ELSEVIER

A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data

E. Rivot [a],[*],[1], E. Prévost [a], E. Parent [b], J.L. Baglinière [a]

[a] *INRA, UMR Ecobiologie et qualité des hydrosystèmes continentaux, 65 rue de St Brieuc, CS 84215, 35042 Rennes Cedex, France*
[b] *ENGREF, Laboratoire de gestion du risque en sciences de l'eau, 19 av. du Maine, 75732 Paris Cedex 15, France*

Abstract

The Bayesian state-space modelling framework allows us to derive inferences on stochastic stage-structured population dynamics models from multiple series of sequential observations with measurement or sampling errors. A state-space model works from the coupling of two models. The first one mimics the dynamics by conditional Markovian transitions between successive hidden states. The second model describes the observation process with random errors. Statistical inferences on states variables and parameters are easily performed via Bayesian Monte Carlo Markov chains (MCMC) methods. The flexibility of MCMC methods allows us to analyse a wide range of state-space models with non-linear relationships in the dynamic and observation equations, and non-Gaussian error structure just as well. The Bayesian paradigm is efficient for deriving quantitative diagnostics on a probability based rationale. We illustrate the value of this framework by fitting a stage-structured life cycle model for Atlantic salmon (*Salmo salar*) to a data set resulting from the survey of the population of the river Oir (Lower Normandy, France) between 1984 and 2001. The system dynamics includes non-linear regulation and has a probabilistic structure accommodating for both the environmental and the demographic stochasticity. The observation model consists in capture-mark-recapture experiments for the evaluation of the runs of downstream migrating juveniles (smolts) and upstream migrating spawners, and random sampling for demographic features. A full Bayesian treatment of the model is carried out by means of Gibbs sampling. Outputs of main interest consist in the joint posterior distribution of all the model parameters and state variables such as the smolt and spawner runs.
© 2004 Elsevier B.V. All rights reserved.

Keywords: State-space model; Bayesian; Markov chains Monte Carlo; Population dynamics; Matrix model; Stage-structured; Life cycle; Atlantic salmon

1. Introduction

The main issue required from the Ecological Detective (Hilborn and Mangel, 1997) when dealing with population dynamics is to reconstruct the best representation of the hidden biological process given the knowledge and the data available. Because this representation may provide guidance for action, the Ecological Detective should work with a rigorous

* Corresponding author. Tel.: +33-2-23-48-54-41;
fax: +33-2-23-48-54-40.
 E-mail address: rivot@roazhon.inra.fr (E. Rivot).
[1] On leave from ENGREF, Ecole Nationale du Génie Rural des Eaux et des Forêts, 19 av. du Maine, 75732 Paris Cedex 15, France.

organization of the available knowledge but also a fair recognition of its limitation. The purpose of this paper is to show how the state-space (SS) modelling framework, coupled with modern Bayesian Monte Carlo Markov chains (MCMC) methods, allows us to achieve these goals. The fitting of an Atlantic salmon (*Salmo salar*) stage-structured life cycle model to a series of observations is used as an illustration.

Stage-structured population dynamics models (PDMs) have become standard tools for population dynamics studies (Tuljapurkar and Caswell, 1997; Caswell, 2001). They are grounded on biological fundamentals and provide some flexibility to mimic complex dynamic behaviors. Such models can also be employed to evaluate the performance of a wide range of management options. Numerical ecology is mainly concerned with drawing conclusions from such PDMs when the structure and the parameters of the model are known. For instance, analytical study of matrix PDMs (Tuljapurkar and Caswell, 1997; Caswell, 2001) or empirical methods based on Monte Carlo simulations (Claytor et al., 1993; Nakaoka, 1996; Brodziak et al., 1998) derive fundamental features of the population renewal. However, as models become more complex, they require more parameters and they inevitably become more sensitive to the parameters values. Therefore, before using the model in a predictive mode, a great challenge is to provide reliable parameter estimates and a fair assessment of their uncertainty. This is especially true from a population management and conservation biology viewpoint. When evaluating the expected consequences of management alternatives, uncertainty is a critical component to take into account (Ludwig et al., 1993; Hilborn and Mangel, 1997). This point has received particular attention in the context of population viability analysis (Ludwig, 1996, 1999; Brook et al., 2000) or in the fisheries literature (Francis and Shotton, 1997; Schnute and Richards, 2001). The derivation of valid statistical inferences for stage-structured PDMs is a challenging task for the Ecological Detective. This is also referred to as an "inverse problem" (Gross et al., 2002; Dowd and Meyer, 2003). Usually the model resorts to various sources of information. Two specific features make that linking the model unknowns with the field data at hand within a rigorous inference procedure is easier said than done: The series of data are inherently dependent and the actual underlying population dy-

namics process is stochastic and not directly observed. As a consequence, biologists are often faced with the dilemma of either sacrificing the model realism for the sake of statistical tractability, or adopting a more complex model formulation which often turns to using ad-hoc statistical approximations when estimating its parameters.

The Bayesian SS modelling framework offers a way out of this dilemma. Firstly, the Bayesian statistical paradigm as been advocated as the natural choice for deriving quantitative diagnostics on a probability based rationale. Hence, it is appealing both to the Ecological Detective and to the population manager (Reckhow, 1990; Ellison, 1996; Hilborn, 1997; Punt and Hilborn, 1997; Germano, 1999; Wade, 2000).

Secondly, the SS modelling framework is a general setting which enable us to perform inferences on stochastic dynamic models from series of sequential observations with error in measurements or sampling (West and Harrisson, 1999). The SS framework organises the hypotheses and the available knowledge and recognizes the uncertainty due to its limitation by splitting the problem into two stochastic models. The first model represents the unobservable stochastic process which governs the dynamics by means of conditional transitions between successive hidden states. The second model describes the observation process with random errors. It defines a likelihood function which connects the data to the underlying hidden states and parameters. Statistical inferences are achieved through the coupling of these two stochastic models. This framework handles time dependence between states and data, works even with complex system dynamics, while allowing us to incorporate any source of observation in the statistical inferences.

Thirdly, deriving statistical inferences from a SS models within the Bayesian framework is straightforward. The main difficulty in deriving inferences within the SS framework arises from the fact that it necessitates an averaging of the likelihood function over the entire space of possible intermediate hidden states. Both the frequentist and the Bayesian framework have proposed general methodology to get over this problem, but the Bayesian methodology appears as the most flexible. Within the frequentist framework, the Kalman filter (KF) algorithms (Kalman, 1960) and related approaches are well known procedures to achieve

this integration. The popularity of the KF stems from the existence of an analytical solution to this integration, in the case where both the process and the observation models are linear and with a Gaussian random structure. It easily provides maximum likelihood estimates of parameters and conditional expectations of the state variables given the observations. Many applications of the KF are found in the fisheries literature (Sullivan, 1992; Red and Simons, 1996; Newman, 1998). However, the use of KF algorithms often sacrifices some biological realism. Indeed, the linearity of state-space equations and the normality of the error structure are generally not warranted for PDMs. The "extended KF" (EKF) uses linear approximations of non-linear process equations (Kitagawa, 1987) and its extensions enable handling non-Gaussian error structure. It has been applied to fisheries models by Gudmundsson (1994), Schnute (1994) and Schnute and Kronlund (2002). However, the EKF may be inappropriate when the approximation is too crude compared with the true non-linear dynamic. Computer intensive approaches now provide alternative to these approximations by numerically keeping track of the complete distributions as the filter algorithm is iterated (De Valpine and Hastings, 2002; Punt, 2003). It is not long ago since modern computational methods based on MCMC sampling (Gelman et al., 1995; Gilks et al., 1996) have made Bayesian analysis of large non-linear non-Gaussian SS models easily workable (Carlin et al., 1992). Numerical integration using MCMC methods now enables to solve any multidimensional integrals needed to derive joint and marginal posterior probability density functions (PDFs) of all unknowns for any complex SS models. Examples of application of MCMC sampling to ecological studies can be found in Gross et al. (2002), Link et al. (2002), White and Lubow (2002) and Braak and Etienne (2003). MCMC methods within the Bayesian SS framework have been applied to fisheries related problems, e.g., biomass production models (Meyer and Millar, 1999a; Millar and Meyer, 2000a), growth models (Dowd and Meyer, 2003) or age-structured models (Virtala et al., 1998; Bjornstad et al., 1999; Millar and Meyer, 2000b). The medical field also provides examples which have many similarities with PDMs, e.g., stage-structured model for cells evolution (Guihenneuc-Jouyaux et al., 2000) or models for the dynamics of epidemic diseases (O'Neill, 2002).

In the present paper, we first describe a general setting for the Bayesian SS approach to stage-structured PDMs fitted to series of sequential observations. We point out how the Gibbs sampler, a specific MCMC algorithm, is particularly convenient to estimate the joint posterior PDF of such models. Then, we illustrate the flexibility of the framework by applying the method to a fully stage-structured model for the Atlantic salmon life cycle with multiple life histories. The model describes the dynamics of the numbers of individuals at various life stages, with a discrete annual time step. It includes non-linear regulation and has a probabilistic structure accommodating for both environmental and demographic stochasticity. Two different hypotheses are tested for the transition between migrating juveniles and spawners. The model is fitted to a data set resulting from the comprehensive survey of the salmon population of the river Oir (Lower Normandy, France) between 1984 and 2001. An observation model is constructed to relate the field data to the hidden states at the various life stages. The observation process corresponds essentially to capture-mark-recapture experiments for the evaluation of migrating juvenile and spawner runs and random sampling for demographic features. The results are derived by a full Bayesian treatment of the model, by means of MCMC Gibbs sampling. The ecological significance of our results and the relevance of the Bayesian SS modelling for the Ecological Detective are discussed at the end of the paper.

2. Materials and methods

We denote random variables by capital or Greek letters, and observed values (i.e., data and constants) by lower case letters. $[A|B]$ is the conditional PDF of the variable A knowing B. $A \sim L(\theta)$ means the random variable A follows the PDF L with parameters θ. $A(t)$ denotes a time dependent variable, where t is a discrete index. A^t denotes the time series of successive variables up to time t, $A^t = \{A(1), \ldots, A(t)\}$.

2.1. SS model: a framework for the Ecological Detective

A SS model consists in the coupling of two fundamental equations: the process equation defines the

"true" dynamic of the system with process noise, and the observation equation relates the data to the hidden dynamics of the system.

