Digging through model complexity: using hierarchical models to uncover evolutionary processes in the wild

M. BUORO*†‡, E. PRÉVOST‡§ & O. GIMENEZ*

*Centre d'Ecologie Fonctionnelle et Evolutive, Campus CNRS, UMR 5175, Montpellier Cedex, France
†Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA
‡INRA, UMR ECOBIOP, INRA/UPPA, Pôle d'Hydrobiologie de l'INRA, St Pée sur Nivelle, France
§Univ Pau & Pays Adour, UMR ECOBIOP, INRA/UPPA, UFR Côte Basque, Anglet, France

Abstract

The growing interest for studying questions in the wild requires acknowledging that eco-evolutionary processes are complex, hierarchically structured and often partially observed or with measurement error. These issues have long been ignored in evolutionary biology, which might have led to flawed inference when addressing evolutionary questions. Hierarchical modelling (HM) has been proposed as a generic statistical framework to deal with complexity in ecological data and account for uncertainty. However, to date, HM has seldom been used to investigate evolutionary mechanisms possibly underlying observed patterns. Here, we contend the HM approach offers a relevant approach for the study of eco-evolutionary processes in the wild by confronting formal theories to empirical data through proper statistical inference. Studying eco-evolutionary processes requires considering the complete and often complex life histories of organisms. We show how this can be achieved by combining sequentially all life-history components and all available sources of information through HM. We demonstrate how eco-evolutionary processes may be poorly inferred or even missed without using the full potential of HM. As a case study, we use the Atlantic salmon and data on wild marked juveniles. We assess a reaction norm for migration and two potential trade-offs for survival. Overall, HM has a great potential to address evolutionary questions and investigate important processes that could not previously be assessed in laboratory or short time-scale studies.

Introduction

Investigation of eco-evolutionary processes in the wild is challenging due to their complex interactions as well as the difficulty in collecting relevant data. Addressing the interplay between life-history traits and associated plasticity remains crucial though, to understand the evolution of life histories, how their variations influence population dynamics (Roff, 1992; Proaktor *et al.*, 2008) and to assess the ability of individuals to adapt to environmental change (Stearns, 1992; Clutton-Brock, 1998; Roff *et al.*, 2006).

Correspondence: Mathieu Buoro, Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA.
Tel.: +1 510 643 9688; fax: +1 510 643 5438;
e-mail: matbuoro@berkeley.edu

Theoretical and manipulative approaches have provided useful information about, for example, state-dependent life-history decisions, reaction norms and trade-offs (Roff, 2002). However, the patterns highlighted in such studies constitute 'potential' evolutionary processes. Working in a controlled environment or isolating part of the process (e.g. a life-history stage) does not capture interactions between ecological and evolutionary processes themselves (Stearns, 1992). Evolutionary studies in the wild are necessary to reveal processes occurring in natural conditions that cannot be easily mimicked in laboratory conditions. In this context, individual lifetime fitness components need to be assessed through continual monitoring of individuals from birth to death, which raises methodological issues, especially in wild animals.

First, the exhaustive monitoring of individuals over time is almost impossible in the wild. The observation of an individual is often a random process, with a probability of detection < 1 (i.e. the probability to observe an individual that is alive and present in the study area). Consequently, life histories and traits are only partially observed: if an individual goes undetected, is it dead or alive? If alive, is it breeding or not? This issue of uncertain detection has long been ignored in evolutionary biology (Clobert, 1995; Cam, 2009; Conroy, 2009), which might have led to flawed inference when addressing evolutionary questions (Gimenez *et al.*, 2008; Hadfield, 2008; Nakagawa & Freckleton, 2008). Second, uncertainty in the observation process may be irreducible when some individual traits cannot be fully observed (e.g. reproductive state) or precisely measured (e.g. size; Catchpole *et al.*, 2008; Hadfield, 2008; King *et al.*, 2008). In this case, field data provide only a noisy or partial measure of the underlying eco-evolutionary processes.

Sound statistical methods, dealing with complex phenomena and uncertainty, are thus needed for addressing eco-evolutionary questions. Hierarchical modelling (HM) has been increasingly recognized as a powerful approach for analysing complex ecological phenomena (Clark *et al.*, 2005; Royle & Dorazio, 2008; Cressie *et al.*, 2009; Link & Barker, 2010; Kéry & Schaub, 2011; Fig. 1). In their review paper about HM, Cressie *et al.* (2009) define HM according to three levels: the data at hand, the underlying process of interest and the parameters governing this process. The process of interest, denoted X, is never fully comprehended. It generates

variability, part of which is due to unknown causes: X is assumed to have some distribution governed by a set of parameters θ_X. The process X is in turn generally not directly or fully observable. It is blurred by randomness in detection or measurement error: data Y are assumed to have some distribution that depends on the process X and on a set of parameters θ_Y governing the random noise in the relationship between Y and X. Combining these two assumptions, HM allows modelling the randomness both in the data and in the underlying process via the joint conditional distribution of Y and X given the set of associated parameters θ_Y and θ_X:

$$[Y, X | \theta_Y, \theta_X] = [Y | X, \theta_Y] \times [X | \theta_X] \quad (1)$$

where $[A|B]$ stands for a set of random variables A distributed conditionally on a set of variables B. This HM formulation is quite generic and encompasses a wide variety of models, including so-called state-space models when the process of interest has a temporal dynamic (Rivot & Prévost, 2002; Clark, 2007; Gimenez *et al.*, 2012). HM offers a clear distinction between the biological process and its observation, and by this mean, it allows a focus on the former while accommodating uncertainties in the latter. Within this general framework, the process can itself be broken into simpler components, of which some are connected to observations, thus facilitating the accommodation of multiple sources of data.

Despite the growing interest in the HM approach for ecological studies, its use in evolutionary ecology is still very limited (Fig. 1). Reasons are unclear but we hypothesize that it stems from the fact that theoretical and/or experimental approaches used in evolutionary studies most often involve organisms that are relatively convenient to monitor. In contrast, ecological studies require monitoring organisms in natural conditions implying numerous constraints in the data collection. We argue that HM is a relevant approach to address evolutionary ecology questions in the wild as it allows the combination of several important ingredients within a single framework: (1) modelling complex phenomena such as complete life histories and associated transitions between states (e.g. alive/dead, breeding/nonbreeding, migrating/resident); (2) integrating underlying evolutionary processes of interest; and (3) accounting for uncertainty in data collection.

Despite the relevance and flexibility of HM, building these models and conducting statistical inference on them is far from trivial (Bolker, 2009; Craigmile *et al.*, 2009; Cressie *et al.*, 2009) and may appear a daunting task for evolutionary biologists with no or little experience in HM. In this paper, we show how the elaboration of a complex HM is facilitated by proceeding step-by-step to the extension of a suite of nested models. We illustrate the HM approach using a case study on Atlantic salmon (*Salmo salar*) dealing with life
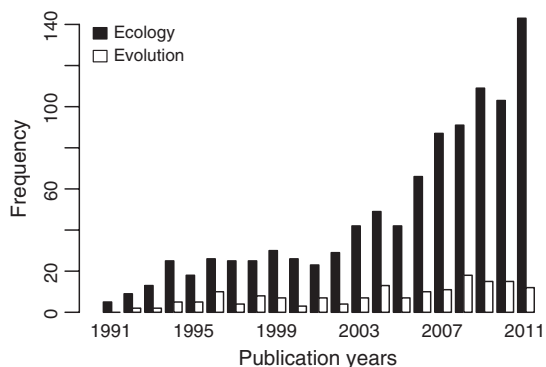

Fig. 1 Time series of the number of papers using or addressing hierarchical modelling (HM) in ecology and evolutionary ecology over the past decade (using ISI web of knowledge citation report; terms employed 'Hierarchical modelling' and 'Multilevel' in subject areas ecology and evolutionary biology). There was no paper on HM in ecology studies until 1991, and the number of papers on HM remains small until 2000; thereafter, it increased constantly. Note the disproportionately small number of papers on HM in evolutionary ecology.

histories of stream-dwelling juvenile salmon in the Scorff River (Southern Brittany, France). We have previously analysed these individual mark–recapture (MR) data to investigate evolutionary trade-offs in the wild while accounting for partial observation with low detectability (Buoro *et al.*, 2010). However successful the HM approach was in terms of results, Buoro *et al.* (2010) did not demonstrate it was worth the complexity and associated modelling effort. Here, we conduct statistical inference at each step and further demonstrate that eco-evolutionary processes would have been poorly estimated or even missed without integrating all life-history events and sources of available information. This case study is well suited for our purpose because (1) Atlantic salmon has a complex life cycle and exhibits a variety of life histories that need to be modelled in a unique framework if we are to understand the processes underlying their variations, (2) juvenile salmon, like fish in general, are not easy to observe in the wild and (3) they have been sampled at several occasions using different sampling methods, hence providing multiple data sets. We first present the HM framework to analyse longitudinal MR data collected at the individual level. Second, we model step-by-step the size-dependent reaction norm for seaward migration and two potential trade-offs, namely a survival cost of migration and a survival cost of reproduction. A suite of four nested models is elaborated that unfold the animal life cycle and successively integrate the various associated sets of observations. The results obtained from this suite of nested models are then compared. Finally, we summarize the lessons learnt from this exercise and discuss the issue of selection among alternative model formulations.

