = M 1|2			
12		=D19/(D$19+D$20+D$21+D$22)	=(D12*F12)/(D$12*F$12+D$13*F$13+D$14*F$14+D$15*F$15)	=MMULT(H12:K12,O$12:O$15)		=0,25+0,75*H19	=(1-I$13)/3	=(1-J$14)/3	=(1-K$15)/3	
13		=D20/(D$19+D$20+D$21+D$22)	=(D13*F13)/(D$12*F$12+D$13*F$13+D$14*F$14+D$15*F$15)	=MMULT(H13:K13,O$12:O$15)		=(1-H$12)/3	=H12	=(1-J$14)/3	=(1-K$15)/3	
14		=D21/(D$19+D$20+D$21+D$22)	=(D14*F14)/(D$12*F$12+D$13*F$13+D$14*F$14+D$15*F$15)	=MMULT(H14:K14,O$12:O$15)		=(1-H$12)/3	=(1-I$13)/3	=H12	=(1-K$15)/3	
15		=D22/(D$19+D$20+D$21+D$22)	=(D15*F15)/(D$12*F$12+D$13*F$13+D$14*F$14+D$15*F$15)	=MMULT(H15:K15,O$12:O$15)		=(1-H$12)/3	=(1-I$13)/3	=(1-J$14)/3	=H12	
16										
17		Evidence		Outcome		Link Strength				
18		e1		y1		h2|1 = h1|2				
19		0,1		1		0,6				
20		0,1		2						
21		0,1		3		outcome link				
22		0,7		4		y2=sqrt(y1)				
23										
24										



	M	N	O	P	Q	R	S	T	U
3		Node 2				Link 3|2, 2|3			
4									
5									
6									
7									
8									
9									
10									
11	p(2)	bel(2)	l(2)			M 3|2 = M 2|3			
12	=MMULT(H12:K12,D$12:D$15)	=(M12*O12)/(M$12*O$12+M$13*O$13+M$14*O$14+M$15*O$15)	=MMULT(O12:T12,X$12:X$15)		=0,25+0,75*O19	=(1-R$13)/3	=(1-S$14)/3	=(1-T$15)/3	
13	=MMULT(H13:K13,D$12:D$15)	=(M13*O13)/(M$12*O$12+M$13*O$13+M$14*O$14+M$15*O$15)	=MMULT(O13:T13,X$12:X$15)		=(1-O$12)/3	=O12	=(1-S$14)/3	=(1-T$15)/3	
14	=MMULT(H14:K14,D$12:D$15)	=(M14*O14)/(M$12*O$12+M$13*O$13+M$14*O$14+M$15*O$15)	=MMULT(O14:T14,X$12:X$15)		=(1-O$12)/3	=(1-R$13)/3	=O12	=(1-T$15)/3	
15	=MMULT(H15:K15,D$12:D$15)	=(M15*O15)/(M$12*O$12+M$13*O$13+M$14*O$14+M$15*O$15)	=MMULT(O15:T15,X$12:X$15)		=(1-O$12)/3	=(1-R$13)/3	=(1-S$14)/3	=O12	
16									
17	Outcome		Evidence		Link Strength				
18	y2		e2		h3|2 = h2|3				
19	=SQRT(F19)		non-		0,4				
20	=SQRT(F20)		informative						
21	=SQRT(F21)				outcome link				
22	=SQRT(F22)				y3=exp(y2)/exp(1)				
23									
24									



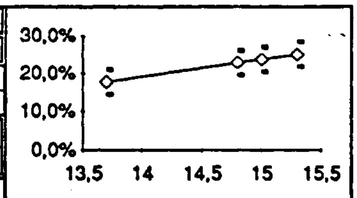
	V	W	X	Y
3		Node 3		
4				
5				
6				
7				
8				
9				
10				
11	p(3)	bel(3)	l(3)	
12	=MMULT(Q12:T12,M$12:M$15)	=(V12*X12)/(V$12*X$12+V$13*X$13+V$14*X$14+V$15*X$15)	=X19/(X$19+X$20+X$21+X$22)	
13	=MMULT(Q13:T13,M$12:M$15)	=(V13*X13)/(V$12*X$12+V$13*X$13+V$14*X$14+V$15*X$15)	=X20/(X$19+X$20+X$21+X$22)	
14	=MMULT(Q14:T14,M$12:M$15)	=(V14*X14)/(V$12*X$12+V$13*X$13+V$14*X$14+V$15*X$15)	=X21/(X$19+X$20+X$21+X$22)	
15	=MMULT(Q15:T15,M$12:M$15)	=(V15*X15)/(V$12*X$12+V$13*X$13+V$14*X$14+V$15*X$15)	=X22/(X$19+X$20+X$21+X$22)	
16				
17	Outcome		Evidence	
18	y3		e3	
19	=EXP(M19)/EXP(1)		0,7	
20	=EXP(M20)/EXP(1)		0,1	
21	=EXP(M21)/EXP(1)		0,1	
22	=EXP(M22)/EXP(1)		0,1	
23				
24				