2.1.1. Process equation

At each time step t, we denote $X(t)$ the state vector, that is the vector of all the unknown quantities of the model which are time dependent. Let θ_1 denote the vector of the parameters for the process equation, that is all the unknowns stationary over time involved in the dynamics. Conditionally on the parameters θ_1, the sequence of unknown states $X^n = \{X(1), \ldots, X(n)\}$ is allowed to follow a multidimensional Markovian chain. The transition Kernel of the Markov process is defined by the population dynamics process equations: the time dependence between states is introduced by successively conditioning the future states on the past and the parameters. Usually, the dependence of the current state $X(t)$ only reduces to that on the most recent state $X(t-1)$. Then, the states process is first-order Markovian. Thanks to conditional independence property, one can split the whole joint PDF into the product of one unit time steps. Once a prior distribution $[X(1)]_0$ is specified for the first state, the process equation can be written as:

$$[X^n|\theta_1] = [X(1)]_0 \prod_{t=2}^{n}[X(t)|X(t-1), \theta_1] \qquad (1)$$

Such a formalization encompasses a wide class of dynamics. Conditional dependence within the Markovian chain is the most general model to describe the causal dependence between events. There is no restriction on the model structure. Conditional PDF $[X(t)|X(t-1), \theta_1]$ may represent any dynamic between the most simple linear deterministic to the most complex non-linear and non-Gaussian stochastic one. One generally distinguishes between the environmental and the demographic stochasticity. The former stands for the unpredictable environmental variability, and the later results from the aggregation of stochastic behavior of each individual. More complicated dynamics, involving time dependence between variables greater than one unit time step, can be written under the form of a one order Markovian model (1) as follows. Let us consider the model where the state $X(t)$ depend upon k ($k > 1$) states in the past, $X(t-1), \ldots, X(t-k)$. Let us de-

note $\Omega(t) = (X(t), \ldots, X(t+k-1))$, and $[\Omega(1)]_0$ the prior PDF assigned to the series of states $X(1)$, $\ldots$, $X(k)$. Then, the full process equation can be written as:

$$[\Omega^n|\theta_1] = [\Omega(1)]_0 \prod_{t=2}^{n-k+1}[\Omega(t)|\Omega(t-1), \theta_1] \qquad (1')$$

2.1.2. Observation equation

The observation equation links the available data to the underlying dynamics. We denote $y(t)$ the quantities actually observed at time t. The observation model is any probabilistic mechanism which would generate data given the system states and parameters. It defines the likelihood function, which gives the probability of the series of observations y^n conditionally on the actual states X^n and on the parameters related to the observation model denoted θ_2. Conditionally on state $X(t)$ and parameters θ_2, the observations $y(t)$ are mutually independent, and the likelihood can be written as:

$$[y^n|X^n, \theta_2] = \prod_{t=1}^{n}[y(t)|X(t), \theta_2] \qquad (2)$$

This framework is general and enables us to incorporate data of different nature and from different sources to perform inferences. In the simplest case, $y(t)$ is a noisy measure of the state $X(t)$. But observation $y(t)$ may be related to the states of interest by a more complex stochastic process (e.g., a capture-mark-recapture model). When the state variables are not directly accessible to measurement, latent variables are introduced in $X(t)$ as intermediate step for conditional modelling to relate observations to the parameters. The dimension of the observation vector can be different between years (because of missing data or different observation process) and can be different from the dimension of the state vector.

2.1.3. Introduction of covariates

The framework described above can be extended to the case where covariates are introduced in the model. Covariates can model the influence of some forcing variables which cannot be controlled (e.g., temperature) or of some control variables (e.g., harvesting) which are of special interest from a decision analysis perspective. Let us denote $U^n = \{U(1), \ldots, U(n)\}$ the series of the forcing or control variables. The process

Eq. (1) then becomes:

$$[X^n|U^n, \theta_1]$$
$$= [X(1)|U(1)]_0 \prod_{t=2}^{n} [X(t)|X(t-1), U(t), \theta_1] \quad (1'')$$

These variables $U(t)$ may be considered as known without error, and the observation Eq. (2) is not modified. But the $U(t)$ can also be latent variables related to some observation denoted $z(t)$, in such a case the observation equation at time t can be written as:

$$[y(t), z(t)|X(t), U(t), \theta_2] \qquad (2')$$

Introduction of covariates in SS modelling will not be further considered in this paper.

2.2. Bayesian inferences from a SS model

2.2.1. Bayesian point of view naturally switches from modelling to statistical inferences

The likelihood principle is a common foundation for both the frequentist and the Bayesian statistical inferences. Both approaches in the SS model framework seek to connect the dynamic model $[X^n|\theta_1]$ with the observation model, which defines the likelihood $[y^n|X^n, \theta_2]$. The inferences simultaneously take into account the dynamic relation between the variables, the process stochasticity and the observation errors.

Frequentist statisticians consider that the parameters are fixed but unknown. Let us denote $\theta = (\theta_1, \theta_2)$ the vector of all parameters in the model. Inferences within the frequentist paradigm are based on the marginal likelihood (3). The basic idea is to get rid of the latent variables by performing the intricate integration of the likelihood over the entire state-space:

$$[y^n|\theta] = \int [y^n, X^n|\theta] \, \mathrm{d}X^n$$
$$= \int [y^n|X^n, \theta] \cdot [X^n|\theta] \, \mathrm{d}X^n \qquad (3)$$

Kalman filtering (Kalman, 1960) and related approaches provide procedures for working out the integration analytically. But numerical algorithms are needed in most cases when the analytical integration is intractable. The marginal likelihood function in (3) is not a PDF of the parameters but the probability of obtaining the observed data for a given value of the

parameters. Thus, from a certain point of view, frequentist approach relies on the modelling reasoning because it proceeds from the "cause" to the "effect": given the model and the parameters, one imagines the mechanisms which could have generated the data. Statistical inferences require an optimization step to find the point value of θ maximizing the marginal likelihood.

By contrast, optimization has nothing to do with inferences in the Bayesian context. The Bayesian treatment of a SS model only relies on the theory of probability. It reconciles modelling and statistical inferences into a unique and coherent rationale. By simply considering the parameter θ as random and assigning a prior PDF $[\theta]$ to it, the Bayesian approach enables us to express the probability of the cause (i.e., the states X^n and the parameters θ) given the effect (i.e., the observed data y^n which are considered fixed):

$$[\theta, X^n|y^n] \propto [\theta] \cdot [X_n|\theta] \cdot [y^n|X^n, \theta] \qquad (4)$$

The constant of proportionality is such that the posterior PDF integrates to 1. The Bayes paradigm is an information processor. When read from the right to the left, the relation (4) highlights how the data, through the likelihood function, update the joint prior law of $[\theta, X^n]$ into the joint posterior $[\theta, X^n|y^n]$. Statistical inferences about the parameters are made by means of the marginal posterior PDF (5), obtained by integrating the joint posterior PDF over the entire state-space:

$$[\theta|y^n] = \int [\theta, X^n|y^n] \, \mathrm{d}X^n \qquad (5)$$

Bayesian treatment of SS models has several advantages over the frequentist approach. First, the relation (4) highlights how modelling and inferences are mutually assembled within a single consistent probabilistic rationale. The Bayesian framework naturally bridges the gap between the likelihood principle and quantitative judgments about the degree of credibility of parameter and state estimates. Second, the states $X(t)$ are considered random and their posterior PDF then makes sense. This is critical when the state variables or function of them have as much biological interest as the parameters θ. By contrast, the nature of the latent states is very ambiguous to the frequentists themselves: while they are clearly random quantities when writing the system dynamic equation, they have also

the same status as parameters, and as such should be considered as fixed but unknown.

2.2.2. Graphic tools and posterior estimation by means of the Gibbs sampler

A directed acyclic graph (DAG) is a useful visual metaphor of a complex high dimensional Bayesian model (Spiegelhalter et al., 1996). Briefly, a DAG represents random quantities as nodes, which appears as ellipses linked by arrows indicating conditional dependencies. The arrows pointing from a node v indicate the nodes under its direct influence, i.e., the children of v (ch(v)). Arrows running into each node v indicate the nodes which directly influence v, i.e., the parents of v (pa(v)). Each node can be stochastic, and then has its PDF defined conditionally on its parents, or logical when it is issued from a deterministic function of its parents. Solid and broken arrows stand for stochastic and logical dependence, respectively. The Fig. 1 provides an illustration of a DAG for a SS model in its general form.

If a Bayesian model can be represented in a DAG form, then the joint PDF of all nodes, denoted [V] has a unique factorization of the form:

$$[V] = \prod_{v \in V} [v|\text{pa}(v)] \tag{6}$$

In the product (6), some conditional PDFs reduce to [v] for initial nodes which have no parent. This is typically the case for parameters which are assigned a prior PDF. A fundamental consequence of the property (6) is that specifying the prior PDFs on initial nodes and the conditional relations between nodes is sufficient to entirely determine the joint PDF of all the

variables of the model. In the context of SS models, the full joint posterior PDF [X^n, θ, y^n] is entirely defined by the specification of the joint prior PDF of all the parameters and initial states, and of the network of conditional relations between all the variables. These include mainly the dynamic relations between the state variables, and the relations between these variables and the observables.

The Gibbs sampler, a special case among the MCMC algorithms (Gelman et al., 1995; Gilks et al., 1996) uses this property to considerably simplify the estimation of the posterior PDF of complex high dimensional SS models (Carlin et al., 1992). One important feature of the Gibbs sampler is its ability to reduce the problem of generating a D-dimensional random variable (where D denotes the total number of unknowns) to that of drawing a sequence of D one-dimensional variables, a comparatively much simpler task when D is high. Instead of sampling directly from the complex multidimensional full joint posterior, as in other MCMC methods, a Markov chain is generated by cyclically sampling from the one-dimensional full conditional posterior PDF of each random variable in the model (i.e., parameters or hidden states). The representation of a Bayesian model under a DAG form highlights how the Gibbs sampling process is considerably simplified. Given its direct parents pa(v) and children ch(v), each node v is independent of all other nodes in the graph. Therefore, the full conditional posterior PDF of each random variable reduces to its PDF conditionally on its direct parents and children.

2.2.3. The WinBUGS® software

The Gibbs sampling algorithm can be easily implemented under the WinBUGS® software (Spiegelhalter et al., 2003). Several case studies in the fisheries and other fields of applied ecology illustrate WinBUGS® potential (Meyer and Millar, 1999a,b; Link et al., 2002; Dowd and Meyer, 2003). WinBUGS® avoids the implementation of a case specific Gibbs sampling algorithm by alleviating the requirement to derive by hand the full conditional PDFs. It only requires the declaration of the Bayesian model, which consists of the prior density of unknown variables, the sampling distributions relating the data and the variables and the deterministic equations linking the variables.
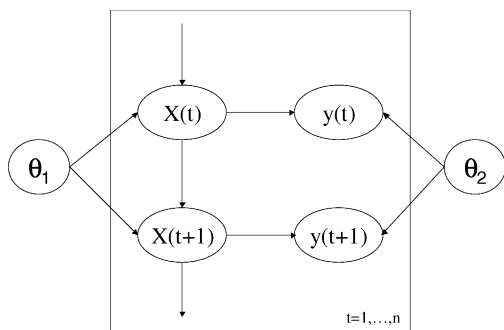

Fig. 1. Directed acyclic graph (DAG) for a general state-space model corresponding to Eqs. (1) and (2) in the text.

2.3. Application: fitting an Atlantic salmon stage-structured life cycle model

We applied this theoretical framework to fitting a stage-structured life cycle model for Atlantic salmon to a data set issued from the survey of the salmon population of the Oir River, Lower Normandy, France. On this study site, both the juvenile and the spawner migrations are monitored thanks to a double downstream and upstream trapping system (see Rivot and Prévost, 2002 for more details). In this section, we detail the components of the $X(t)$, $y(t)$, θ and their relations in Eqs. (7)–(20). The aim is to provide the full DAG of the SS model which is nothing but a graphical translation of the Eqs. (7)–(20) ready for statistical inferences.