Materials and methods

Longitudinal data at the individual level

In the wild, evolutionary ecologists often rely upon individual MR protocols (Lebreton *et al.*, 1992) for estimating important fitness components, for example survival, dispersal and reproduction (Gimenez *et al.*, 2008). MR data result from the partial observation (detection or not) of events that are generated from a sequence of life stages. HM has been proposed as a flexible framework to deal with MR data (Rivot & Prévost, 2002; Gimenez *et al.*, 2007; Schofield *et al.*, 2007). For the sake of illustration, we first go through a simple example with survival while coping with imperfect detection. Let us focus on the case of an individual *i* between two sampling occasions *t*−1 and *t* (Fig. 2). Conditional on its state at time *t*−1 (alive or dead), this individual may be alive or dead at the following sampling occasion with some probability. Formally, we denote $X_{i,t}$ a binary random variable corresponding to the state of the individual *i* at time *t*, which takes value 1 if the individual

is alive at *t* and 0 otherwise. Then, $X_{i,t}$ given $X_{i,t-1}$ is distributed according to a Bernoulli distribution with probability depending on the survival probability Φ (Gimenez, 2007; Royle, 2008). Note that the survival probability Φ corresponds to the associated parameters θ_X in eqn (1). This leads to the state equation:

$$\left[X_{i,t} \mid X_{i,t-1}, \Phi_{i,t-1}\right] \sim \text{Bernoulli}\left(X_{i,t-1} \times \Phi_{i,t-1}\right) \quad (2)$$
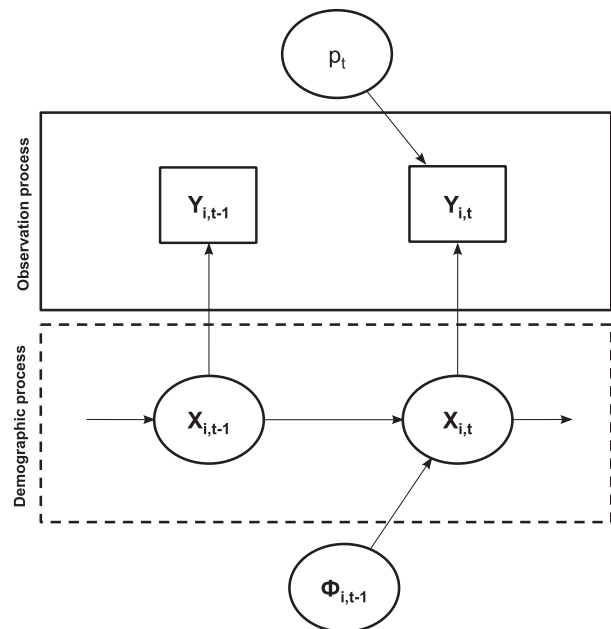

Fig. 2 Graphical representation of a hierarchical mark–recapture model for an individual *i* between two sampling occasions *t*−1 and *t* (see eqns 2 and 3). Each quantity in the model corresponds to a node (e.g. latent states or parameters), and links between nodes show direct dependence. Rectangular and elliptical nodes denote known and unknown quantities, respectively. The first component is a demographic process (dashed box) characterized by a succession of hidden states (solid circles), also called latent states. The demographic process depends on parameters corresponding to transition probabilities between successive states (solid ellipses). The unknown state of individual *i* at time *t* ($X_{i,t}$) is drawn from a Bernoulli distribution depending on its state at time *t*−1 ($X_{i,t-1}$) and the probability of transition between these two states (e.g. the survival probability $\Phi_{i,t-1}$). The observational data (solid square) through the observation process (solid box) are the visible part of the demographic process. Observations are obtained conditionally on latent states and the parameters of the observation process associated (solid ellipses). The observation or not of individual *i* at time *t* ($Y_{i,t}$) is drawn from a Bernoulli distribution that depends on the detection probability p_t at time *t* and conditional on individual *i* being alive at time *t* ($X_{i,t} = 1$). This formulation separates the nuisance parameters (detection probabilities) from the parameters of interest, for example survival probability, the latter being involved exclusively in the state equation. The resulting hierarchical modelling (HM) is a combination of a demographic process and an observation process.

When individual i is alive at t ($X_{i,t} = 1$), it can be observed or not, whereas when dead ($X_{i,t} = 0$), it necessarily goes undetected. We denote $Y_{i,t}$ a binary random variable corresponding to the observation of the individual i at time t, which takes value 1 if the individual is observed and 0 otherwise. Given the state $X_{i,t}$, $Y_{i,t}$ is distributed according to a Bernoulli distribution with probability p_t depending on the detection probability at time t (Gimenez *et al.*, 2007; Royle, 2008). Note that the detection probability p_t corresponds to the associated parameters θ_Y in eqn (1). This leads to the observation equation:

$$[Y_{i,t} | X_{i,t}, p_t] \sim \text{Bernoulli}(X_{i,t} \times p_t) \qquad (3)$$

Besides the estimation of transition probabilities between the demographic states, HMs allow inferring the state (e.g. survivor, migrant or breeder) of any individual at a given time while acknowledging that they may be only partially observed. By extension to the full eco-evolutionary process (e.g. life-history transitions), the life history of every individual can be estimated. In the sequel, we extend the simple example above to a real case study and show how building increasingly complex and comprehensive HMs can limit the risk of flawed inference in eco-evolutionary studies. We focus on the study of phenotypic plasticity and selective survival in the juvenile phase of Atlantic salmon life cycle using MR data.

Step-by-step modelling of juvenile Atlantic salmon MR data

Atlantic salmon life cycle

Atlantic salmon is an anadromous fish species. Its life cycle unfolds both in freshwater and in the ocean (Webb *et al.*, 2007). The juvenile phase takes place in freshwater. In Brittany, it lasts 1 or 2 years (Fig. 3). Thereafter, the fish migrate to the ocean. This migration is accompanied by a smolting process preparing individuals for seawater life. Fish return after 1 or 2 years to their native stream to breed. Males may breed before undertaking seaward migration. During the juvenile phase of the life cycle, individuals adopt alternative life-history tactics. First, they have to decide whether to migrate to the ocean after their first year of life or to reside in the freshwater an additional year. Second, they have to decide whether to mature or not before migrating to the ocean. The latter choice involves only males during their second year in freshwater. These life-history tactics depend on, and modify, the way energy is acquired, stored and used by individuals (Thorpe *et al.*, 1998). Evolutionary trade-offs are thus expected in this species, that is, migrating and maturing costs for survival.

In the following, we use the term '0+' for individuals of < 1 year of age in freshwater, '1+' for those of more than 1 year of age and '2+' for those of more than 2 years of age. Juveniles are named 'parr' if resident in freshwater and 'smolts' when they migrate to the sea.
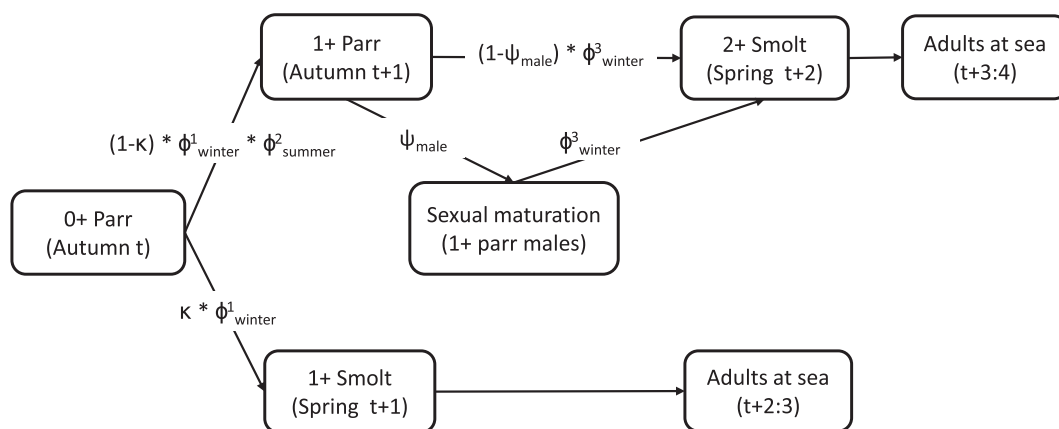

