This preprint is a PDF of a manuscript that has been accepted for publication in an ESA journal. It is the final version that was uploaded and approved by the author(s). While the paper has been through the usual rigorous peer review process of ESA journals, it has not been copy-edited, nor have the graphics and tables been modified for final publication. Also note that the paper may refer to online Appendices and/or Supplements that are not yet available. We have posted this preliminary version of the manuscript online in the interest of making the scientific findings available for distribution and citation as quickly as possible following acceptance. However, readers should be aware that the final, published version will look different from this version and may also have some differences in content.

The doi for this manuscript and the correct format for citing the paper are given at the top of the online (html) abstract.

Once the final published version of this paper is posted online, it will replace the preliminary version at the specified doi.
TESTING HYPOTHESES IN EVOLUTIONARY ECOLOGY WITH IMPERFECT DETECTION: STRUCTURAL EQUATION MODELING OF CAPTURE-RECAPTURE DATA

Sarah Cubaynes\textsuperscript{1,2}, Claire Doutrelant\textsuperscript{1,3}, Arnaud Grégoire\textsuperscript{1,4}, Philippe Perret\textsuperscript{1,5}, Bruno Faivre\textsuperscript{6,7} and Olivier Gimenez\textsuperscript{1,8}.

\textsuperscript{1}Centre d’Ecologie Evolutive et Fonctionnelle
UMR 5175, 1919 Route de Mende
34293 Montpellier, Cedex 5 - France

\textsuperscript{2}sarah.cubaynes@cefe.cnrs.fr, \textsuperscript{3}claire.doutrelant@cefe.cnrs.fr, \textsuperscript{4}arnaud.gregoire@cefe.cnrs.fr, \textsuperscript{5}philippe.perret@cefe.cnrs.fr, \textsuperscript{6}bruno.faivre@u-bourgogne.fr, \textsuperscript{8}olivier.gimenez@cefe.cnrs.fr

\textbf{Abstract:} Studying evolutionary mechanisms in natural populations often requires testing multifactorial scenarios of causality involving direct and indirect relationships among individual and environmental variables. It is also essential to account for imperfect detection of individual to provide unbiased demographic parameter estimates. To cope with these issues, we develop a new approach combining structural equation models with capture–recapture models (CR–SEM) that allows the investigation of competing hypotheses about individual and environmental variability observed in demographic parameters. We employ Markov chain Monte Carlo sampling in a Bayesian framework to i) estimate model parameters, ii) implement a model selection procedure to evaluate competing hypotheses about causal mechanisms and iii) assess the fit of models to data using posterior predictive checks. We illustrate the value of our approach using two case studies on wild bird populations. We first show that CR–SEM can be useful to quantify the action of selection on a set of phenotypic traits with an analysis of selection gradients on morphological traits in Common blackbirds (\textit{Turdus merula}). In a second case study on Blue tits
(Cyanistes caeruleus), we illustrate the use of CR–SEM to study evolutionary trade-offs in the wild, while accounting for varying environmental conditions.

**Keywords:** Capture–recapture models; Evolutionary ecology; Individual heterogeneity; Life-history trade-offs; Selection gradient analyses; State–space models; Structural equation models.

**INTRODUCTION**

A key issue in ecology is to disentangle the multiple interacting factors driving animal demography, e.g. environmental forcing, individual variability or evolutionary trade-offs. Because biological phenomena are often the result of complex interactions between living organisms and their environment, the underlying mechanisms are multifactorial and usually involve numerous variables interacting directly or indirectly through other variables. Understanding such mechanisms and factors thus requires identifying the relevant variables, describing the structural links connecting them and quantifying the shape and the strength of the relationships among them. To do so, one needs to carry out a series of experimental manipulations, in which some variables are fixed to constant values, thus minimizing uncontrolled variation. Because manipulative experiments are often difficult or impossible to conduct in natural populations, studies of mechanisms driving natural populations dynamics rely on observational data.