2.3.1. Life cycle of Atlantic salmon

The model focuses on two of the fundamental events in the life history of Atlantic salmon: smoltification and spawning migration (Fig. 2). Spawning occurs in the river in late fall and early winter and most spawners die shortly after their first reproduction. Eggs hatch un-

der gravels and alevins emerge during the next spring. After emergence, young-of-the-year salmons (denoted 0+) spend one or several years in the river before undergoing smoltification in spring. At this time, they migrate downstream into the sea as smolt. Adults return to their home river for spawning after one or several winters spent in the sea. Several simplifying hypotheses have been made. We only consider the four main life histories encountered in the northwest of France. A major portion of the juveniles smoltify at one year of river age (1+ Smolt), in the spring following that of their emergence, leaving behind the smaller juveniles (1+ Parr) to spend an additional year in the river before seaward migration (2+ Smolt). Fish from the two smolt age class can return as freshwater spawners after one or two winters spent at sea. One and two sea winter spawners issued from 1+ smolts are denoted 11 SW and 12 SW, respectively, and those issued from 2+ smolts are denoted 21 SW and 22 SW, respectively. Other possible life histories which are rare or have negligible influence on the population dynamics are not represented. In particular, the precocious maturation of male in freshwater before seaward migration
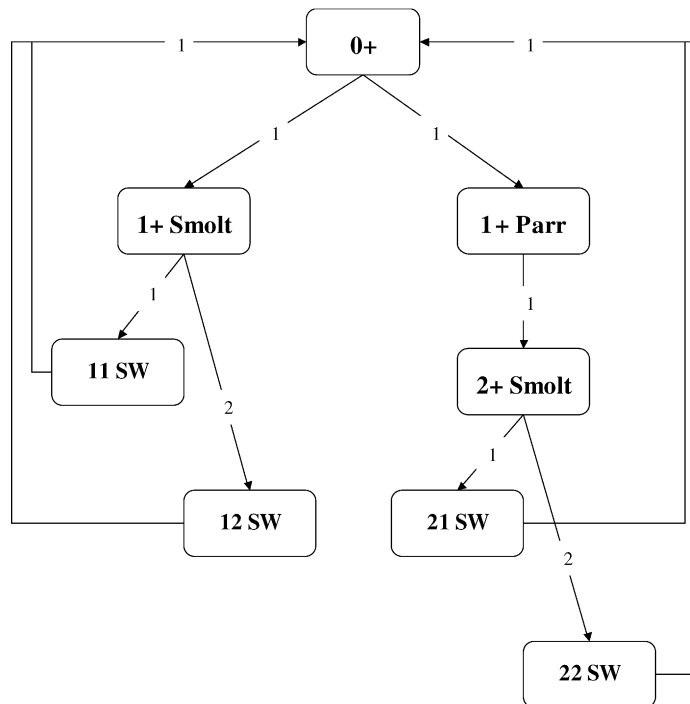

Fig. 2. Simplified life cycle model for Atlantic salmon (*S. salar*) (see text). Number on each transition indicate the time lag (year).

is not explicitly modeled, adults spending more than two winters at sea are ignored and all spawners are consider to die after first spawning. Three other simplifying hypotheses have been made: (i) we suppose the life histories is not an inheritable character; (ii) we use a common life cycle for male and female; (iii) we assume that all eggs carried by females are spawned and fertilized. Males are not considered as limiting for the reproduction success.

2.3.2. Dynamic equations

The choice of dynamic equations has been partly driven by the data available. The population is divided into the two main stages which are quantified in the survey of the Atlantic salmon population of the Oir R.: downstream migrating smolts and upstream migrating spawners. The 0+ and 1+ Parr development stages are not represented such as all parameters and development stages are identified from data alone. As a consequence, transitions between stages are not always one year step. The number of 1+ and 2+ smolts are denoted S_1 and S_2, respectively. The number of spawners are denoted R_{rs}, with the indices standing for the river age ($r = 1, 2$) and the sea age ($s = 1, 2$) respectively. One sea winter and two sea winters spawners (irrespective of the river age) are denoted $R_1 = R_{11} + R_{21}$ and $R_2 = R_{12} + R_{22}$, respectively.

Spawners → eggs. The number of eggs spawned by the adults returning year t, $W(t)$, is a deterministic function of the number of one and two sea winters females (irrespective of the river age) that spawn, denoted $F_1(t)$ and $F_2(t)$, respectively, and of the mean fecundity of these females, considered constant, denoted fec_1 and fec_2, respectively (7). According to Prévost et al. (1996), $fec_1 = 4635$ and $fec_2 = 7965$ eggs per female.

$$W(t) = F_1(t)fec_1 + F_2(t)fec_2 \qquad (7)$$

For the sea age class s, $F_s(t)$ writes as the sum of two terms (8). The first one denoted $f_s^m(t)$ is the observed number of females captured, marked and released. The second term is the number of females that escaped to the trap. This writes as the number of returning spawners $R_s(t)$ minus the observed number of spawners captured at the trap, denoted $c_{Rs}(t)$, times the parameter ρ_{fs} standing for the proportion of female

within the sea-age class s.

$$F_s(t) = f_s^m(t) + \rho_{fs}(R_s(t) - c_{Rs}(t)) \qquad (8)$$

Eggs → smolts. A stochastic density dependent process models the freshwater production of smolts issuing from the reproduction of the spawners returning year t. It relates the number of eggs $W(t)$ to the smolt production $J(t) = S_1(t + 2) + S_2(t + 3)$. To work with variables scaled free from the system studied, the egg deposition and the smolt production are standardized for river size by dividing by the surface area of habitat available for juvenile production, denoted h. According to Prévost et al. (1996), $h = 25\ 229\ m^2$. We denote $W^*(t) = W(t)/h$ the standardized egg deposition, and $J^*(t) = J(t)/h$ the standardized smolt production. Density dependence is modeled by the widely used dome-shaped Ricker curve with unknown parameters (α, β). Environmental variability is introduced via independent and identically distributed log-normal errors, $\varepsilon(t)$, with mean 0 and variance σ_r^2.

$$\begin{aligned} J^*(t)|W^*(t), \alpha, \beta, \sigma_r^2 \\ \sim \log normal(\log(W^*(t)\,e^{(\alpha - \beta W^*(t))}), \sigma_r^2) \end{aligned} \qquad (9)$$

Smolts → 1+ smolts and 2+ smolts. In each cohort, demographic stochasticity acts on the distribution of $J(t)$ into $S_1(t + 2)$ and $S_2(t + 3)$. Each smolt in $J(t)$ has a probability ρ_{s1} to migrate as a 1+ smolt, and a subsequent probability $(1 - \rho_{s1})$ to migrate as a 2+ smolt. The unknown parameter ρ_{s1} is assumed constant over time. $S_1(t + 2)$ is modeled as the result of a binomial process with the parameter ρ_{s1} and $S_2(t + 3)$ is the remaining number of juveniles:

$$S_1(t + 2)|\rho_{s1}, J(t) \sim Binomial(\rho_{s1}, J(t)) \qquad (10a)$$

$$S_2(t + 3) = J(t) - S_1(t + 2) \qquad (10b)$$

Smolts → returning spawners. Two alternative models for the smolt-to-spawner transition were tested. Both rely on the premise of a strict homing of adults to their natal stream. The first model, M_0, assumes that the spawners are issued from the smolts through a multinomial process with transition rates constant over time (11). $\pi_r = (\pi_{r1}, \pi_{r2}, \pi_{r3})$ are unknown parameters with the constraint $\pi_{r1} + \pi_{r2} + \pi_{r3} = 1$ and $D_r(t)$ denotes the total number of dead

smolts (π_{r3} is the mortality rate):

$$(R_{r1}(t+1), R_{r2}(t+2), D_r(t))|S_r(t), \pi_r$$
$$\sim \text{Multinomial}(\pi_r, S_r(t)) \qquad (11)$$

The second model, M_1, encompasses both demographic and environmental stochasticity, by allowing an interannual variability of the transition rates. This is modeled within an exchangeable hierarchical structure in (12a, b) which consists in modelling the $\pi_r(t) = (\pi_{r1}(t), \pi_{r2}(t), \pi_{r3}(t))$ as independent random variables issued a priori from a Dirichlet PDF with three-dimensional unknown parameter Ψ_π (Gelman et al., 1995). We consider that Ψ_π is common for the two river ages.

$$(\pi_{r1}(t), \pi_{r2}(t), \pi_{r3}(t))|\Psi_\pi \sim \text{Dirichlet}(\Psi_\pi) \qquad (12a)$$

$$(R_{r1}(t+1), R_{r2}(t+2), D_r(t))|S_r(t), \pi_r(t)$$
$$\sim \text{Multinomial}(\pi_r(t), S_r(t)) \qquad (12b)$$

2.3.3. Observation equations

The observation model that links the dynamic equations to the data is common to both models M_0 and M_1. The observations series (from 1984 to 2001) has been gathered under an homogeneous experimental design and few data are missing. Each annual survey provides two complementary sources of information about both the population size and the demographic structure (age class and sex-ratio).

Population size at the various life stages. The information available to estimate the size of the smolt and spawner runs is aggregated. It does not distinguish between the age classes. Capture-mark-recapture (CMR) experiments provide data to update the aggregated variables $S(t) = S_1(t) + S_2(t)$ and $R(t) = R_{11}(t) + R_{12}(t) + R_{21}(t) + R_{22}(t)$. For both the smolts and the spawners, the CMR experiments are analogous to the two stages Petersen experiments: a proportion of the migrating population is captured at a partial counting fence and marked for further recapture. The experimental design, the trapping methodology and the CMR model used are described in detail in Rivot and Prévost (2002). We denote $E_S(t)$, $E_R(t)$ the annual trapping efficiencies for smolts and spawners, respectively. We used a hierarchical structure for the prior PDF on $E_S(t)$ and on $E_R(t)$ with hyper-parameters denoted Ψ_S and Ψ_R, respectively (Rivot and Prévost,

Table 1
Capture-mark-recapture data for spawners and smolts by migration year

Year	Spawners							Smolts		
	c	x	m	r	rm	f_1^{m}	f_2^{m}	c	m	rm
1984	167	13	154	22	12	36	18	NA	NA	NA
1985	264	48	216	25	21	56	24	439	NA	NA
1986	130	37	93	9	5	24	16	887	135	91
1987	16	4	12	24	2	1	0	283	31	24
1988	226	43	183	12	12	43	21	307	59	43
1989	235	36	199	56	56	52	11	553	65	43
1990	15	8	7	17	2	3	1	746	38	35
1991	44	0	44	24	23	13	1	151	35	27
1992	31	11	20	9	4	8	0	580	50	43
1993	100	19	81	7	4	12	0	209	26	24
1994	32	14	18	5	1	9	0	329	17	10
1995	109	7	102	46	39	41	2	618	63	53
1996	70	15	55	82	25	18	2	767	76	58
1997	56	22	34	15	12	10	1	205	63	31
1998	34	4	30	36	6	8	0	511	63	31
1999	154	6	148	35	13	42	9	195	59	45
2000	53	0	53	37	4	26	2	1849	300	232
2001	160	1	159	35	31	50	3	688	264	123

See Rivot and Prévost (2002) for more details about the experiments. Spawners: c, number of spawners caught at the upstream trap; x, number which died during manipulation or were removed for experimental use or hatchery production; m, number marked and released ($m = c - x$); r, recapture sample; rm, number of fish marked among r; f_1^{m} and f_2^{m}, number of females, one sea winter and two sea winter, respectively, which are released upstream from the trap. Smolts: c, number of smolts caught at the downstream trap; m, number marked and released; rm, number of marks recaptured; NA, missing data.

2002). Annual CMR data, denoted $\text{cmr}_S(t)$ and $\text{cmr}_R(t)$, enable us to estimate the trapping efficiency and the smolt and spawner runs (Table 1). Each year t, the CMR observation model defines the likelihood functions (17) and (18) (not detailed in this paper):

$$[\text{cmr}_S(t)|S(t), E_S(t)] \qquad (13)$$

$$[\text{cmr}_R(t)|R(t), E_R(t)] \qquad (14)$$

Demographic structure. In addition to the CMR data, sampling among the fish caught at the trap enables us to update the yearly proportions of the age classes which compose the variables $S(t)$ and $R(t)$ and the mean sex ratio in the spawner run. Ageing of the smolts (river age) and the spawners (river and sea age) is done by scale reading (Prévost et al., 1996). The sex of the spawners is identified visually based

Table 2
Data from sampling of the spawner population at the upstream trap and of the smolts population at the downstream trap by migration year

Year	Spawners									Smolts	
	c_R	c_{R1}	c_{R2}	c_{f1}	c_{f2}	c_{R1}^*	c_{R11}	c_{R2}^*	c_{R12}	c_S	c_{S1}
1984	167	141	26	40	21	133	113	26	23	NA	NA
1985	264	203	61	69	42	166	116	45	43	439	232
1986	130	93	37	31	26	80	61	31	24	887	848
1987	16	15	1	1	1	15	13	1	1	283	146
1988	226	182	44	63	31	159	85	38	36	307	282
1989	235	197	38	64	21	183	129	37	31	553	495
1990	15	9	6	3	4	5	3	4	3	746	708
1991	44	42	2	13	1	39	38	2	1	151	101
1992	31	28	3	11	2	26	22	2	2	580	571
1993	100	93	7	22	3	91	85	7	7	209	171
1994	32	30	2	18	2	27	24	2	2	329	323
1995	109	106	3	45	3	105	88	3	3	618	541
1996	70	67	3	27	3	57	48	3	2	767	684
1997	56	55	1	21	1	51	47	1	1	205	186
1998	34	33	1	10	1	27	22	1	1	511	438
1999	154	136	18	43	13	123	105	17	12	195	43
2000	53	49	4	26	2	47	45	4	2	1849	1835
2001	160	151	9	51	3	134	120	6	5	688	636

Spawners: c_R, number of spawners examined for sea age; c_{R1}, c_{R2}, number of one sea winter (1 SW) and two sea winter (2 SW) respectively, among c_R; c_{f1}, c_{f2}, number of females among c_{R1} and c_{R2}, respectively. c_{R1}^*, c_{R2}^*, number of 1 SW and 2 SW, respectively, examined for river age; c_{R11}, c_{R12}, number of fish having spent one year in the river among c_{R1}^* and c_{R2}^* respectively. Smolts: c_S, number of smolts examined for river age; c_{S1}, number of 1-year-old smolts among c_S; NA, missing data.

on the sexual dimorphism of salmon at the time they are caught at the trap. We suppose the age and the sex determination is done without error. The data are given in Table 2. For each demographic feature, the sampling is modeled by independent binomial processes, detailed in Eqs. (15)–(20).