Fig. 3 Life cycle of the Atlantic salmon in the Scorff, Brittany (France). Reproduction occurs in freshwater in December, and eggs are buried in the river bed gravel. Fry emerge from the spawning in early spring. After a few months of life, juveniles, then called '0+ parr', choose between migrating to sea the following spring (1+ smolt stage) with a probability κ or staying another year in freshwater (1+ parr) with a probability $1-\kappa$. The probability of winter survival of the 0+ parr between the first autumn and the following spring is Φ_1 winter. The probability of summer survival of the 1+ parr is Φ_2 summer. Some of the males remaining in freshwater become sexually mature at the 1+ parr stage with a probability of maturing ψ. The probability of winter survival of the 1+ parr between the second autumn and the following spring is Φ_3 winter. Virtually, all surviving juveniles (previously mature or not) will migrate to the sea in the following spring (2+ smolt). Migration to the sea is accompanied by physiological, morphological and behavioural changes (i.e. smolting process), which prepares individuals for seawater life.

Study site and MR data collection

The Scorff River is a small coastal river of Southern Brittany (France). Atlantic salmon colonization is essentially restricted to the main river over a 50-km stretch. In autumn 2005, 0+ parr were sampled by electrofishing at 39 stations along the main course of the Scorff. Every fish captured was measured (fork length) and individually marked with a passive integrative transponder (PIT) for subsequent identification. In spring 2006, downstream migrating 1+ smolts were captured at two successive traps located at the lower end of the river system below all sites where juveniles were marked. Untagged fish caught at the upstream trap (Leslé Mill) were marked by removing a small piece of a pelvic fin. At the second trap (Princes Mill) situated downstream, all individuals previously fin-clipped were identified. In autumn 2006, the 1+ parr were sampled by electrofishing. Marked fish were identified and untagged fish were PIT-tagged. Sexually maturing and already spermating males were detected by pressing their belly. In spring 2007, the 2+ smolts were trapped, checked for PIT tags and fin-clipped if unmarked. These data are summarized in Table S1. More details on the MR protocol can be found in the study by Buoro *et al.* (2010).

Nested statistical models

Our HM approach combines a demographic process model and an observation model. Both are made up of several components, each corresponding to a life-history transition or a source of information. Life-history events are binary and modelled as random variables with Bernoulli distribution (see eqns 1 and 2 above; Table S2). To illustrate the usefulness of increasing complexity and comprehensiveness of HM, we incorporate the different parts in four steps (noted A–D) sequentially and compare the results between steps.

Each model is represented in two ways. First, we opt for a directed acyclic graph (DAG) that is convenient to display HM conditioning structure (Lunn *et al.*, 2000). Second, we provide the corresponding equations needed to describe the model based on extensions of eqns (2) and (3).

Model A: 0+ Parr to 1+ smolt stage transition. Demographic process. We start with model A focusing on the first two life-history events in the first year of life of Atlantic salmon juveniles (Fig. 3): (1) the decision taken in the first autumn of either smolting the following spring at 1 year of age or to stay an additional year in freshwater (1+ parr) and (2) the survival of the 0+ parr between the first autumn and the following spring (winter survival). The DAG of model A is displayed in Fig. 4, whereas equations are given in Table S2.

First, we assume that age at smolting depends positively on growth during the first months of life in fresh-

water (Baglinière & Maisse, 1993; Thorpe & Metcalfe, 1998). Using the conceptual framework of probabilistic reaction norms (Heino, 2002), we use a logit-linear relationship to represent the link between the individual i probability of smolting at age 1+ (κ_i) and its size at the 0+ parr stage:

$$\mathrm{logit}(\kappa_i) = \alpha_1 + \alpha_2 \times \mathrm{Fl}_i \qquad (4)$$

where Fl_i is the individual fork length (mm) centred on the sample mean. We use a logit-link function to ensure that probabilities lie on [0, 1]. Parameter α_2 controls the influence of size at 0+ parr stage on smolting and corresponds to the selection gradient of the probabilistic reaction norm for smolting.

Second, we assume the decision of smolting at age 1+ modifies survival. During the first winter, future migrants adopt a very different behaviour from those intended to reside an additional year in the river as they try to maximize their growth (Metcalfe & Thorpe, 1992; McCormick *et al.*, 1998). We model this differential survival at the individual level by linking the probabilities of winter survival to the smolting decision.

$$\mathrm{logit}(\Phi_{1,i}) = \beta_1 + \beta_2 \times \mathrm{Parr}_{0,i}^{\mathrm{smolt}} \qquad (5)$$

where $\Phi_{1,i}$ stands for the probability of first winter survival (0+ parr) of an individual i, $\mathrm{Parr}_{0,i}^{\mathrm{smolt}}$ is the smolting indicator taking value 1 if individual i is smolting and 0 otherwise. Parameter β_2 reflects the influence of the decision of smolting on winter survival at 0+ parr stage. For instance, if β_2 is negative, then the winter survival $\Phi_{1,i}$ of a smolting 0+ parr ($\mathrm{Parr}_{0,i}^{\mathrm{smolt}} = 1$) is lower than that of a resident 0+ parr ($\mathrm{Parr}_{0,i}^{\mathrm{smolt}}$), revealing a survival cost of smolting.

Observation. The first recapture event after tagging was the trapping of the 1+ smolts (spring 2006) both at the Leslé Mill with probability pL_1 and at the Princes Mill with probability pP_1. At the individual level, capture was modelled as a Bernoulli distribution (Table S2) with associated probability of capture assumed fixed across individuals.

Model B: Incorporating summer survival and maturation process at 1+ parr stage. Demographic process. To unfold the life cycle, summer survival and maturation decision of the 1+ parr need to be modelled (Fig. 4, model B). Survival of resident 1+ parr between their initial marking in autumn 2005 and their first recapture in autumn 2006 is made of two successive survival events: winter survival (from autumn 2005 to spring 2006) and summer survival (from spring 2006 to autumn 2006). The explicit distinction of these two survival events allows assessing the first winter survival probability of the future 1+ parr, despite the absence of recaptures for 1+ parr in spring 2006.
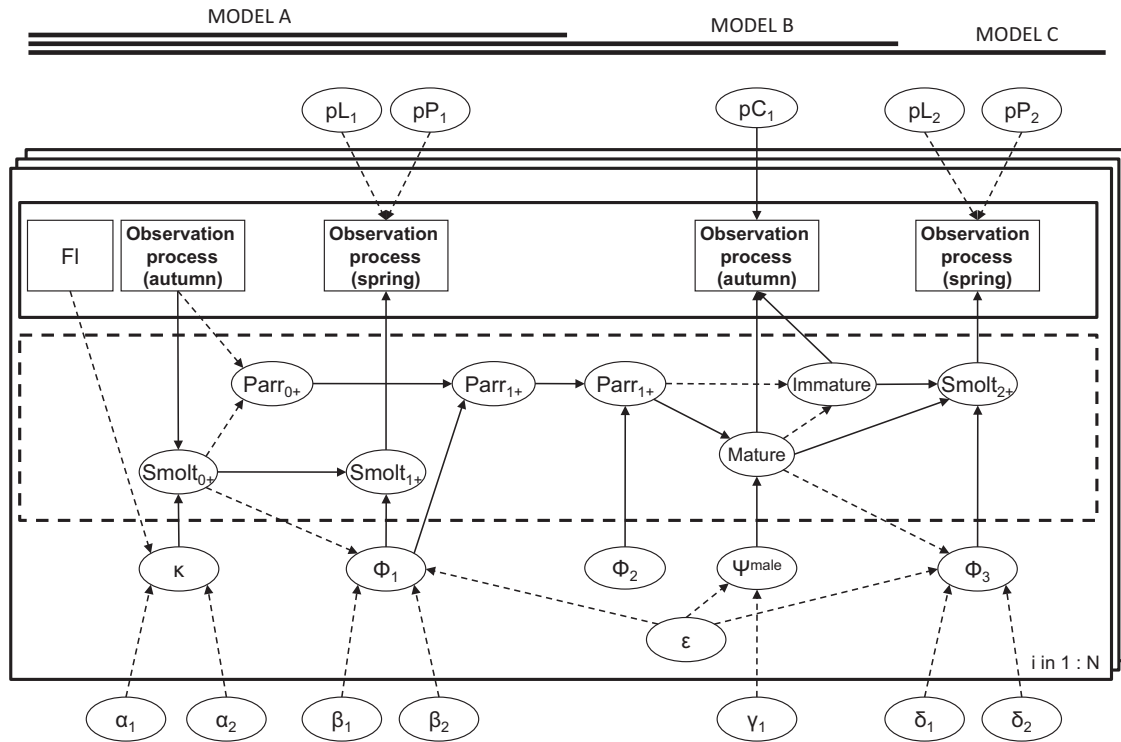
Fig. 4 Graphical representation (DAG) of hierarchical models A–C for life histories and mark–recapture of Atlantic salmon juvenile. Model D was not included as it corresponds to the same observation models than model C with additional data. Notations are given in the text. As described in Fig. 2, we distinguished the demographic process (dashed box) from the observation process (solid box) and hidden states (solid ellipses) from observational data (solid rectangular). Observations are obtained conditionally on latent states and the parameters of the associated observation process (solid ellipses). Each quantity in the model corresponds to a node (e.g. latent states or parameters), and links between nodes show direct dependence. Rectangular nodes and elliptical nodes denote known and unknown quantities, respectively. Stochastic dependence and deterministic dependence are denoted by single arrows and dashed arrows, respectively. Repetitive structures, such as the loop *i* from 1 to *N*, are represented by overlapping frames; DAG, directed acyclic graph.