An alternative and complementary approach for evaluating causal assumptions based on observational data is to formalize and confront different scenarios of causality using structural equation modeling (SEM; Shipley, 2002; Grace, 2006). SEM is a multivariate regression framework that allows the evaluation of direct and indirect relationships among a set of variables, including variables that cannot be directly observed and measured, so-called latent variables. Typically, a SEM is built by specifying a set of pathways describing how variables may affect each other. If the model is not consistent with the observations, the corresponding scenario is rejected and an alternative hypothesis about the underlying mechanism has to be considered. Shipley (2002) defines SEM as 'models representing translations of a series of hypothesized
cause-effect relationships between variables into a composite hypothesis concerning patterns of statistical dependencies’. Rather than being strictly causal, SEM allows testing competing causal assumptions and refuting an unlikely hypothesis, based on cor relational evidence. Hence, the term ‘causal model’ should be understood as ‘a model that conveys causal assumptions’ and not necessarily as a model that produces validated causal conclusions (see Shipley (2009) and Pearl (2000) for further discussion about causal inference).

The flexibility of SEM to represent complex scenarios involving several observed and latent variables has lead to an increasing number of applications in ecology and evolution (Shipley, 2002; Pugesek et al., 2003; Grace, 2006). In particular, several examples include studies about natural selection and life–history strategies in natural and experimental populations (Mitchell, 1992; Pugesek and Tomer, 1996; Kirk et al., 2001). To address important questions linked to natural selection, one needs to estimate fitness components like demographic parameters, e.g. survival. However, these traits are difficult to obtain because of the imperfect detection of individuals inherent to monitoring in the wild. Besides, if ignored, the issue of detectability less than one can lead to flawed inference about the relationship between survival and explanatory variables (Gimenez et al., 2008). To deal with this issue, capture–recapture models (Lebreton et al., 1992) provide unbiased estimators of demographic parameters while explicitly coping with imperfect detection inherent to wild populations. CR models allow estimating the effect of temporal and individual covariates on demographic parameters and testing their significance in a regression-like framework (Pollock, 2002). However, these models can only detect correlations and do not inform on causal pathways in the (possibly indirect) relationships between demographic parameters and the factors possibly explaining their variability, neither do they allow the incorporation of latent variables.

Here, we develop a method to test competing hypotheses about individual and environmental variability observed in demographic parameters using CR data in a SEM framework. We show how to combine structural equation and capture–recapture models (CR–SEM hereafter) using a hierarchical approach. To illustrate our method, we consider two case studies in evolutionary ecology on wild bird populations with an analysis of selection gradients on morphological traits
in Common blackbirds (*Turdus merula*) and a study of trade-offs between survival and reproduction in Blue tits (*Cyanistes caeruleus*). Adopting a Bayesian approach, we estimate parameters using Markov chain Monte Carlo methods, implement a model selection procedure to evaluate competing hypotheses about causal mechanisms and assess the fit of the CR–SEM.

**Structural Equation Modeling of CR Data**

In this section, we introduce the general procedure for processing CR–SEM. We used a hierarchical approach in which we first modeled CR data to obtain survival (level 1), then we plugged a SEM involving survival and other variables, measured or unmeasured, linked with direct or indirect relationships to explain variations in survival (level 2). We estimated model parameters using a Bayesian approach implemented with a MCMC sampling procedure, which provides powerful computer-intensive methods for handling complex models and are becoming increasingly popular in SEM (Lee, 2007; Congdon, 2006; Palomo et al., 2007). We compared competing models using Bayesian model selection. We calculated posterior model probabilities using the method developed by Kuo and Mallick (1998) (see Royle (2008) for an example of implementation in the CR framework). For each parameter for which we wanted to test the relevance, we introduced an indicator variable \( w \) having a Bernoulli(0.5) prior distribution, and premultiplied the parameter by \( w \). We computed the posterior model probability for a particular model from the MCMC histories, using the ratio between the number of iterations giving this model over the total number of iterations. We also reported the relative importance of a particular factor by calculating the number of iterations giving a model containing the corresponding parameter over the total number of iterations (Appendix 1). In addition, we calculated posterior predictive checks to evaluate the fit of the model to the data (Appendix 2).