River age in $S(t)$:

$$c_{S1}(t)|c_S(t), \rho_{s1}^*(t) \sim \text{Binomial}(\rho_{s1}^*(t), c_S(t)) \text{ with } \rho_{s1}^*(t) = \frac{S_1(t)}{S(t)} \tag{15}$$

Sea age in $R(t)$:

$$c_{R1}(t)|c_R(t), \rho_1(t) \sim \text{Binomial}(\rho_1(t), c_R(t)) \text{ with } \rho_1(t) = \frac{R_1(t)}{R(t)} \tag{16}$$

River age in $R_1(t)$ and $R_2(t)$:

$$c_{R11}(t)|c_{R1}^*(t), \rho_{11}(t) \sim \text{Binomial}(\rho_{11}(t), c_{R1}^*(t)) \text{ with } \rho_{11}(t) = \frac{R_{11}(t)}{R_1(t)} \tag{17}$$

$$c_{R12}(t)|c_{R2}^*(t), \rho_{12}(t) \sim \text{Binomial}(\rho_{12}(t), c_{R2}^*(t)) \text{ with } \rho_{12}(t) = \frac{R_{12}(t)}{R_2(t)} \tag{18}$$

Females in $R_1(t)$ and $R_2(t)$:

$$c_{f1}(t)|c_{R1}(t), \rho_{f1} \sim \text{Binomial}(\rho_{f1}, c_{R1}(t)) \tag{19}$$

$$c_{f2}(t)|c_{R2}(t), \rho_{f2} \sim \text{Binomial}(\rho_{f2}, c_{R2}(t)) \tag{20}$$

2.3.4. Bayesian inferences

The dynamics of models M_0 and M_1, and the observation model are coupled into the framework of a Bayesian SS model. The following stands for the most general model M_1. The vector of parameters, θ, contains all the time independent unknown

Table 3
Initial probability distribution functions (PDFs) on some components of the state vector $X(1)$, $X(2)$ and $X(3)$ needed to define the initial PDF $[X(1), X(2), X(3)|\theta]_0$

Components	Initial distribution		
	$t = 1$	$t = 2$	$t = 3$
R_{11}	$\sim$Uniform(1,500)	No	No
R_{21}	$\sim$Uniform(1,100)	No	No
R_{12}	$\sim$Uniform(1,100)	//	No
R_{22}	$\sim$Uniform(1,50)	//	No
S_1	$\sim$Uniform(1,3000)	//	No
S_2	$\sim$Uniform(1,300)	//	//

Symbol // indicates the same PDF are used as for $t = 1$. Initial PDFs are independent bounded uniform distributions defined over a sufficiently wide range to minimize their influence on the sequence of successive states.

quantities in the model: the parameters standing for the freshwater recruitment phase, α, β, σ_r^2, and ρ_{s1}, the three-dimensional hyperparameter Ψ_π, the mean proportions of females ρ_{f1}, ρ_{f2}, and the two-dimensional hyperparameters related to the observation process Ψ_S and Ψ_R.

We define the state vector $X(t)$ as all the states variables related to the observation at time t. Thus, some of the components of $X(t)$ directly depend upon previous states $X(t - 1)$, $X(t - 2)$ and $X(t - 3)$. For instance, $R_{11}(t)$ depends on $S_1(t - 1)$, $S_1(t)$ depends on $J(t - 2)$, and $S_2(t)$ depends on $J(t - 3)$ and $S_1(t - 2)$. As a consequence, the elementary terms of the general process equation are of the form (21). They are obtained by combining all the conditional PDFs defined by Eqs. (7)–(12).

$$[X(t)|X(t - 1), X(t - 2), X(t - 3), \theta] \tag{21}$$

Note that by simply defining an increased state vector $\Omega(t) = (X(t), X(t + 1), X(t + 2))$, the time dependence between states $\Omega(t)$ can be written as one order Markovian model (see Eq. (1')). Because of this three steps backward dependence of the states $X(t)$, initial PDFs must be specified on some components of the first three states $X(1)$, $X(2)$ and $X(3)$ in order to define an initial PDF $[X(1), X(2), X(3)|\theta]_0$. Vague (uniform) and mutually independent initial PDF were assigned to each component (Table 3). Finally, the general process equation of the model writes:

$$
\begin{aligned}
[X^n|\theta] = {} & [X(1), X(2), X(3)|\theta]_0 \\
& \times \prod_{t=4}^{n} [X(t)|X(t - 1), X(t - 2), X(t - 3), \theta]
\end{aligned}
\tag{22}
$$

Each likelihood term $[y(t)|X(t),\theta]$ is obtained by the product of the stochastic observation relations (13)–(20). The observation equation can simply be written as the product of all the likelihood terms $[y(t)|X(t), \theta]$ for all years.

To complete the full specification of the Bayesian model, the prior PDF on parameters must be specified. For simplicity, parameters are set a priori independent. The joint prior PDF for θ is the product of the priors PDFs for each component. Although Bayesian statistics allow prior knowledge to be introduced through so called "informative prior", we rather used uninformative proper prior PDFs in order to reflect our initial lack of knowledge and to let the data speak for themselves (Table 4).

The full joint posterior PDF can be written as:

$$
\begin{aligned}
[\theta, X^n|y^n] \propto {} & [\theta] \cdot [X(1), X(2), X(3)|\theta]_0 \\
& \times \prod_{t=4}^{n} [X(t)|X(t - 1), X(t - 2), X(t - 3), \theta] \\
& \times \prod_{t=1}^{n} [y(t)|X(t), \theta]
\end{aligned}
\tag{23}
$$

2.3.5. Model comparison and model checking

We used the Bayes factor (BF) to compare the two alternatives formulations for the smolt-to-spawner transition in M_0 and M_1. We calculated $BF_{1,0}$ defined as the ratio of the marginal likelihood of the data y^n under the model M_1 to that obtained under the model M_0, denoted $[y^n|M_1]$ and $[y^n|M_0]$ respectively. It measures the relative credibility of model M_1 compared with the model M_0 given the data. According to Kass and Raftery (1995), $BF_{1,0}$ was estimated by the ratio (25), where $[\widehat{y^n|M_i}]$ is an MCMC approximation of $[y^n|M_i]$ in (24). It is easily calculated from a sample $\{\theta^{(1)}, \ldots, \theta^{(G)}\}$ drawn in the posterior PDF of θ in the model M_i. It is well known that the MCMC estimate of BF in (25) may be quite unstable. This is because the harmonic mean of the likelihood function is highly sensitive to the small values of the likelihood which may occur during the MCMC simulation.

Table 4
Prior probability density functions (PDFs) on parameters for both models M_1 and M_0

Parameters	Form of the prior	
	M_0	M_1
α	$\sim$Uniform$(-10, 5)$	//
β	$\sim$Uniform$(0, 2)$	//
σ_r^2	$\propto 1/\sigma_r^2$ on $[1.0E - 6, 1.0E + 6]$	//
ρ_{s1}	$\sim$Beta$(1/2, 1/2)$	//
$\psi_\pi = (\psi_{\pi1}, \psi_{\pi2}, \psi_{\pi3})$	No	$\psi_{\pi i} \sim$ Gamma$(0.01, 0.01)$ for $i = 1, 2, 3$
$\pi_1 = (\pi_{11}, \pi_{12}, \pi_{13})$	$\sim$Dirichlet$(1, 1, 10)$	No
$\pi_2 = (\pi_{21}, \pi_{22}, \pi_{23})$	$\sim$Dirichlet$(1, 1, 10)$	No
ρ_{f1}	$\sim$Beta$(1/2, 1/2)$	//
ρ_{f2}	$\sim$Beta$(1/2, 1/2)$	//
$M_S = \alpha_S/(\alpha_S + \beta_S)$	$\sim$Uniform$(0, 1)$	//
$U_S = 1/(\alpha_S + \beta_S + 1)$	$\sim$Beta$(0.001, 1)$	//
$M_R = \alpha_R/(\alpha_R + \beta_R)$	$\sim$Uniform$(0, 1)$	//
$U_R = 1/(\alpha_R + \beta_R+1)$	$\sim$Beta$(0.001, 1)$	//

Symbol // indicates the same prior PDF are used as for M_0 and M_1. Priors are all independent. Some priors are restricted to a finite domain to ensure they are all proper and thereby avoid improper posterior PDF (Gelman et al., 1995). Priors for α, β, ρ_{s1}, ρ_{f1}, ρ_{f2} and σ_r^2 are standard uninformative (Gelman et al., 1995). Model M_0: A slightly informative Dirichlet prior is used for parameters $\pi_1 = (\pi_{11}, \pi_{12}, \pi_{13})$ and $\pi_2 = (\pi_{21}, \pi_{22}, \pi_{23})$. It gives a priori little weight on values of $\pi_{11}, \pi_{12}, \pi_{21}$ and π_{22} larger than 0.3 and give strong weight to total mortality rate π_{13} and π_{23} around 0.8. Model M_1: Vague gamma priors were assigned to each component of ψ_π, the parameter vector of the Dirichlet PDF. Hyperparameters $\psi_S = (\alpha_S, \beta_S)$ and $\psi_R = (\alpha_R, \beta_R)$ are assigned a diffuse hyperprior. Instead of defining a prior on the initial parameters (α_S, β_S) and (α_R, β_R), we rather set a non-informative prior on the relevant one-to-one transformed parameters (M_S, U_S) and (M_R, U_R) and go back to initial parameters by the inverse transformation (Rivot and Prévost, 2002).

However, this approximation is likely to give results which are accurate enough for the interpretation in the logarithm scale (Kass and Raftery, 1995).

$$\widehat{[y^n | M_i]} = \left(\frac{1}{G} \sum_{g=1}^{G} [y^n | \theta^{(g)}, M_i]^{-1} \right)^{-1} \tag{24}$$

$$\widehat{B_{1,0}} = \frac{\widehat{[y^n | M_1]}}{\widehat{[y^n | M_0]}} \tag{25}$$

To check the adequacy of the fit of the model to the data, we performed statistical tests analogous to cross-validation (CV) tests. CV is a general method of approximating the errors in forecasting a new, un-observed data point by leaving out some data, fitting the model to the rest and forecasting the data left out (Gelman et al., 1995). We used the following simple-point-deletion (SPD) test adapted to our case study. We proceeded separately for spawners and smolts. For the spawners of each year t in $5, \ldots, n$, we fit the model to the altered data set denoted $y^n_{(t-, R)}$ obtained from the entire data set y^n by deleting the spawners CMR data $(c(t), rm(t))$. The data left out are treated as random variables and are estimated. We start

deletion at $t = 5$ because some age classes of spawners returning the year before $t = 5$ are issued from smolts which are not updated by observations and hence directly depend upon the initialization. We used the same procedure for the smolts. For each year t in $3, \ldots, n$, let $y^n_{(t-, S)}$ denotes the data set obtained from y^n by deleting the CMR data $(c(t), rm(t))$. We start deletion at $t = 3$ because some age classes of smolts migrating the year before $t = 3$ are issued from the reproduction of adults which are not updated by observations. We compared the marginal posterior PDFs of some components of interest of the state vector obtained by fitting the model to the complete data set y^n (denoted the CDS estimate in the following) with that obtained by fitting the model to the altered data set $y^n_{(t-)}$ (the SPD estimate). This provides indication of the ability of the model for one step ahead forecasting. We compare SPD estimates with CDS estimates of the $S(t)$'s and the $R(t)$'s.