First, we assume that summer survival of 1+ parr was higher than previous winter survival in the Scorff River (Baglinière *et al.*, 1994). We incorporate this information by specifying summer survival probability $\Phi_{2,i}$ conditionally on winter survival $\Phi_{1,i}$ as

$$\Phi_{2,i} = \Phi_{1,i} + \left(1 - \Phi_{1,i}\right) \times \Delta_{\text{survival}} \quad (6)$$

where Δ_{survival} is an unknown parameter between 0 and 1, which determines the survival difference between first winter and the following summer. From this model B, we introduce an additive random effect ε_i in eqn (5), to account for heterogeneity in individual quality affecting survival. This effect would reflect variations in energy storage among individuals. As survival is energy demanding, higher survival should be related to higher energy stores. ε_i is normally distributed with 0 mean and unknown standard deviation σ_ε. In accordance with Cam *et al.* (2002), this unobservable individual quality affects every subsequent survival event. By this way, for a given individual, having a higher survival probability during the first winter reveals a

better chance to stay alive during the following survival events [i.e. summer survival and second winter survival (see model C)].

Second, we assume that males 1+ parr have to decide to mature or not prior to ocean migration during this summer transition (Thorpe *et al.*, 1998). Maturation state is only observed for male 1+ parr captured in autumn and detected as spermating. We model sexual maturation of males at the 1+ parr stage using a Bernoulli distribution with probability of maturing at 1+ parr stage being the product of the probabilities of sexual maturation for a male and of being a male. We assume that the probability for a 1+ parr to be a male was 0.5 (balanced sex ratio). A male with a high level of energy storage should have a higher probability of maturing at the 1+parr stage (Prévost *et al.*, 1992; Duston & Saunders, 1997). The probability of sexual maturation for a male ψ_i^{male} is assumed to depend on the unobserved individual quality:

$$\text{logit}\left(\psi_i^{\text{male}}\right) = \gamma_1 + \varepsilon_i \quad (7)$$

Observation. The 1+ parr remaining in freshwater (autumn 2006) were captured by electrofishing with probability pC_1 (Table S2). Sampling by electrofishing was conducted in the same sites each year. There is evidence of site fidelity in Atlantic salmon parr (Stickler *et al.*, 2008). In our study, almost all the 1+ parr recaptured were caught at the same station where they were marked at 0+ stage. As a consequence, we consider that the probability of capture for 1+ parr in 2006 was higher than the probability of capture for 0+ parr in 2005, and specified the probability of detection of 1+ parr in autumn 2006, pC_1, conditionally on the probability of detection of 0+ parr in autumn 2005, pC_0:

$$pC_1 = pC_0 + (1 - pC_0) \times \Delta_{\text{capture}} \qquad (8)$$

where Δ_{capture} is an unknown parameter between 0 and 1, which determines the difference between pC_1 and pC_0.

Model C: Incorporating 2+ smolts stage and cost of reproduction for survival. *Demographic process.* Sexual maturation and reproduction of resident males lead to reduced survival (Myers, 1984; Myers & Hutchings, 1987; Fleming & Reynolds, 2004). We model this cost of reproduction at the individual level by linking the probabilities of the second winter survival to the state indicator of the decision of maturation ($\text{Parr}_{1,i}^{\text{mature}}$) (Fig. 4, model C):

$$\text{logit}(\Phi_{3,i}) = \delta_1 + \delta_2 \times \text{Parr}_{1,i}^{\text{mature}} + \varepsilon_i \qquad (9)$$

where $\Phi_{3,i}$ stands for the probability of second winter survival (1+ parr) of individual i. $\text{Parr}_{1,i}^{\text{mature}}$ is the maturing indicator taking value 1 if individual i is mature and 0 otherwise. Parameter δ_2 reflects the influence of the decision of maturing on winter survival at the 1+ parr stage. If δ_2 is negative, then the winter survival $\Phi_{3,i}$ of a maturing male 1+ parr ($\text{Parr}_{1,i}^{\text{mature}} = 1$) is lower than that of a immature 1+ parr ($\text{Parr}_{1,i}^{\text{mature}} = 0$), evidencing a survival cost of reproduction.

Observation. The 2+ smolts (spring 2007) were trapped at the Leslé Mill with probability pL_2 and at the Princes Mill with probability pP_2 according to the same protocol used for the 1+ smolt the previous year (see 'Observation' section in model A).

Model D: Incorporating additional information. Additional information about life-history events and smolt trapping probabilities is available. First, the 1+ parr sampled by electrofishing in autumn 2006, but untagged at 0+ parr stage, were also PIT-tagged. We assume that sexual maturation and second winter survival are the same whether 1+ parr had been tagged or untagged at the 0+ parr stage. As a consequence, we use the same eqns (7) and (9) with common parameters as presented before (models B–C). Second, we take advantage of ancillary fin-clipping data collected at the smolt traps in 2006 and 2007 (see 'Study site and MR data collection' section; Table S1) to improve the estimation of smolt trapping probabilities. We assume that the probability of capture at both traps is the same for PIT-tagged, fin-clipped and untagged smolts.

Statistical inference in a Bayesian framework

We fit our HMs to data by Bayesian statistical inference using Markov chain Monte Carlo (MCMC) sampling. This approach is the most widely used for fitting complex HMs (Cressie *et al.*, 2009; see also Gelman *et al.*, 2004; Ellison, 2004; King *et al.*, 2009; and Link & Barker, 2010 for more details about the Bayesian statistical modelling and associated computational issues). Besides handling the complexity of our model, it facilitates the combination of multiple sources of data, that is, from the PIT-tag protocol and from the ancillary fin-clipping experiment (see previous section). Under the Bayesian approach, the information available in the observed data is conveyed by the likelihood. The latter is combined with the prior distribution of the unknown quantities of the model to obtain their joint posterior probability distribution. This joint posterior probability distribution is the outcome of the Bayesian statistical inference. It provides the comparative degrees of credibility of the possible values of all the model unknowns (i.e. individual states, transition probabilities between states, random effects, observation probabilities and other parameters) conditionally on the observed data (i.e. the PIT-tagging data and the fin-clipping data).