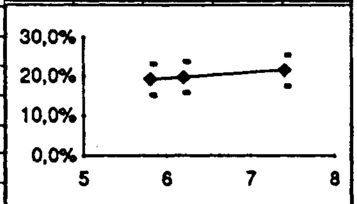
APPENDIX 2

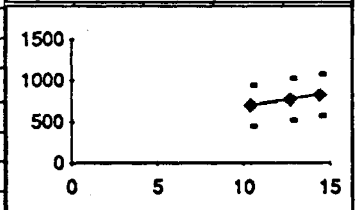
User interfaces of the four steps of the assessment procedure.

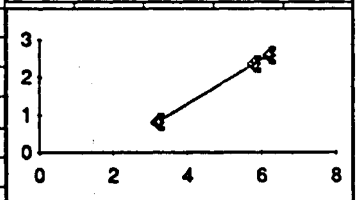
(a) Regression models (cf. Table 2 and Eq. 2). Numbers in bold are inputs.

	1989	1990	1991	1992
T [C]	14,8	15,3	13,7	15
S|T	s(S)	r2	n	
	0,059	0,73	10	
μ(S|T)	22,9%	25,2%	17,8%	23,8%
s(S|T)	3,3%	3,3%	3,3%	3,3%
G [cm]	5,8	6,2	7,4	
S|G	s(S)	r2	n	
	0,059	0,6	10	
μ(S|G)	19,2%	19,8%	21,5%	
s(S|G)	4,0%	4,0%	4,0%	
CPUE (U)	14,4	12,7	10,4	
N:A1|U	s(N)	r2	n	
	263,9	0,23	10	
μ(N1|U)	830,1	774,1	698,3	
s(N1|U)	245,6	245,6	245,6	
G	3,2	5,8	6,2	
F:A1|G	s(G)	r2	n	
	0,298	0,704	9	
μ(F|G)	0,828	2,362	2,598	
s(F|G)	0,173	0,173	0,173	









APPENDIX 2

User interfaces of the four steps of the assessment procedure.

(b) Estimation and tuning of the VPA and linking it with the belief network. μ shows the mean and *cv* shows the coefficient of variation of the prior distribution to the belief network. In that distribution, the regression information is included. *y* shows the three outcomes of the VPA distribution, each with equal probability in the pure VPA calculation. *bel* indicates the posterior probabilities, calculated by the belief network, for those outcomes. μ' and *cv'* are mean and coefficient of variation of the posterior distribution. Numbers in bold are inputs.

1989	μ	y	bel	μ'
S	26,8%	32,5%		26,8%
stock	cv	37,1%		cv'
5230	0,09	41,8%		0,09
0,1	4,115	-0,016		6,389
m	cv	0,007		cv'
F	0,007	0,036		0,006
A0	μ	y	bel	μ'
N	1942	1700		1659
catch	cv	1942		cv'
11,2	0,129	2185		0,248
0,1	0,099	0,674		0,182
m	cv	0,735		cv'
F	0,743	0,817		0,647
A1	μ	y	bel	μ'
N	1250	1450		1251
catch	cv	1628		cv'
805,6	0,122	1807		0,121

Input cv & w		Links			
cv(m)	**0,4**	S|N0	**0,8**	S|S*	**0,5**
cv(C)	**0,2**	F0|N0	**0,8**	F0|F0*	**0,5**
w(S|T)	**0,9**	F0|F1	**0,1**	N0|N1*	**0,8**
w(S|G)	**0,9**	F1|N1	**0,8**	F1|F1*	**0,3**
w(N|U)	**0,9**	F1|F2	**0,1**	N1|N2*	**0,8**
w(F|G)	**0,01**	F2|N2	**0,8**	F2|F2*	**0,1**

1990	μ	y	bel	μ'
S	20,8%	14,1%		19,6%
stock	cv	17,9%		cv'
4389	0,109	21,7%		0,033
0,1	3,637	-0,026		3,224
m	cv	0,015		cv'
F	0,015	0,083		0,015
A0	μ	y	bel	μ'
N	784,6	618,4		832,7
catch	cv	784,6		cv'
11,1	0,218	950,7		0,185
0,1	0,269	0,28		0,417
m	cv	0,363		cv'
F	0,396	0,487		0,352
A1	μ	y	bel	μ'
N	1286	1483		1279
catch	cv	1746		cv'
504,7	0,15	2009		0,16
0,1	**0,3**	1,51		0,297
m	cv	2,13		cv'
F	**2,13**	2,75		2,135
A2	μ	y	bel	μ'
N	706,6	579,7		626,1
catch	cv	706,6		cv'
587,5	0,185	833,5		0,308