2.3.6. Posterior estimation using the WinBUGS® software

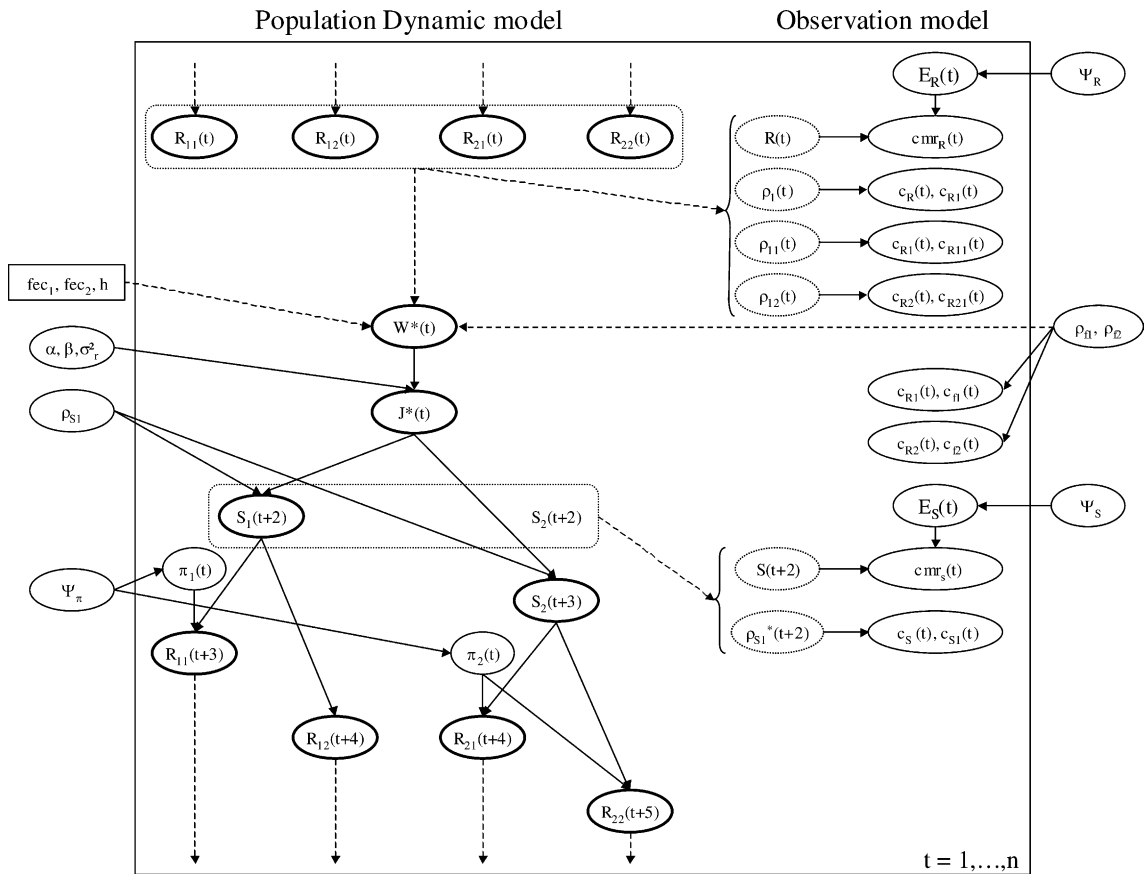Our model is easily implemented under the WinBUGS® software. The DAG of the model is

Fig. 3. Directed acyclic graph for a state-space representation of the Atlantic salmon life cycle model (model M_1). Graphical conventions are described in the text. Ellipse in dotted line indicate state variables which are useful to connect the dynamics with the observation model. fec_1, fec_2, h appear as rectangles because they are constant.

represented in the Fig. 3. The WinBUGS® code is available upon request to the authors. A critical issue is whether the simulated MCMC sample has converged to its ergodic target distribution, which is the posterior PDF. We used the Gelman–Rubin (G–R) diagnostic (Spiegelhalter et al., 2003 and references therein) which is based on the comparison of k chains ($k > 1$) obtained by starting the sampling process at contrasted initial values spread over the variable space. This diagnostic is routinely implemented by WinBUGS® and considered to be sufficient in most practical cases. Three different chains were run from different starting points. The first 1000 iterations of MCMC chains were discarded as an initial burn-in period. Then, further simulations were performed, and convergence of the chains were

checked using the G–R diagnostics. For some of our MCMC simulations, we encountered very poor mixing between chains and very slow convergence. This problem is common in the implementation of Bayesian SSM and arises because of high correlation of the states variables in the time series. Thus, in order to obtain satisfying G–R convergence diagnostics, 30 000 random draws per chains were systematically performed. All inferences were derived using the size-90 000 sample obtained from concatenating the three chains. Marginal posterior PDFs were summarized by statistics directly computed from the MCMC samples: the mean, median and the 5th and 95th percentile which define the 90% Bayesian posterior credibility interval (90% BPI in the following).

3. Results

3.1. Model comparison

We assessed how the instability of the harmonic mean of the likelihood (24) impacts the estimation of the BF. We first computed $\log(\widehat{BF}_{1,0})$ from the size-30 000 MCMC sample obtained from the first chain. Then, a new draw was sampled and added to the previous sample, and we computed $\log(\widehat{BF}_{1,0})$ from the size-30 001 MCMC sample. We iterated the procedure up to the computation of $\log(\widehat{BF}_{1,0})$ from a size-90 000 MCMC sample. We thus obtained 60 000 different values of $\log(\widehat{BF}_{1,0})$. These values vary between 176 and 180. Despite this slight instability, the evidence in favor of the model M_1 is very strong. The hypothesis of constant smolt-to-spawner transition rate can not be supported by the data. In the following, we only report the results obtained with model M_1.

3.2. Spawner returns and smolt output

The posterior estimates of the $R(t)$'s and $S(t)$'s are characterized by a high between year variability (Fig. 4). The posterior means of the $R(t)$'s range between 47 (1991) and 403 (1999). The posterior means of the $S(t)$'s range between 223 (1991) and 2298 (2000). The uncertainty around the estimates of the $R(t)$'s and $S(t)$'s varies between years and partly stems from CMR experiments. Imprecise inferences about $R(t)$ in 1987 are due to sparse mark-recapture data. No data were available for smolts in 1984, what leads to the large imprecision about the estimation of $S(t)$ that year. Detailed interpretation of the CMR experiments and further results concerning the estimation of the trapping efficiencies are given in Rivot and Prévost (2002).

The posterior PDFs of the main demographic features of the returning spawners exhibit moderate within and between year variability. A great part of the variations may be due to chance because of small sample size. Thus, we only comment on the weighted average calculated across the years. Most of the returning spawners are 1 SW fish (posterior mean = 86.5%, 90% BPI = [84.7, 88.4%]). Returning spawners issued from 1+ smolts represent the main part of the spawning population. Those fish represent 81.8% (90% BPI = [79.5, 84.0%]) and 83.5%

(90% BPI = [78.2, 87.9%]) of the 1 SW and 2 SW fish, respectively. The average proportion of female in 1 SW and 2 SW, respectively, ρ_{f1} and ρ_{f2}, have their posterior mean at 0.34 and 0.68, respectively, and very low variances (Table 5). On average, 69.4% of the eggs spawned are from 1 SW females.

3.3. Eggs spawned and subsequent smolts production

The relationship between the number of eggs spawned per unit of area ($W^*(t)$) and the subsequent number of smolts per unit of area ($J^*(t)$) is characterized by a great variability around the mean deterministic Ricker relationship (Fig. 5). The posterior means of the $W^*(t)$'s are highly contrasted between years, ranging from 2.9 (1997) to 30.7 egg m^{-2} (1999). As $W^*(t)$ is calculated from the number of spawners, the time series of the $W^*(t)$'s exhibits within and between years patterns similar to that of the $R(t)$'s. The number of smolts per unit of area $J^*(t)$ corresponds to the smolts which migrate downstream in years $t + 2$ and $t + 3$. J^* are not provided for the last 3 years 1999, 2000 and 2001 because they are not updated by any data. The posterior mean of $J^*(t)$ ranges between 0.005 (1997) and 0.099 smolt m^{-2} (1998). The $J^*(t)$'s exhibit the same pattern of within and between year variations than the $S(t + 2)$'s because the pre-recruitment essentially consists of 1+ smolts.

The main statistics of the posterior PDFs of the recruitment process parameters (α, β, σ_r^2, ρ_{s1}) are given in Table 5. The estimated variance, σ_r^2, is large. This reflects the wide dispersion around the Ricker curve (Fig. 5). The parameters of the Ricker curve (α, β) are highly correlated a posteriori (Fig. 6). The posterior PDF of β seems to be non-null in the neighborhood

Table 5

Main statistics of the marginal posterior distribution for the main parameters in the model M_1 estimated from a size-90 000 MCMC sample

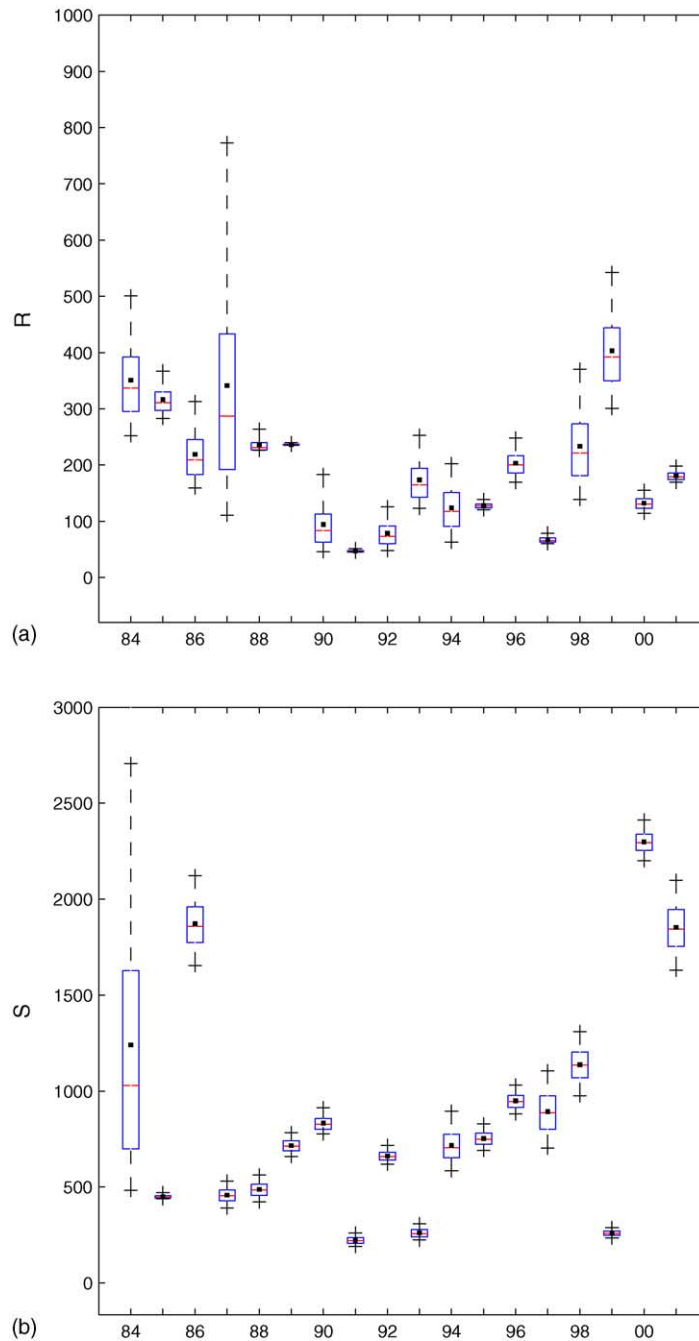|	Mean	5%	25%	Median	75%	95%
α	−5.51	−6.08	−5.77	−5.53	−5.27	−4.86
β	0.04	0.008	0.024	0.038	0.054	0.08
σ_r^2	0.68	0.34	0.47	0.61	0.81	1.25
ρ_{s1}	0.89	0.88	0.89	0.89	0.89	0.90
ρ_{f1}	0.34	0.32	0.33	0.34	0.35	0.36
ρ_{f2}	0.68	0.63	0.66	0.68	0.70	0.72