The Bayesian approach requires assigning a prior distribution to model parameters, that is, unknown quantities that do not depend on any other in the model. This prior represents available information apart from the observed data used to derive the likelihood. By default, weakly informative priors are used in order to 'let the observed data speak for themselves' (Van Dongen, 2006; Gelman *et al.*, 2004). Priors can also be used to incorporate information brought by supplementary data, extracted from the existing literature or eliciting expert knowledge (Kuhnert, 2011), hence improving the precision of parameter estimates (McCarthy & Masters, 2005). In agreement with what was known about the species biology, we consider the probability of surviving in freshwater as being neither null nor equal to 1 between two consecutive stages. Consequently, we chose prior distributions such that less importance was given to extreme values of survival probabilities (see 'Choice of prior distribution' section in Appendix S1 for more details). The probability of capture of 0+ parr in autumn 2005 could be assessed from the smolt trapping data in 2006. We summarize this available information by assigning a beta prior distribution to pC_0 (eqn 8) whose parameters are the

number of smolts marked in the previous autumn and captured at smolt traps (Leslé Mill and Princes Mill), and the number of untagged smolts captured at smolt traps. Note that the data used here are different from the observations used elsewhere in the model. For all the other parameters, we use weakly informative priors (Appendix S1). MCMC sampling of the joint posterior probability distribution is implemented using the Open-BUGS software (Lunn *et al.*, 2009). The OpenBUGS codes of models are available in the Dryad repository: http://dx.doi.org/doi:10.5061/dryad.f05mk. OpenBUGS uses a model statement syntax similar to the popular R software (R Development Core Team, 2012), which should make our code self-reading for familiar users. We run three parallel MCMC chains and retained 50 000 iterations after an initial burn-in of 10 000 iterations for each model. Convergence of MCMC sampling was assessed by means of the Brooks–Gelman–Rubin diagnostic (Brooks & Gelman, 1998).

Results

In the following, medians and 95% credible intervals from the posterior distribution are reported for unknowns of interest (Table 1). We also provide the posterior probability for a quantity of being positive calculated as the proportion of sampled values higher than zero. Results for the probabilities of sexual maturation and winter survival are given considering a random effect $\varepsilon_i = 0$ (eqns 5, 7 and 9).

Observation probabilities

Recapture probabilities are *a posteriori* well estimated (Fig. 5) in comparison with their weakly informative prior distributions (see Appendix S1). Using ancillary fin-clipping data (model D), uncertainty is most significantly reduced compared to PIT-tag data only (models A–C). Smolt trap efficiencies varied from 2006 (pL_1 and pP_1) to 2007 (pL_2 and pP_2) due to their sensitivity to hydrological conditions (Rivot & Prévost, 2002).

Choice between alternative life-history tactics

Whatever the model, the gradient of the probabilistic reaction norm for age at smolting is strictly positive (Table 1). The decision of smolting at 1 year of age is strongly size dependent (Fig. 6). The estimation of the reaction norm is little affected by increasing model complexity and amount of data assimilated. Using the full model D, uncertainty in the reaction norm parameters is slightly reduced compared to the simplest model A considering the 1+ smolt stage only (Table 1). At the same time, the gradient α_2 is somewhat larger and the parameter α_1 somewhat smaller.

The probability of sexual maturation for a male at the 1+ parr stage is high for both models C and D but

poorly estimated. The uncertainty is marginally reduced in model D [0.72 (0.14; 0.99) vs. 0.58 (0.19; 0.91) for models C and D, respectively].

Selective survival and cost of reproduction

The 95% posterior probability interval of β_2 includes 0 except in model D combining all life-history events and sources of data (Table 1). The gradual reduction in uncertainty in the estimation of β_2 along the suite of models finally reveals, under the most comprehensive model D, a selective survival in first winter in favour of 0+ parr that decided to smolt at 1 year of age the following spring (Fig. 7a). Note that, from model C to D, it is the incorporation of data not directly linked to the first winter survival event that allows evidencing the cost associated with freshwater residency decision.

Using PIT-tag data only (model C), parameter δ_2 was estimated negative with probability 0.72 (Table 1). Combining additional sources of data (model D), parameter δ_2 was unambiguously estimated negative with probability 0.99, thus evidencing a selective survival depending on the sexual maturation status of the 1+ parr, that is, a cost of reproduction on the second winter survival (Fig. 7b).

Discussion

There is growing interest in studying eco-evolutionary processes in the wild (Schoener, 2011). The two main challenges to address are to account for (1) the rather complex nature of these processes involving various life-history traits (and stages) and their associated inter-actions and (2) the fact that they are partially observed and/or with measurement error. It is now well recognized that the HM approach offers a generic framework for meeting these challenges in ecological studies (Clark *et al.*, 2005; Clark & Gelfand, 2006; Cressie *et al.*, 2009). By explicitly distinguishing the observation procedures from the processes of ultimate interest, HM is a powerful way to deal with uncertainties inherent in data collection and hence to focus on the underlying complex eco-evolutionary mechanisms. In the present paper, by adopting a step-by-step approach, we further demonstrated how a complex life-history model is relatively easily built from successive integration of simple components. We also demonstrated the payoffs of increasing complexity of models consisting in improving estimates precision and revealing eco-evolutionary processes of interest not identified with simpler models.

Modelling complex eco-evolutionary processes using HM

The incorporation step-by-step of successive life stages of Atlantic salmon juveniles helps combining several life-history decisions (i.e. migration and maturation decisions) resulting from evolutionary processes (i.e.

Table 1 Summary of posterior distributions (medians, 95% posterior credible intervals and probability to be positive) for demographic process parameters of interest (i.e. probabilistic reaction norm and selective survival).

| Parameter | Definition | Model | Posterior distribution | | $Pr([\theta \| Y] > 0)$ |
			Median	95% credible interval	
Demographic process					
α_1	Intercept of the size-dependent probabilistic reaction norm for smolting (eqn 4)	A	−2.23	−2.98; −1.33	0
		B	−2.10	−2.82; −1.35	0
		C	−1.83	−2.55; −1.13	0
		D	−1.91	−2.56; −1.24	0
α_2	Selection gradient of the size-dependent probabilistic reaction norm for smolting (eqn 4)	A	0.13	0.09; 0.19	0.99
		B	0.14	0.10; 0.19	0.99
		C	0.16	0.11; 0.21	0.99
		D	0.16	0.12; 0.21	0.99
β_1	First winter survival for futures 1+ parr (eqn 5)	A	0.02	−3.68; 3.78	0.50
		B	−2.04	−4.53; 0.46	0.07
		C	−2.05	−4.53; −0.63	0
		D	−1.53	−2.80; −0.73	0
β_2	Effect of the decision of smolting at 1 year of age on the first winter survival (eqn 5)	A	0.51	−3.60; 4.96	0.60
		B	2.29	−1.07; 5.63	0.90
		C	1.84	−0.20; 5.44	0.95
		D	1.64	0.43; 3.56	0.99
δ_1	Second winter survival for immature 1+ parr (males and females; eqn 9)	A	–	–	–
		B	–	–	–
		C	0.55	−2.06; 4.18	0.66
		D	−0.49	−2.29; 0.53	0.19
δ_2	Effect of the decision of maturing on the second winter survival (cost of reproduction for survival; eqn 9)	A	–	–	–
		B	–	–	–
		C	−1.28	−5.87; 3.52	0.28
		D	−1.51	−3.12; −0.28	0.01

phenotypic plasticity, selective survival, cost of reproduction). Adopting a HM framework has also the advantage of treating unobserved individual states (also called latent states) as any other unknown quantity to be estimated. This adds to the flexibility of HM because an unknown state can be used as a covariate within a model. Taking advantage of this feature, we were able to explore the effect of migration decision on winter survival, although this decision could not be observed at the onset of the winter transition. Indeed, migration is decided before the first winter by 0+ parr but it was only partially observed in the following spring through the recapture at smolts stage during their downstream migration. Thus, migration decision was an unknown to estimate for any fish not recaptured at the smolt stage (or later on). The unknown latent state indicator for migration decision before first winter was estimated thanks to the probabilistic reaction norm for migration, and was used in turn as a covariate in eqn 5 to highlight a selective winter survival. By combining two eco-evolutionary processes (life-history decision and potential evolutionary trade-off), our analysis revealed a positive relationship between the first winter survival of the 0+ parr and their decision of smolting the following spring. We used the same approach for combining the life-history decision for maturing with the second

winter survival event to highlight a survival cost of reproduction. HM allowed the estimation of the maturation state of all 1+ parr tagged, even though it was only observable for the spermating males captured at the 1+ parr stage in autumn 2006. Our results confirmed that mature male 1+ parr had a lower probability of winter survival (post-reproductive survival) than their immature counterparts (Myers, 1984; Baglinière & Maisse, 1993; Whalen & Parrish, 1999).