1991	μ	y	bel	μ'	
S	20,8%	14,8%		21,3%	
stock	cv	23,0%		cv'	
4014	0,127	31,1%		0,083	
0,1	**0,3**	0,008		0,292	
m	cv	0,011		cv'	
F	**0,011**	0,014		0,011	
A0	μ	y	bel	μ'	
N	922,5	595		906,5	
catch	cv	922,5		cv'	
8,55	0,366	1250		0,081	
0,05	**0,3**	0,17		0,219	0,273
m	cv*	0,24		cv'	cv
F	**0,24**	0,31		0,263	0,263
A1	μ*	y	bel	μ'	μ
N	698,7	514,9		744,9	
catch	cv	699		cv'	
150,9	0,223	883,1		0,229	
0,1	**0,3**	0,454		0,297	
m	cv	0,64		cv'	
F	**0,64**	0,826		0,643	
A2	μ	y	bel	μ'	
N	1099	824,3		971,2	
catch	cv	1099		cv'	
474,1	0,258	1374		0,441	

APPENDIX 2

User interfaces of the four steps of the assessment procedure.

(c) Forecasts and the TAC decision. Means and coefficients of variation of the VPA forecast (μ and cv, respectively), and of the belief network model forecast (μ' and cv', respectively). At the lower right hand corner, the users of the system are allowed to fix target levels for the spawning wild females, and the cumulative probability is shown for the realization of the goal. For more details, see the text. Numbers in bold are inputs.

Section 3 — Forecasts for the present and target years — Section 4 — TAC decision

1992	μ	y	bel	μ'	1993
PSM	23,4%	14,8%		23,2%	
stock	cv	23,0%		cv'	
4000	0,142	31,1%		0,106	

A0	μ	cv	μ'	cv'	A0	μ	cv	μ'	cv'
F	0,016	0,3	0,016	0,292	F	0,025	0,3	0,016	0,292
N	840,1	0,119	834,9	0,091	N	840,1	0,119	834,9	0,091
catch	6,483	0,269	6,461	0,251	catch	6,483	0,269	6,461	0,251

											beginning		autumn	
A1	μ	cv	μ'	cv'	A1	μ	cv	μ'	cv'	wild%	**10%**	**0,4**		
F	0,408	0,273	0,408	0,219	F	0,633	0,363	0,408	0,219	fem%	**60%**	**0,2**		
N	696,9	0,3	684,7	0,08	N	507,6	0,136	503,8	0,108	w spw	30,46	0,383	17,75	0,342
catch	145,6	0,363	143,2	0,176	catch	266,8	0,206	265,8	0,163	w spw'	30,23	0,376	20,69	0,31

A2	μ	cv	μ'	cv'	A2	μ	cv	μ'	cv'	wild%	**10%**	**0,4**		
F	0,992	0,3	0,996	0,297	F	1,538	0,3	1,544	0,297	fem%	**40%**	**0,2**		
N	430	0,245	457,3	0,214	N	252,7	0,345	247,2	0,246	w spw	10,11	0,469	3,19	0,479
catch	241,5	0,297	258,2	0,27	catch	422,4	0,286	416,1	0,159	w spw'	9,887	0,421	3,103	0,446

A3	μ	cv	μ'	cv'	A3	μ	cv	μ'	cv'	wild%	**10%**	**0,4**		
F	1,29	0,3	1,295	0,297	F	1,999	0,3	2,007	0,297	fem%	**40%**	**0,2**		
N	565,6	0,254	498,5	0,386	N	112,3	0,372	118,7	0,356	w spw	4,491	0,484	0,991	0,546
catch	453,6	0,28	402,2	0,404	catch	294,3	0,237	313,8	0,215	w spw'	4,748	0,475	1,04	0,538

											autumn			
1992	μ	cv	μ'	cv'	1993	μ	cv	μ'	cv'	w spw	μ	cv	p(<L1)	p(<L2)
Σcatch	847,1	0,183	810	0,221	Σcatch	990	0,151	1002	0,104	Σ	21,93	0,287	0,029	0,379
TAC-Σc	-47,08	TAC	-10,02	B	TAC-Σc	9,982	TAC	-2,207	B	Σ'	24,83	0,265	0,012	0,231
Δ [%]	106%	**800**	101%	**0,69**	Δ [%]	99%	**1000**	100%	**1,55**	Δ [%]	-12%	L1, L2	**10**	**20**
P(Σc<T)	0,381		0,478		P(Σc<T)	0,527		0,492		K max	**60**	<- figure scale