The simulations were performed using JAGS (http://calvin.iarc.fr/martyn/software/jags/) called from program R (Ihaka and Gentleman, 1996) using package rjags (Plummer, 2003).
Level 1: modeling survival using CR data

To account for the issue of detectability less than one, we used CR data collected under the form of 1’s and 0’s corresponding to a detection or not of I individuals over T sampling occasions. CR models can be formulated as state-space models (SSM: or hierarchical models) (Gimenez et al., 2007; Royle, 2008; Schofield and Barker, 2008) to distinguish the underlying demographic process from the observation process, which provides much flexibility in the modeling of demographic parameters.

Here, we focused on survival and the reasoning was conditional on first capture of individuals. The model had two main components. The state model specified the dynamic process. Let $X_{i,t}$ be a binary random variable representing the demographic process, with $X_{i,t} = 1$ if individual $i$ was alive and available for detection at time $t$ and 0 if it was dead. The state process in the SSM formulation stipulates that if individual $i$ was alive at time $t - 1$, it survived until time $t$ with survival probability $\phi_{i,t}$ or died with a probability $1 - \phi_{i,t}$; in other words, $X_{i,t}$ is distributed as a Bernoulli random variable with parameter $\phi_{i,t}$ given $X_{i,t-1} = 1$. The observation model connects the demographic process to its observation through the detection of individuals. Let $Y_{i,t}$ be a binary random variable taking values 1 if individual $i$ was encountered at time $t$ and 0 otherwise. If individual $i$ was alive at time $t$, then it had a probability $p_{i,t}$ of being encountered and a probability $1 - p_{i,t}$ otherwise; in other words, the link between survival and the detection of individuals is made through the observation equation, which states that $Y_{i,t}$ is distributed as a Bernoulli random variable with parameter $p_{i,t}$ given $X_{i,t} = 1$. In the following, we assumed homogeneous detection among individuals, so that the index $i$ was dropped in $p_{i,t}$. Temporal variation (index $t$) was assessed specifically for each case study.

Level 2: structural equation model (SEM) for survival

SEM is a multivariate regression framework that allows the evaluation of direct and indirect relationships among a set of correlated variables, including variables that cannot be directly observed and measured, so-called latent variables, while taking measurement errors into account (more details can be found in Pugesek et al. (2003); Grace (2006); Lee (2007)).
In the SEM process, causal assumptions are first translated into models for evaluation. Model specification consists of formulating latent variables, hypothesizing their interdependencies, and choosing their indicators by specifying a set of pathways describing how variables may affect each other, based on theory or prior knowledge of the process. Observed variables (e.g. morphological measures or index of food abundance in examples below) are referred to as indicators of the latent variables (e.g. individual fitness or environmental quality in examples below).

We made the distinction between independent latent variables – exogenous variables – vs. factors that might appear as dependent variables in the model – endogenous variables.

The SEM is typically composed of a measurement model which specifies the relationships among the latent and the observed variables, and a structural model which specifies the relationships among the latent variables. For a discussion of the concept of latent variables, see Grace et al. (2010). We used the LISREL formalism (e.g. Bollen, 1989; Grace et al., 2010) to write down the SEM. The measurement model was split into two parts, depending on whether the observed variables were connected to exogenous or endogenous latent variables:

\[ x = \Theta \xi + \delta \]  

where Eq. (1) related the observed variables \( x \) to a vector of latent exogenous variables \( \xi \) through the matrix \( \Theta \) of regression parameters; we also had

\[ y = \Lambda \eta + \varepsilon \]  

where Eq. (2) related the observed variables \( y \) to a vector of latent endogenous variables \( \eta \) through the matrix \( \Lambda \) of regression parameters. The measurement error \( \delta \) and \( \varepsilon \) was assumed to be normally distributed with mean 0 and variance parameters to be estimated.

In the specific case of SEM-CR models, survival could appear in the SEM via the first or the second measurement equation, as it could be connected to exogenous or endogenous variables. In the first example on black birds (see Applications), survival was used as a proxy for the latent
variable representing individual fitness. Hence, survival was introduced in Eq. (2) as an indicator of an endogenous latent variable $y$. In the second example on blue tits (see Section 5), survival was related to an exogenous latent variable representing parental investment in reproduction. Hence, survival was introduced in Eq. (1), as a part of vector $x$.