Dealing with various sources of available information

In line with Cressie *et al.* (2009), we showed HM is a powerful approach to deal with various sources of information within a unique framework, hence improving our ability to study complex eco-evolutionary processes (Ogle, 2009; Buoro *et al.*, 2010). In our Atlantic salmon case study, the step-by-step modelling of eco-evolutionary processes is accompanied by the sequential use of additional sources of data. By assimilating additional data sets, that is, the 1+ parr untagged at 0+ parr stage and the smolt fin-clipping experiment, into model D, we significantly improved the estimation of the observation parameters (detection probabilities; Fig. 5). More importantly, expanding the basic model
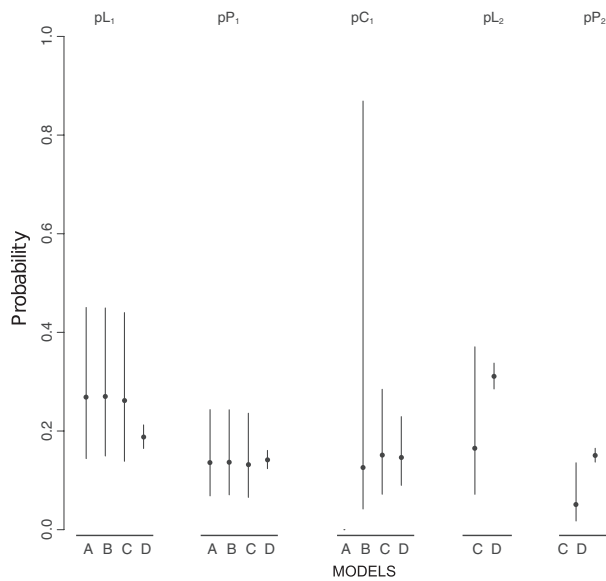
Fig. 5 Posterior distributions of the probabilities of capture for each model (also noted from 'A' to 'D'; based on 50 000 MCMC samples). Notation: *pL* and *pP* are the probabilities of capture at Leslé Mill and Princes Mill, respectively, at both occasions (1: spring 2006 and 2: spring 2007), and *pC* is the probability of capture in autumn at 1+ parr stage. The median (black point) and the 95% posterior probability interval (PPI; solid lines) are displayed. MCMC, Markov chain Monte Carlo.

and incorporating additional data sets was crucial in identifying eco-evolutionary processes of ultimate interest (Fig. 7). It was only from model C (and beyond) that the survival advantage of the 0+ parr having taken the decision to migrate the following spring, over those remaining an additional year in freshwater, could be ascertained. In the same way, it is only model D that allowed revealing the survival cost of reproduction for
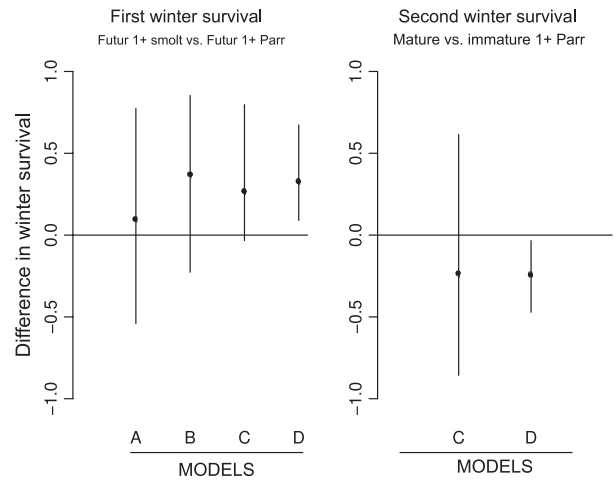

Fig. 7 Posterior distributions of the difference in survival probability (considering a zero individual random effect) for each model from 'A' to 'D'. Panel (a) Future 1+ smolts vs. future 1+ parr during the first winter and panel (b) mature 1+ parr vs. immature 1+ males during the second winter (cost of reproduction for survival). The median (black point) and the 95% posterior probability interval (PPI; solid lines) are displayed.

the 1+ parr. In the first instance, it is worth noting that the selective survival over the first winter was revealed by unfolding the model beyond the stages at stake. Indeed, the 1+ old juveniles having survived the first winter were observed in spring and autumn and were already taken into account in model B. In the second instance, the evidence of a cost of reproduction was, at least in part, due to the use of ancillary data (fin clipping) that aimed primarily at improving the estimation of the observation parameters. Our results showed the payoff of increasing model complexity, including by expanding models beyond the stages and processes of
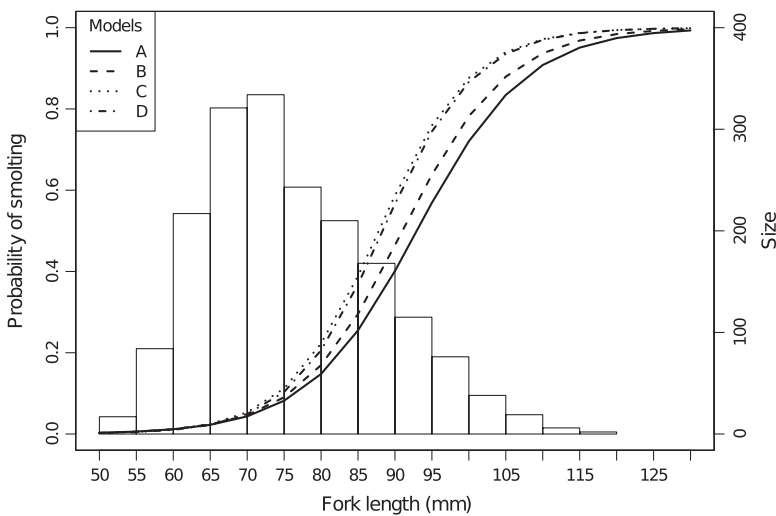

Fig. 6 Probabilistic reaction norms for the age at smolting for each model (noted from A to D). The posterior medians of the probability of smolting at 1 year of age vs. fork length are shown. A histogram of the size distribution of the 0+ parr sampled in autumn 2005 is also displayed.

focal interest. It is the inclusion of the various eco-evolutionary and observation components within a single model, which connects all unknown quantities by means of conditional relationships, that allows the information brought by the various data sets collected at different life stages to potentially contribute to the estimation of every unknown quantity in the model.

Comparing alternative models

To illustrate the relevance of HM in revealing eco-evolutionary processes using noisy data, we developed a sequence of four nested models of increasing complexity. Although we acknowledge that true processes giving rise to the field data of our case study remain elusive, we based our biological interpretation of the results on model D because (1) simpler models A–C were less integrative in terms of eco-evolutionary processes and (2) they restricted data assimilation by excluding informative data sets. However, alternative models of similar complexity could be envisaged. Of great interest for evolutionary biologists is the question of contrasting alternative eco-evolutionary hypotheses from observation data collected in the wild. In this context, model selection has become popular in ecology and evolution and allows the simultaneous confrontation of several competing hypotheses using the data at hand (Burnham & Anderson, 2002; Johnson & Omeland, 2004; Hobbs & Hilborn, 2006; Link & Barker, 2006). Many criteria have been proposed (among others, the Bayes Factor, the Deviance Information Crtierion, the Bayes Information Criterion and the Akaike Information Criterion) but from a practical point of view, model selection for complex HMs is still an open question in statistical modelling and faces several issues (Gelman & Rubin, 1996; Spiegelhalter et al., 2002 and following discussants; Celeux et al., 2006; Link & Barker, 2006; De Valpine, 2009). An alternative to model selection is to conduct careful posterior model checking within complex models. It consists in comparing replicated data obtained a posteriori from the estimated model with the observed data (Craigmile et al., 2009; Cressie et al., 2009; Brun et al., 2011; Gelman & Shalizi, 2012). Although posterior model checking techniques are still in their infancy, useful suggestions and guidelines can be found in the study by Gelman et al. (2004), Marshall & Spiegelhalter (2007) and Kerman et al. (2008).

Advances and constraints of HM

The flexibility of HM does not come without costs. With the rapid development of statistical software such as BUGS (Lunn et al., 2009) and the increase in computer power, one may be tempted to incorporate more and more details, hence increasing model complexity. This raises several issues. First, even though proceeding step-by-step as demonstrated in this paper makes model specification easier (see also Craigmile et al., 2009), it may not be straightforward for users without any programming skills. Solutions can be found by making BUGS codes publicly available, by fostering collaborations between statisticians and evolutionists and by training in HM (Gimenez, 2008; Ogle, 2009). Second, there is a risk of overparameterized models with nonestimable parameters. Rather than fitting the complete model at first, incorporating the different relevant components, piece-by-piece, starting with a simple model and then increasing complexity is here again a good practice (Craigmile et al., 2009; Brun et al., 2011). Third, to cope with complexity, one often resorts to a Bayesian approach using MCMC algorithms to fit HM to data. Despite being very flexible, these computational methods require sufficient training to be applied correctly. As an alternative to the Bayesian approach, several approaches are being developed (Lele et al., 2007; Lele & Dennis, 2009; De Valpine, 2009, 2011) that deserve further exploration.

HM as a generic framework for studying evolutionary processes in the wild?