Fig. 4. Box and whisker plots of a size-90 000 MCMC sample from the posterior distribution of (a) the number of returning spawners (*R*) by year of upstream migration and (b) the number of migrating smolts (*S*) by year of downstream migration. The boxes indicate the inter quartile range and the median. The broken lines extending from each end of a box show the 90% Bayesian posterior interval. (■, mean; +, 5 and 95% percentiles).
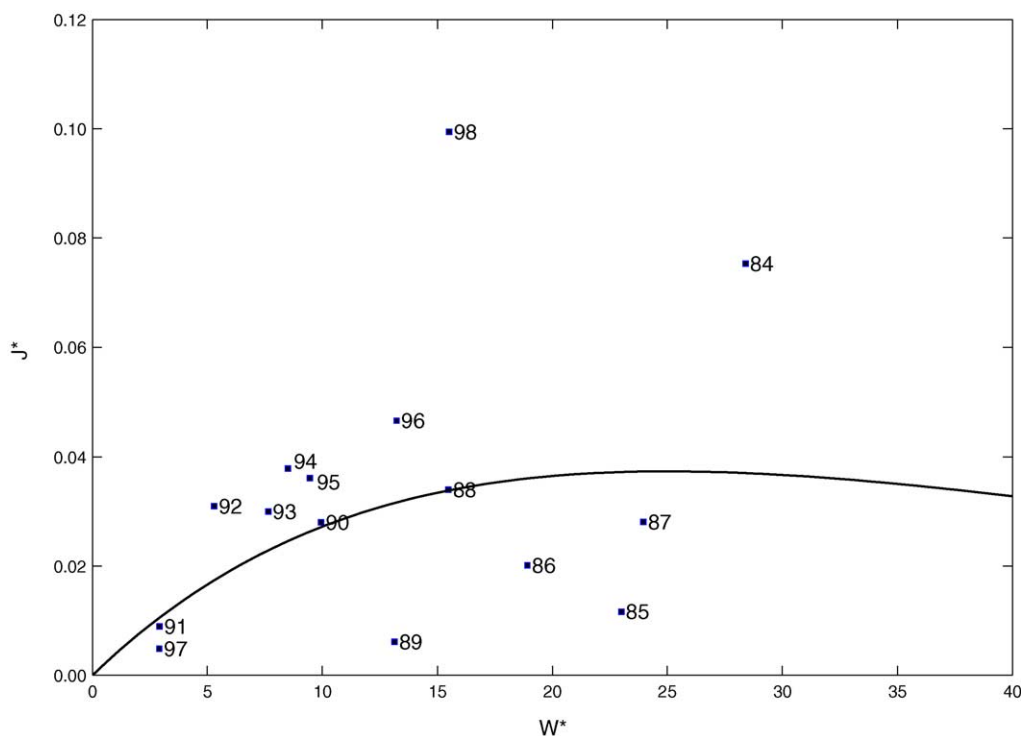
Fig. 5. Relation between the posterior means of the eggs spawned per unit of area (W^*) and of the subsequent number of smolts per unit of area (J^*). The year of upstream migration is indicated near each point. Years 1999, 2000 and 2001 are not represented because the number of smolts cannot be updated due to incomplete data. The Ricker curve (solid line) corresponds to parameters α and β set to their posterior means, -5.51 and 0.04, respectively.

of $\beta = 0$. The posterior mean of $\exp(\alpha)$, the maximum average egg-to-smolt survival rate when the egg deposition tends towards 0, is 0.0044 with 90% BPI = [0.0023, 0.0077]. The average egg-to-smolt survival rate calculated across year is 0.0025 with 90% BPI = [0.0021, 0.003]. The posterior mean of ρ_{s1} indicates that on average, within a cohort, 89.2% of the smolts are 1-year-old. The recruitment success essentially depends upon a single age class of smolts.

3.4. Smolt-to-spawner transition rates

Posterior PDFs of the $\pi_{rs}(t)$'s, i.e., the transition rates from smolt to spawner in the model M_1, vary highly between years (Fig. 7). The posterior PDFs of π_{11}, π_{21}, π_{12} and π_{22} for year 2001 and of π_{12} and π_{22} for year 2000 are highly dispersed because there are no data to update these rates. The posterior means of the $\pi_{rs}(t)$'s vary without any time trend. This between year variability is consistent with the bad fit of the

model M_0 assuming the π_{rs} are constant across years. The estimated means range between 0.06 (1990) and 0.86 (1999) for the $\pi_{11}(t)$'s, between 0.003 (1989) and 0.14 (1987) for the $\pi_{12}(t)$'s, between 0.06 (1990) and 0.83 (1988) for the $\pi_{21}(t)$'s, and between 0.07 (1995) and 0.14 (1997) for the $\pi_{22}(t)$'s. The weighted averages calculated across years for the $\pi_{11}(t)$'s, $\pi_{12}(t)$'s, $\pi_{21}(t)$'s and the $\pi_{22}(t)$'s have their posterior means at 0.18, 0.03, 0.31 and 0.04, respectively. The transition rate from smolt to 1 SW spawner is much higher than the transition rate to 2 SW spawners. The transition rates from 2+ smolt to 1 SW and 2 SW spawner are, on average, higher than the transition rates from 1+ smolt to 1 SW and 2 SW.

3.5. Model checking

The SPD estimates of the $R(t)$'s (Fig. 8a) and the $S(t)$'s (Fig. 8b) are generally more dispersed and higher than the CDS estimates. For the $R(t)$'s, the
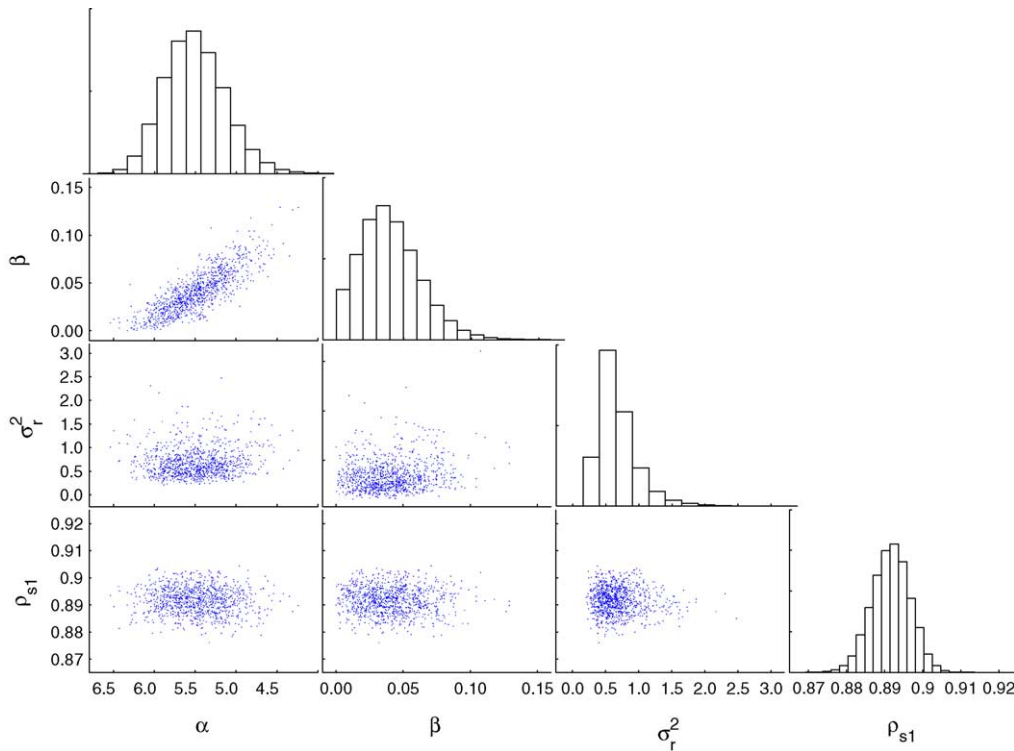
Fig. 6. Scatter plot of a size-1000 MCMC sample from the joint posterior distribution of the parameters (α, β, σ_r^2, ρ_{S1}). Histograms approximate the marginal posterior distribution of each one dimensional parameter.

SPD estimates are inconsistent with the CDS ones for 5 of the 14 years tested. Indeed, for 1991, 1997 and 2000, the 90% BPI's of SPD and CDS estimates do not overlap. Only a partial overlapping is obtained for 1988 and 1996. The SPD estimates of the $S(t)$'s are inconsistent with the CDS ones for 7 of the 16 years tested. For 1987, 1989, 1998, 1999 and 2000, the 90% BPI's do not overlap. A partial overlapping is observed for 1997 and 2001.

4. Discussion

4.1. A powerful data synthesis method

A few attempts to build stage-structured life cycle models for the Atlantic salmon, with parameters obtained from temporal population surveys, have already been made (Cunjack and Therrien, 1998; Dumas and Prouzet, 2003). However, our SS mod-

elling approach is the first which uses a rigorous and coherent probabilistic framework to fit a realistic life cycle model. The statistical analysis enables us to synthesize the 18-years time series of field data from the Oir R. with regards to the structure and the functioning of its salmon population. The demographic features and the parameters of the various between-stage transitions are well determined, and in some cases their inter-annual variation is assessed. The results are consistent with recent reviews about Atlantic salmon life history by Hutchings and Jones (1998) and Dumas and Prouzet (2003), given the Southern European location of the Oir R.