The structural model allows the study of relationships among latent variables, here $\eta$ (endogenous) and $\xi$ (exogenous), via the relationship:

$$\eta = B \eta + \Gamma \xi + \zeta$$

(3)

where the matrix $B$ captures the relationships among the latent endogenous variables, while the matrix $\Gamma$ captures the effect of $\xi$ on $\eta$. The $\zeta$ stands for the unexplained variability in $\eta$, and was assumed to be normally distributed with mean 0 and variance parameters to be estimated. We assumed that errors terms were all independent of each other, and that, conditional on the latent variables being known, the observed variables were independent. The later means that we considered that there was no unmodeled common cause influencing simultaneously several variables in the model, i.e. no unobserved ‘confounder’ effect.

APPLICATIONS

Selection gradient analysis in common blackbirds

Assessing the relationship linking (components of) fitness (e.g., survival) to a suite of phenotypic traits (e.g., various morphological traits) is an important step in describing selection pressure on phenotypic traits. To do so, the standard second-order polynomial regression method is usually used (Lande and Arnold, 1983), and has recently been integrated in CR models (Gimenez et al., 2009a). However, this approach was not designed to understand how selection acts on combinations of traits through direct or indirect relationships, nor was it developed to deal with latent variables.

Here, the aim was to investigate the action of selection on a set of morphologic traits in a natural
population of blackbirds. We used a dataset on the relationship between morphological traits (i.e. tarsus, phalanx, wing and tail length) and survival of adult blackbirds in an urban park in Dijon, France. From 1998 to 2002, 84 female birds were banded, measured for morphological traits, and released. These morphological traits have been demonstrated to have a significant heritable component in birds and are therefore potentially subject to natural selection.

We considered a model in which the morphological variables were connected to ‘fitness’ via a latent variable which we referred to ‘overall size’ (Fig. 1A). Morphological traits were different ways of measuring body size which served as indicators for this latent variable. Fitness in turn was assumed to be perfectly related to ‘survival’ which was estimated from CR data.

Although survival is known to be highly related to fitness, we acknowledge that assuming survival is perfectly related to fitness was a strong assumption. If survival is not a good proxy for fitness, e.g. individuals surviving better have low reproductive values, this could lead to misleading conclusions about the relationship between size and fitness. Using data on individual reproductive output as a second indicator for latent variable fitness would allow to relax this assumption and to specify a more biologically plausible model (Shipley, 2002).

[Figure 1 about here]

Using the formulation provided above, we then wrote down the measurement and structural equations of the model.

Concerning the first measurement equation relating the observed variables to the exogenous latent variable, Eq. (1), the SEM involved a unique exogenous latent variable $\xi$ representing ‘overall size’, with four indicators $\mathbf{x} = (x_1, x_2, x_3, x_4)'$, namely ‘tarsus length’, ‘phalanx length’, ‘wing length’ and ‘rectrice length’. Hence, we had $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4)'$ the regression parameters relating the morphological variables to ‘overall size’ and $\delta = (\delta_1, \delta_2, \delta_3, \delta_4)'$ the associated measurement errors. We assumed $\delta_1, \delta_2 \sim N(0, \sigma_1^2)$ as $x_1$ and $x_2$ were measured with the same tool, and $\delta_3, \delta_4 \sim N(0, \sigma_2^2)$ for the same reason. Then, the SEM involved a unique endogenous latent variable $\eta = \eta$ representing individual ‘fitness’, assumed to be perfectly related to its unique indicator $y = \phi$, namely ‘survival’. Consequently, the measurement equation relating the
observed variable to the endogenous latent variable, Eq. (2), involved only scalars with the regression coefficient $A = 1$ and the error associated with survival estimated from the CR model $\varepsilon = \varepsilon$. In the structural model, we specified a linear relationship between the latent variables size, $\xi$, influencing fitness, $\eta$: $\eta = \gamma \xi + \zeta$ with $\xi \sim N(0, \sigma^2_\xi)$. For identifiability issues, we specified $\zeta \sim N(0, 1)$ and $\theta_1 = 1$.