Overall, the costs associated with HM are compensated by benefits stemming from their ability both to comprehensively integrate various life-history events and to combine multiple sources of available information. These two features are key for drawing accurate statistical inference about eco-evolutionary processes in the wild from field data. These processes operating at the level of individuals, populations and communities generate and maintain biodiversity (Pressey et al., 2007). Evolutionary processes and individual variation, raw material for natural selection, have ecological consequences and vice versa (Bolnick et al., 2011), hence a growing interest in integrating ecological and evolutionary tools and concepts. Long-term individual-based studies are becoming more frequent, providing a unique opportunity to address evolutionary questions in the wild and illustrate processes that could not previously be assessed in laboratory or with short time-scale studies (Clutton-Brock & Sheldon, 2010). HM holds great promise by offering a generic framework to combine empirical and theoretical backgrounds and explore eco-evolutionary dynamics (Ezard et al., 2009; Pelletier et al., 2009; Carlson et al., 2011). In this regard, introducing quantitative genetics into eco-evolutionary models would help understanding the evolution of traits and their ecological consequences (Ozgul et al., 2009; Coulson et al., 2010).

Acknowledgments

Pée sur Nivelle), Y. Guilloux (Federation de pêche du Morbihan, Pont-Scorff) and other technical staff members for their help in collecting field data. Mathieu Buoro and Olivier Gimenez were supported by a grant from the French Research National Agency (ANR), reference ANR-08-JCJC-0028-01.

References

Baglinière, J.L. & Maisse, G. 1993. Comparison of two methods of estimating Atlantic salmon (*Salmo salar*) wild smolt production. In: *Production of Juvenile Atlantic Salmon, Salmo salar, in Natural Waters* (R.J. Gibson & R.E. Cutting, eds). *Can. Spec. Publ. Fish. Aquat. Sci.* No. 118, pp. 189–201.

Baglinière, J.L., Prévost, E. & Maisse, G. 1994. Comparison of population dynamics of Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*) in a small tributary of the River Scorff (Brittany, France). *Ecol. Freshw. Fish* **3**: 25–34.

Bolker, B.M. 2009. Learning hierarchical models: advice for the rest of us. *Ecol. Appl.* **19**: 588–592.

Bolnick, D.I., Amarasekare, P., Levine, J.M., Novak, M., Rudolf, V.H.W., Schreiber, S.J. *et al.* 2011. Why intraspecific trait variation matters in community ecology. *Tree* **26**: 183–192.

Brooks, S.P. & Gelman, A. 1998. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* **7**: 434.

Brun, M., Abraham, C., Jarry, M., Dumas, J., Lange, F. & Prévost, E. 2011. Estimating an homogeneous series of a population abundance indicator despite changes in data collection procedure: a hierarchical Bayesian modelling approach. *Ecol. Model.* **222**: 1069–1079.

Buoro, M., Prévost, E. & Gimenez, O. 2010. Investigating evolutionary trade-offs in wild populations of Atlantic salmon (*Salmo Salar*): incorporating detection probabilities and individual heterogeneity. *Evolution* **64**: 2629–2642.

Burnham, K.P. & Anderson, D.R. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, New York, NY.

Cam, E. 2009. Contribution of capture-mark-recapture modeling to studies of evolution by natural selection. In: *Modeling Demographic Processes in Marked Populations. Environmental and Ecological Statistics*, Vol. 3 (D.L. Thomson, E.G. Cooch & M.J. Conroy, eds), pp. 83–129. Springer US, Boston, MA.

Cam, E., Link, W.A., Cooch, E.G., Monnat, J.-Y. & Danchin, E. 2002. Individual covariation in life-history traits: seeing the trees despite the forest. *Am. Nat.* **159**: 96–105.

Carlson, S.M., Quinn, T.P. & Hendry, A.P. 2011. Eco-evolutionary dynamics in Pacific salmon. *Heredity* **106**: 438–447.

Catchpole, E.A., Morgan, B.J.T. & Tavecchia, G. 2008. A new method for analysing discrete life history data with missing covariate values. *J. R. Stat. Soc. B.* **70**: 445–460.

Celeux, G., Forbes, F., Robert, C.P. & Titterington, D.M. 2006. Deviance information criteria for missing data models. *Bayesian Anal.* **1**: 651–674.

Clark, J.S. 2007. *Models for Ecological Data: An Introduction*. Princeton University Press, Princeton, NJ.

Clark, J.S. & Gelfand, A.E. 2006. A future for models and data in environmental science. *Tree* **21**: 375–380.

Clark, J.S., Ferraz, G., Oguge, N., Hays, H. & DiCostanzo, J. 2005. Hierarchical Bayes for structured, variable populations: from recapture data to life-history prediction. *Ecology* **86**: 2232–2244.

Clobert, J. 1995. Capture-recapture and evolutionary ecology: a difficult wedding? *J. Appl. Stat.* **22**: 989–1008.

Clutton-Brock, T. 1998. Introduction: studying reproductive costs. *Oikos* **83**: 421–423.

Clutton-Brock, T. & Sheldon, B.C. 2010. Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Tree* **25**: 562–573.

Conroy, M.J. 2009. Application of capture-recapture to addressing questions in evolutionary ecology. In: *Modeling Demographic Processes in Marked Populations. Series: Environmental and Ecological Statistics*, Vol. 3 (D.L. Thomson, E.G. Cooch & M.J. Conroy, eds), pp. 131–156. Springer, New York, NY.

Coulson, T.N., Tuljapurkar, S. & Childs, D.Z. 2010. Using evolutionary demography to link life history theory, quantitative genetics and population ecology. *J. Anim. Ecol.* **79**: 1226–1240.

Craigmile, P.F., Calder, C.A., Li, H., Paul, R. & Cressie, N. 2009. Hierarchical model building, fitting, and checking: a behind-the-scenes look at a Bayesian analysis of arsenic exposure pathways. *Bayesian Anal.* **4**: 1–36.

Cressie, N., Calder, C.a., Clark, J.S., Ver Hoef, J.M. & Wikle, C.K. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecol. Appl.* **19**: 553–570.

De Valpine, P. 2009. Shared challenges and common ground for Bayesian and classical analysis of hierarchical statistical models. *Ecol. Appl.* **19**: 584–588.

De Valpine, P. 2011. Frequentist analysis of hierarchical models for population dynamics and demographic data. *J. Ornithol.* **152**: 393–408.

Duston, J. & Saunders, R. 1997. Increases in Atlantic salmon smolt production by elevated winter temperature and size-grading are limited by sexual maturation. *Environ. Biol. Fishes* **50**: 149–166.

Ellison, A.M. 2004. Bayesian inference in ecology. *Ecol. Lett.* **7**: 509–520.

Ezard, T.H.G., Côté, S.D. & Pelletier, F. 2009. Eco-evolutionary dynamics: disentangling phenotypic, environmental and population fluctuations. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**: 1491–1498.

Fleming, I.A. & Reynolds, J.D. 2004. Salmonid breeding systems. In: *Evolution Illuminated: Salmon and Their Relatives* (A.P. Hendry & S.C. Stearns, eds), pp. 264–294. Oxford University Press, Oxford.

Gelman, A. & Rubin, D.B. 1996. Markov chain Monte Carlo methods in biostatistics. *Stat. Methods Med. Res.* **5**: 339–355.

Gelman, A. & Shalizi, C.R. 2012. Philosophy and the practice of Bayesian statistics. *Br. J. Math. Stat. Psychol.* doi: 10.1111/j.2044-8317.2011.02037.x.

Gelman, A., Carlin, J.B., Stern, H. & Rubin, D.B. 2004.*Bayesian Data Analysis*, 2nd edn, pp. 696. Chapman & Hall/CRC, Boca Raton, FL.

Gimenez, O. 2008. Discussion: towards a Bayesian analysis template? *Can. J. Stat.* **36**: 21–24.

Gimenez, O., Rossi, V., Choquet, R., Dehais, C., Doris, B., Varella, H. *et al.* 2007. State-space modelling of data on marked individuals. *Ecol. Model.* **206**: 431–438.

Gimenez, O., Viallefont, A., Charmantier, A., Pradel, R., Cam, E., Brown, C.R. *et al.* 2008. The risk of flawed inference in

evolutionary studies when detectability is less than one. *Am. Nat.* **172**: 441–448.

Gimenez, O., Lebreton, J.-D., Gaillard, J.-M., Choquet, R. & Pradel, R. 2012. Estimating demographic parameters using hidden process dynamic models. *Theor. Popul. Biol.* (in press).