4.2. Limits and improvements of the model

Our model tries to balance biological realism and parsimony to ensure that all parameters are statistically identifiable from the data. Some intermediate development stages (e.g., 0+ juveniles) have not been
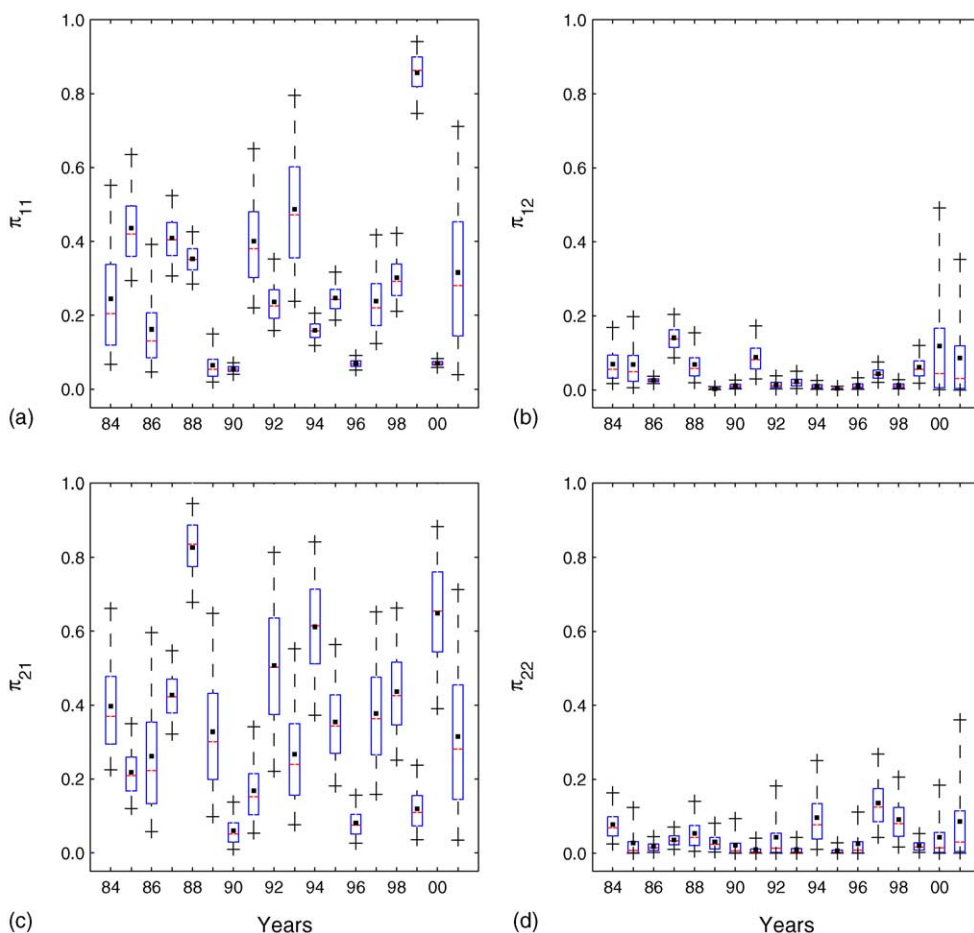
Fig. 7. Box and whisker plot for a size-90 000 MCMC sample from the posterior distribution of the annual smolt-to-spawner transition rates (a) π_{11}, (b) π_{12}, (c) π_{21}, (d) π_{22}, by year of downstream migration of the smolts. The boxes indicate the inter-quartile range and the median. The broken lines extending from each end of the box show the 90% Bayesian posterior interval. (■ , mean; +, 5 and 95% percentiles).

represented, some transitions have been simplified and most of the parameters have been considered to be constant over time. Under the current model specification, the data are informative enough to disentangle observation errors from process stochasticity without a fixed hypothesis on the relative importance of these two sources of noise. But our analysis also stresses the limited ability of our model to reproduce the evolution of the population abundance over time. The marginal likelihood is very low and the SPD tests show that the predictive power of our model concerning the smolt outputs and the spawner returns is rather poor. The cumulative effects of two types of misspecification explain this lack of fit. First, the influence of

non-modelled environmental factors blurs the demographic relation between some stages. Second, some key processes have been missed. The results are useful in setting priorities for model enhancement because they point out the portions of the model which particularly failed and should be improved.

4.2.1. Improving the egg-to-smolt transition

We restrict our analysis to the dome-shaped Ricker curve with constant parameters and an i.i.d. log-normal error structure. The variance around the Ricker function is very large (Table 5 and Fig. 5). This is indicative of a departure from the specified log-normal Ricker model. It is needed to bet-
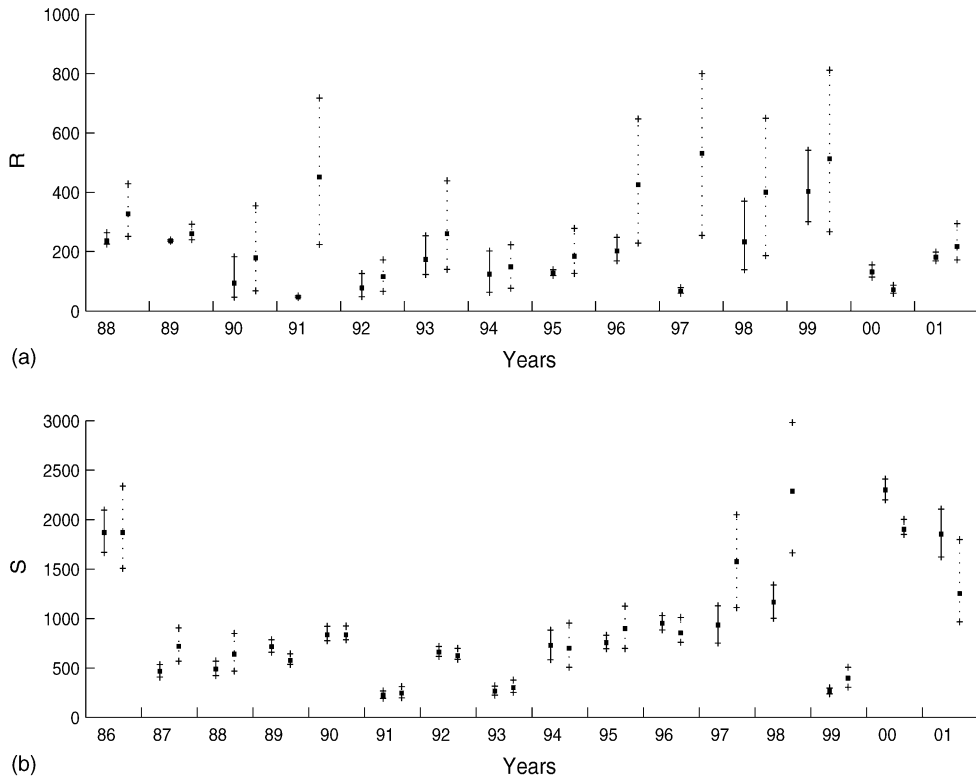
Fig. 8. Comparison between the estimates obtained with the complete data set (solid line) and the simple-point-deletion estimates (broken line) for spawner returns (a) and smolt migrations (b). Posterior distributions are summarized by the posterior mean (■) and the 5 and 95% percentiles (+). The x-axis indicates the year of upstream migration of spawners (a) and downstream migration of smolts (b).

ter understand and to improve the modelling of the density-dependent regulation mechanisms during the juvenile phase. The posterior PDF of the β parameter does not provide strong evidence for over-compensatory mortality of the eggs. Alternative shapes could be tested, such as the Beverton and Holt, the Deriso-Schnute or the Shepherd curve (Schnute and Kronlund, 2002; Milner et al., 2003). But whatever the deterministic function used, the variation around it will remain large. Hence, it is clearly another major challenge to unravel the density-independent mechanisms which influence the egg-to-smolt survival rate (Jonsson et al., 1998; Milner et al., 2003). Accounting for the influence of environmental factors such as the temperature or the flow rate looks appealing. However, it is worth to reveal deceptive and useless in the perspective of using the model for predictions (Walters and Korman, 2001).

4.2.2. Improving the smolt-to-spawner transition

The estimates of the $\pi_{rs}(t)$'s are characterized by a large between year variability. This variability is considered purely random in our model, and it plays a major role in its poor prediction power as revealed by the SPD tests. Some key demographic processes explaining these yearly variations in the smolt-to-spawner transition may have been missed. The spawner returns in the Oir R., and thus the $\pi_{rs}(t)$'s, result from at least three combined processes: (i) the natural mortality of smolts during their stay at sea; (ii) the fishery mortality; (iii) the emigration/immigration processes when the adults return to freshwater. Unfortunately, the data available do not enable us to unambiguously disentangle these three components. Although not explicitly incorporated in our model, our results stress that the emigration/immigration processes are likely to be non-negligible at the spatial scale of our study. Indeed,

available knowledge about the natural and the fishery mortality rates (Hutchings and Jones, 1998; Prévost, unpublished data) indicates that the range of variation of the estimated $\pi_{rs}(t)$'s cannot result from the combination of these two sources of mortality alone. In particular, return rates exceeding 30% (Fig. 7) are extremely unlikely, and a significant net positive immigration of fish issued from other rivers than the Oir R. must be accounted for to explain these high values. Atlantic salmon is known for its homing behavior. But straying, i.e., reproduction in a stream other than the original one, is also recognized as a fundamental life history trait. It can potentially disconnect adult returns from smolt runs (Altukhov and Salmenkova, 1994). The particular context of our study may be favorable to a high level of straying during the estuarine and riverine phase of the spawning migration. Indeed, the Oir R. is a tributary of the Sélune R., which shares a common estuary with the Sélune R. The model in its current form may yield misleading forecasts and diagnostics about the population renewal, because it does not take into account the immigration/emigration process which occur between the different parts of this river network. Further research will aim at extending the model to unravel the patterns of immigration/emigration so as to provide a better understanding of the variation of spawner returns in the Oir R.

4.3. A flexible modelling framework for the Ecological Detective

Beyond the limitations of the illustrative case study discussed above, the Bayesian SS modelling framework we implemented has been revealed to be efficient. It is a valuable contribution regarding the following two important challenges for the Ecological Detective.

4.3.1. Handling the link between PDMs and field data

The SS modelling framework is a highly flexible modelling framework which can handle both process and observation errors within a single consistent probabilistic framework. Thanks to the flexibility offered by Bayesian MCMC methods, complex system dynamics including many latent variables can be combined with sophisticated descriptive statistical models for various sources of information. MCMC algorithms are "model independent", i.e., they can handle any

kind of likelihood function (Gelman et al., 1995; Gilks et al., 1996). Hence, it becomes workable to analyse a wide range of SS models with non-linear relationships in the dynamic and observation equations, and non-Gaussian error structure just as well. The limitations of the traditional Kalman filtering and related approaches relying on linear Gaussian models or linear approximations are overcome. The Ecological Detective can now compare its data with the model he actually wants to use rather than with a model which has been imposed by some mathematical convenience.

A few papers have presented the use of the MCMC sampling methods to estimate the parameters of non-linear SS models within the Bayesian framework. Most often, simple model formulations have been used, based on highly aggregated state variables such as the total biomass of a population (Meyer and Millar, 1999a; Millar and Meyer, 2000a) or based on a single age class (Schnute and Kronlund, 2002). Age-structured models have also been fitted using the same technique (Bjornstad et al., 1999; Millar and Meyer, 2000b). Our paper pointed out that the Bayesian SS modelling framework can be applied to stage-structured life cycle models including multiple life histories and complex stage-structured interactions.

The specific observation model we developed avoids the simplifying assumptions which are made in traditional filtering methods to ensure the observation errors and the process stochasticity can be disentangled. The form of the distribution of observation errors is often assumed known and constant across the years, a log-normal structure being a classical choice (Millar and Meyer, 2000a; Schnute and Kronlund, 2002). The ratio of the variances of the process versus the observation errors is often fixed arbitrarily (Kimura et al., 1996; Schnute and Kronlund, 2002), or an informative prior PDF is set for the variance of the process errors (Millar and Meyer, 2000a). These hypotheses may be impossible to justify or unfortunate. The measurement errors may vary over time, particularly if sample sizes or the data gathering method change during the observation period (Rivot and Prévost, 2002; Maunder and Starr, 2003). Some authors have stressed that the results from statistical filtering methods are sensitive to the ratio of variances (Kimura et al., 1996; Ludwig, 1996, 1999; Schnute and Kronlund, 2002). In the detailed observation

model we used, observation errors are specified on a yearly basis from a realistic stochastic observation model, based on random sampling process and CMR experiments.

Because Bayesian SS modelling frees the Ecological Detective from many modelling restrictions, it encourages the investigation of several competing models. This point is illustrated by the basic but pedagogic model comparison we carried out regarding the smolt-to-spawner transition. It is possible to improve, extend or even replace the dynamic model while keeping the data assimilation scheme unchanged. Supplementing the data does not cause any further complication as it requires only the specification of their relations to the state variables by means of stochastic observation equations. Incomplete series of data can be used as well. The treatment of missing data is straightforward as demonstrated by the SPD approach we implemented.

4.3.2. Setting diagnostics and management decisions in a consistent probabilistic framework

In recent years, Bayesian SS modelling has been increasingly applied to fish stock and fisheries assessment (Punt and Hilborn, 1997; Schnute and Richards, 2001). This is most likely related to the prevalence of uncertainties in the models and the data used in fisheries science, and to the potentially disastrous biological and socio-economical consequences of overlooking major uncertainties. But a fair evaluation of the balance between ignorance and knowledge is also warranted in the management of all renewable resources (Ludwig et al., 1993), or in other fields of applied population dynamics such as the population viability analysis (Ludwig, 1996).