Gimenez, O., Rossi, V., Choquet, R., Dehais, C., Doris, B., Varella, H. *et al.* 2007. State-space modelling of data on marked individuals. *Ecol. Modell.* **206**: 431–438.

Hadfield, J.D. 2008. Estimating evolutionary parameters when viability selection is operating. *Proc. R. Soc. Lond. B. Biol. Sci.* **275**: 723–734.

Heino, M. 2002. Estimating reaction norms for age and size at maturation with reconstructed immature size distributions: a new technique illustrated by application to Northeast Arctic cod. *ICES J. Mar. Sci.* **59**: 562–575.

Hobbs, N.T. & Hilborn, R. 2006. Alternatives to statistical hypothesis testing in ecology: a guide to self teaching. *Ecol. Appl.* **16**: 5–19.

Johnson, J.B. & Omland, K.S. 2004. Model selection in ecology and evolution. *Trends Ecol. Evol.* **19**: 101–108.

Kerman, J., Gelman, A., Zheng, T. & Ding, Y. 2008. Visualization in Bayesian data analysis. In: *Handbook of Data Visualization* (C.-H. Chen, W.K. Härdle & A. Unwin, eds), pp. 709–724. Springer US, Boston, MA.

Kéry, M. & Schaub, M. 2011. *Bayesian Population Analysis Using WinBUGS – A Hierarchical Perspective*. Academic P., Burlington, VT.

King, R., Brooks, S.P. & Coulson, T. 2008. Analyzing complex capture-recapture data in the presence of individual and temporal covariates and model uncertainty. *Biometrics* **64**: 1187–1195.

King, R., Morgan, B.J.T., Gimenez, O. & Brooks, S.P. 2009. *Bayesian Analysis of Population Ecology*. Chapman and Hall/CRC, Boca Raton, FL.

Kuhnert, P.M. 2011. Four case studies in using expert opinion to inform priors. *Environmetrics* **22**: 662–674.

Lebreton, J.-D., Burnham, K.P., Clobert, J. & Anderson, D.R. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol. Monogr.* **62**: 67–118.

Lele, S.R. & Dennis, B. 2009. Bayesian methods for hierarchical models: are ecologists making a Faustian bargain? *Ecol. Appl.* **19**: 581–584.

Lele, S.R., Dennis, B. & Lutscher, F. 2007. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecol. Lett.* **10**: 551–563.

Link, W. & Barker, R.J. 2006. Model weights and the foundations of multimodel inference. *Ecology* **87**: 2626–2635.

Link, W.A. & Barker, R.J. 2010. *Bayesian Inference: With Ecological Applications*. Academic P., Burlington, VT.

Lunn, D.J., Thomas, A., Best, N. & Spiegelhalter, D.J. 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* **10**: 325–337.

Lunn, D.J., Spiegelhalter, D.J., Thomas, A. & Best, N. 2009. The BUGS project: evolution, critique and future directions. *Stat. Med.* **28**: 3049–3067.

Marshall, E.C. & Spiegelhalter, D.J. 2007. Simulation-based tests for divergent behaviour in hierarchical models. *Bayesian Anal.* **2**: 409–444.

McCarthy, M.A. & Masters, P. 2005. Profiting from prior information in Bayesian analyses of ecological data. *J. Appl. Ecol.* **42**: 1012–1019.

McCormick, S.D., Hansen, L.P., Quinn, T.P. & Saunders, R.L. 1998. Movement, migration, and smolting of Atlantic salmon (*Salmo salar*). *Can. J. Fish. Aquat. Sci.* **55**: 77–92.

Metcalfe, N.B. & Thorpe, J.E. 1992. Anorexia and defended energy levels in over-wintering juvenile salmon. *J. Anim. Ecol.* **61**: 175–181.

Myers, R.A. 1984. Demographic consequences of precocious maturation of Atlantic salmon (*Salmo salar*). *Can. J. Fish. Aquat. Sci.* **41**: 1349–1353.

Myers, R.A. & Hutchings, J.A. 1987. Mating of anadromous Atlantic salmon, *Salmo salar* L., with mature male parr. *J. Fish Biol.* **31**: 143–146.

Nakagawa, S. & Freckleton, R.P. 2008. Missing inaction: the dangers of ignoring missing data. *Tree* **23**: 592–596.

Ogle, K. 2009. Hierarchical Bayesian statistics: merging experimental and modeling approaches in ecology. *Ecol. Appl.* **19**: 577–581.

Ozgul, A., Tuljapurkar, S., Benton, T.G., Pemberton, J.M., Clutton-Brock, T.H. & Coulson, T. 2009. The dynamics of phenotypic change and the shrinking sheep of St. Kilda. *Science* **325**: 464–467.

Pelletier, F., Garant, D. & Hendry, A.P. 2009. Eco-evolutionary dynamics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**: 1483–1489.

Pressey, R.L., Cabeza, M., Watts, M.E., Cowling, R.M. & Wilson, K.A. 2007. Conservation planning in a changing world. *Tree* **22**: 583–592.

Prévost, E., Chadwick, E.M.P. & Claytor, R.R. 1992. Influence of size, winter duration and density on sexual-maturation of Atlantic salmon (*Salmo salar*) juveniles in little Codroy river (southwest Newfoundland). *J. Fish Biol.* **41**: 1013–1019.

Proaktor, G., Coulson, T. & Milner-Gulland, E.J. 2008. The demographic consequences of the cost of reproduction in ungulates. *Ecology* **89**: 2604–2611.

R Development Core Team, 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

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

Roff, D.A. 1992. *The Evolution of Life Histories: Theory and Analysis*. Chapman an., New York.

Roff, D.A. 2002. *Life History Evolution*. Sinauer Associates, Sunderland, MA.

Roff, D.A., Heibo, E. & Vøllestad, L.A. 2006. The importance of growth and mortality costs in the evolution of the optimal life history. *J. Evol. Biol.* **19**: 1920–1930.

Royle, J.A. 2008. Modeling individual effects in the Cormack-Jolly-Seber Model: a state-space formulation. *Biometrics* **64**: 364–370.

Royle, J.A. & Dorazio, R.M. 2008. *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations, and Communities*, pp. 444. Academic P., San Diego, CA.

Schoener, T.W. 2011. The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *Science* **331**: 426–429.

Schofield, M.R., Barker, R.J. & MacKenzie, D.I. 2007. Flexible hierarchical mark-recapture modeling for open populations using WinBUGS. *Environ. Ecol. Stat.* **16**: 369–387.

Spiegelhalter, D., Best, N., Carlin, B. & Van der Linde, A. 2002. Bayesian measures of model complexity and fit. *J. Roy. Stat. Soc. B* **64**: 583–639.

Stearns, S.C. 1992. *The Evolution of Life Histories*. Oxford Uniersity, New York.

Stickler, M., Enders, E.C., Pennell, C.J., Cote, D., Alfredsen, K. T. & Scruton, D.A. 2008. Habitat use of Atlantic salmon *Salmo salar* parr in a dynamic winter environment: the influence of anchor-ice dams. *J. Fish Biol.* **73**: 926–944.

Thorpe, J.E. & Metcalfe, N.B. 1998. Is smolting a positive or a negative developmental decision? *Aquaculture* **168**: 95–103.

Thorpe, J.E., Mangel, M., Metcalfe, N.B. & Huntingford, F.A. 1998. Modelling the proximate basis of salmonid life-history variation, with application to Atlantic salmon, *Salmo salar* L. *Evol. Ecol.* **12**: 581–599.

Van Dongen, S. 2006. Prior specification in Bayesian statistics: three cautionary tales. *J. Theor. Biol.* **242**: 90–100.

Webb, J., Verspoor, E., Aubin-Horth, N., Romakkaniemi, A. & Amiro, P. 2007. The Atlantic salmon. In: *The Atlantic Salmon: Genetics, Conservation and Management* (E. Verspoor, L. Stradmeyer & J.L. Nielsen, eds), pp. 17–56. Blackwell, Oxford.

Whalen, K.G. & Parrish, D.L. 1999. Effect of maturation on parr growth and smolt recruitment of Atlantic salmon. *Can. J. Fish. Aquat. Sci.* **56**: 79–86.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1 Summary of tag/recapture data for each observation events and each method of tagging (PIT and fin-clipping) for one cohort of Atlantic salmon.

Table S2 States (indicators), transition probabilities (life-history traits of ultimate interest) and equations of the demographic process accounting for the sequence of events corresponding to the life history of Atlantic salmon from the 0+ parr stage in autumn to the migration to the ocean in the second year of life.

Appendix S1 Choice of prior distribution.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.