The Bayesian SS modelling framework offers many advantages for dealing with uncertainty and is naturally connected to the assessment of population status and of management actions within a formal decision theoretic approach (Ellison, 1996; Ludwig, 1996; Punt and Hilborn, 1997; McAllister and Kirkwood, 1998; Germano, 1999; Wade, 2000). Posterior PDFs provide readily interpretable probability statements of any quantities derived from the model. Bayesian approach is well suited for simulation based studies that aim to predict possible results of alternative management actions in a probabilistic framework that accounts for all sources of uncertainty. MCMC samples from the joint posterior PDFs of parameters can be used as direct inputs for Monte Carlo forward projections, thus allowing a proper incorporation of our ignorance about the parameters. Bayesian methods for model averaging (Kass and Raftery, 1995) enables to integrate the uncertainty about the model structure. The Bayesian SS modelling framework is also well suited for the introduction of control variables within the dynamic model (West and Harrisson, 1999; Dorazio and Johnson, 2003), and for finding the optimal control conditionally on the data (Berger, 1985).

The challenge for applied population dynamics in the future lies in two areas. First, improving the understanding of the processes regulating populations through the analysis of data issued from long term surveys. Second, incorporating uncertainties into modelling and management decisions making. We bet that the Bayesian SS modelling framework described and illustrated in this paper will bring major future contributions to these issues.

Acknowledgements

References

Altukhov, Y.P., Salmenkova, E.A., 1994. Straying intensity and genetic differentiation in salmon populations. Aquacult. Fish. Manage. 25 (Suppl. 2), 99–120.

Berger, J.O., 1985. Statistical Decision Theory and Bayesian Analysis, 2nd ed. Springer-Verlag, New York, 617 pp. (Springer series in statistics).

Bjornstad, O.N., Fromentin, J.M., Stenseth, N.C., Gjosaeter, J., 1999. Cycles and trends in cod populations. PNAS 96, 5066–5071.

Braak, C.J.F.T., Etienne, R.S., 2003. Improved Bayesian analysis of metapopulation data with an application to a tree frog metapopulation. Ecology 84 (1), 231–241.

Brodziak, J., Rago, P., Conser, R., 1998. A general approach for making short term stochastic projections from an age-structured fisheries assessment model. In: Funk, F., et al. (Eds.), Fishery Stock Assessment Models. University of Alaska Sea Grant College Program, AK-SG-98-01, pp 933–954.

Brook, B.W., O'Grady, J.J., Chapman, A.P., Burgman, M.A., Akçakaya, H.R., Frankham, R., 2000. Predictive accuracy of population viability analysis in conservation biology. Nature 404, 385–387.

Carlin, B.P., Polson, N.G., Stoffer, D.S., 1992. A Monte Carlo approach to non-normal and non-linear state-space modeling. J. Am. Stat. Assoc. 87, 493–500.

Caswell, H., 2001. Matrix Population Models. Construction, Analysis and Interpretation, 2nd ed. Sinauer Associates Inc. Publishers, 722 pp.

Claytor, R.R., Nielsen, G.A., Shelton, P.A., 1993. Using jacknife and Monte Carlo simulation techniques to evaluate forecast models for Atlantic salmon (*Salmo salar*). In: Smith, S.J., Hunt, J.J., Rivard, D. (Eds.), Risk Evaluation and Biological Reference Points for Fisheries Management. Can. Spec. Publ. Fish. Aquat. Sci. 120, 203–219.

Cunjack, R.A., Therrien, J., 1998. Inter stage survival of wild juvenile Atlantic salmon, *Salmo salar* L. Fish. Manage. Ecol. 5, 209–223.

De Valpine, P., Hastings, A., 2002. Fitting population model incorporating process noise and observation errors. Ecol. Monogr. 72 (1), 57–76.

Dorazio, R.M., Johnson, F.A., 2003. Bayesian inference and decision theory—a framework for decision making in natural resource management. Ecol. Appl. 13 (2), 556–563.

Dowd, M., Meyer, R., 2003. A Bayesian approach to the ecosystem inverse problem. Ecol. Model. 168, 39–55.

Dumas, J., Prouzet, P., 2003. Variability of demographic parameters and population dynamics of Atlantic salmon (*Salmo salar* L.) in a southwest French river. ICES J. Mar. Sci. 60, 356–370.

Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecol. Appl. 6 (4), 1036–1046.

Francis, R.I.C.C., Shotton, R., 1997. Risk in fisheries management: a review. Can. J. Fish. Aquat. Sci. 54, 1699–1715.

Gelman, A., Carlin J.B., Stern H., Rubin D.B., 1995. Bayesian Data Analysis. Texts in Statistical Science. Chapman & Hall, 526 pp.

Germano, J.D., 1999. Ecology, statistics and the art of misdiagnosis the need for a paradigm shift. Environ. Rev. 7, 167–190.

Gilks, W.R., Richardson, S., Spiegelhalter, D., 1996. Markov Chain Monte Carlo in Practice. Chapman & Hall, 512 pp.

Gross, K., Craig, B.A., Hutchison, W.D., 2002. Bayesian estimation of a demographic matrix model from stage-frequency data. Ecology 83 (12), 3285–3298.

Gudmundsson, G., 1994. Time series analysis of catch-at-age data. Appl. Stat. 43 (1), 117–126.

Guihenneuc-Jouyaux, C., Richardson, S., Longini Jr., I.M., 2000. Modeling markers of deasease progession by a hidden Markov process: application to characterizing CD4 cell decline. Biometrics 56, 733–741.

Hilborn, R., 1997. Statistical hypothesis testing and decision theory in fisheries science. Fisheries 22 (10), 19–20.

Hilborn, R., Mangel, M., 1997. The Ecological Detective. Confronting Models with Data. Monographs in Population Biology, vol. 28. Princeton University Press, 315 pp.

Hutchings, J.A., Jones, M.E.B., 1998. Life history variation and growth rate thresholds for maturity in Atlantic salmon, *Salmo salar*. Can. J. Fish. Aquat. Sci. 55 (Suppl. 1), 22–47.

Jonsson, N., Jonsson, B., Hansen, L.P., 1998. The relative role of density-dependent and density-independant survival in the life cycle of Atlantic salmon *Salmo salar*. J. Anim. Ecol. 67, 751–762.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Trans. ASME J. Basic Eng. 82, 35–45.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90 (430), 773–795.

Kimura, D.K., Balsiger, J.W., Ito, D.H., 1996. Kalman filtering the delay-difference equation: practical approaches and simulations. Fish. Bull. 94, 678–691.

Kitagawa, G., 1987. Non Gaussian state-space modeling of non-stationary time series (with discussion). J. Am. Stat. Assoc. 82, 1032–1063.

Link, W.A., Cam, E., Nichols, J.D., Cooch, E.G., 2002. Of bugs and birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research. J. Wildlife Manage. 66 (2), 277–291.

Ludwig, D., Hilborn, R., Walters, C.J., 1993. Uncertainty. Science 260, 17–36.

Ludwig, D., 1996. Uncertainty and the assessment of extinction probabilities. Ecol. Appl. 6 (4), 1067–1076.

Ludwig, D., 1999. Is it meaningful to estimate a probability of extinction? Ecology 80 (1), 298–310.

Maunder, M.N., Starr, P.J., 2003. Fitting fisheries model to standardised CPUE abundance indices. Fish. Res. 63, 43–50.

McAllister, M.K., Kirkwood, G.P., 1998. Using Bayesian decision analysis to help achieve a precautionary approach for managing developing fisheries. Can. J. Fish. Aquat. Sci. 55, 2642–2661.

Meyer, R., Millar, R.B., 1999a. Bayesian stock assessment using a state-space implementation of the delay difference model. Can. J. Fish. Aquat. Sci. 56, 37–52.

Meyer, R., Millar, R.B., 1999b. BUGS in bayesian stock assessments. Can. J. Fish. Aquat. Sci. 56, 1078–1086.

Millar, R.B., Meyer, R., 2000a. Non linear state-space modeling of fisheries biomass dynamic using Metropolis-Hastings within Gibbs sampling. Appl. Stat. 49, 327–342.

Millar, R.B., Meyer, R., 2000b. Bayesian state-space modelling of age-structured data: fitting the model is just the beginning. Can. J. Fish. Aquat. Sci. 57, 43–50.

Milner, N.J., Elliott, J.M., Armstrong, J.D., Gardiner, R., Welton, J.S., Ladle, M., 2003. The natural control of salmon and trout populations in streams. Fish. Res. 62, 111–125.

Nakaoka, M., 1996. Dynamics of age- and size-structured populations in fluctuating environments: applications of stochastic matrix models to natural populations. Res. Popul. Ecol. 38 (2), 141–152.

Newman, K.B., 1998. State-space modeling of animal movement and mortality with application to salmon. Biometrics 54, 1290–1314.

O'Neill, P.D., 2002. A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods. Math. Biosci. 180, 103–114.

Prévost, E., Baglinière, J.L., Maisse, G., Nihouarn, A., 1996. Premiers éléments d'une relation stock/recrutement chez le Saumon atlantique (*salmo salar*) en France. Cybium 20 (Suppl. 3), 7–26.

Punt, A., 2003. Extending production models to include process error in the population dynamics. Can. J. Fish. Aquat. Sci. 60, 1217–1228.

Punt, A.E., Hilborn, R., 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish. Biol. Fish. 7, 35–63.

Reckhow, K., 1990. Bayesian inferences in non-replicated ecological studies. Ecology 71 (6), 2053–2059.

Red, W.J., Simons, C.M., 1996. Analyzing catch-effort data by means of the Kalman filter. Can. J. Fish. Aquat. Sci. 53, 2157–2166.

Rivot, E., Prévost, E., 2002. Hierarchical bayesian analysis of capture-mark-recapture data. Can. J. Fish. Aquat. Sci. 59, 1768–1784.

Schnute, J.T., 1994. A general framework for developing sequential fisheries models. Can. J. Fish. Aquat. Sci. 51, 1676–1688.

Schnute, J.T., Kronlund, A.R., 2002. Estimating salmon stock-recruitment relationships from catch and escapement data. Can. J. Fish. Aquat. Sci. 59, 433–449.

Schnute, J.T., Richards, L.J., 2001. Use and abuse of fishery models. Can. J. Fish. Aquat. Sci. 58, 1–8.

Spiegelhalter, D.J., Thomas, A., Best, N.G., 1996. Computation on graphical Bayesian models. In: Bernardo, J.M., et al. (Eds.), Bayesian Statistics, vol. 5. Oxford University Press, pp. 407–425.

Spiegelhalter, D., Thomas, A., Best, N., 2003. WinBugs, Version 1.4. User Manual. MRC and Imperial College of Science, Technology and Medicine (available at: http://www.mrc-bsu.cam.ac.uk/bugs).

Sullivan, P.J., 1992. A Kalman filter approach to catch-at-length analysis. Biometrics 48, 237–257.

Tuljapurkar, S., Caswell, H., 1997. Structured-population Models in Marine, Terrestrial and Freshwater Systems. Chapman & Hall, 656 pp.

Virtala, M., Kuikka, S., Arjas, E., 1998. Stochastic virtual population analysis. ICES J. Mar. Sci. 55, 892–904.

Wade, P.R., 2000. Bayesian methods in conservation biology. Conserv. Biol. 14, 1308–1316.

Walters, C., Korman, J., 2001. Analysis of stock-recruitment data for deriving escapement reference points. In: Prévost, E., Chaput, G. (Eds.), Stock, Recruitment and Reference Points. Assessment and Management of Atlantic Salmon. INRA Editions, pp. 67–94.

West, M., Harrisson, J., 1999. Bayesian Forecasting and Dynamic Models, 2nd ed. Springer-Verlag, 680 pp. (Springer series in statistics).

White, G.C., Lubow, B.C., 2002. Fitting population models to multiple sources of observed data. J. Wildlife Manage. 66 (2), 300–309.