These constraints were arbitrary values used to achieve identifiability, but were not intended to have any impact on model interpretation. By setting $\theta_1$ to 1, we stipulated that $x_1$ was our reference for interpreting $\zeta$; by doing so, it defined the scale of $\zeta$ meaning that $\zeta$ had to be interpreted in the same unity as $x_1$ (tarsus length measured in centimeter). In other words, this value was used to identify the scale of the latent variables. Scale identification was needed because latent variables were unmeasured and could have arbitrary scales of measurement.

Besides directional selection, we were also interested in stabilizing selection that could be an alternative. We therefore considered nonlinear effects by writing $\eta = \gamma_1 \xi + \gamma_2 \xi^2 + \zeta$ in the structural model. Model selection was performed on the $\theta$'s and the $\gamma$'s. A similar model was considered by Pugesek and Tomer (1996, see also Shipley (2002)) in a situation where detectability was assumed $= 1$. Here, a preliminary analysis using program E-SURGE (Choquet et al., 2009) suggested that detectability $< 1$, and that the recapture probability could be considered constant over time.

In order to completely specify the Bayesian model, we provided prior distributions for all parameters. Specifically, we chose uniform distributions for the detection probabilities ($U(0, 1)$) and normal distributions with large variance for the regression parameters ($N(0, 1000)$). We assigned uniform distributions to the standard deviation of the random effects ($U(0, 100)$) (Gelman, 2006). Convergence was assessed using the Gelman and Rubin statistic which compares the within to the between variability of chains started at different and dispersed initial values (Gelman, 1996). The quality of mixing was assessed by visually inspecting the chains.

We used two MCMC–chains of 15,000 iterations, a burn–in of 5,000 iterations which resulted in acceptable mixing and convergence. Model code is provided in Appendix 3.

All the morphological variables considered were positively related to the latent fitness variable.
(Table 1) and were selected very often in the model (relative importance for \( \theta_2 \) is 1, for \( \theta_3 \) is 0.73 and for \( \theta_4 \) is 0.99), showing a clear size effect. The correlation of tarsus and phalanx length with fitness was stronger, suggesting a higher selection pressure on these traits. One explanation for this difference could originate in the period when both traits are grown. Contrary to feather structures (i.e. wing and tail) which are renewed each year during moult, tarsus and phalanx end their growth before chicks leave their nest. These traits are thus strongly associated with early life stages and condition. This period has long term consequences (Lindström, 1999), in agreement with the stronger link observed between survival and tarsus and phalanx length.

The posterior model probability of the quadratic model was 0.32 for both coefficients non-zero vs. 0.54 for quadratic term only non-zero. These results revealed a quadratic effect of the factor size on survival, which suggested a disruptive selection, favoring smallest and biggest individuals rather than medium-sized individuals (the posterior distribution of the quadratic coefficient is centered on positive values, so that the curve is concave). A previous study on this population showed that female investment in reproduction in terms of breeding attempts per season was positively associated with female size (Faivre et al., 2001). Bigger females might be considered better females, investing in both reproduction and survival. In contrast, smaller females with high survival would invest less in reproduction, which might reflect a trade-off between survival and reproduction. Detection probability was relatively high. Posterior distributions are displayed in Appendix 4.

The posterior distributions of the replicated survival looked similar to the posterior distributions of the estimated survival, and the replicated morphological measures simulated from the posterior distributions were centered on the observed values, which suggested an adequate fit of the model (Appendix 2).
Evolutionary trade-offs in Blue tits

Life history theory predicts the existence of evolutionary trade-offs among traits closely related to fitness (Stearns, 1992). As these traits are all dependent on the same limited resources (e.g. time, energy), investment in one trait should have consequences on investment on another traits within a same season or over life. Although trade-offs between e.g. reproduction and survival have been demonstrated in experimental studies, evidence of such costs in natural conditions remains scarce possibly due to unaccounted environmental conditions affecting individual strategies of resources acquisition and allocation (van Noordwijk and de Jong, 1986), and the imperfect detection of individuals which makes the detection of trade-offs difficult (e.g. Buoro et al., 2010).

Here, we used CR–SEM to explore a cost of reproduction on survival in blue tits while accounting for complex interactions among environmental conditions and individual life histories and detectability less than one.

We used data on 755 breeding individuals that were banded, released and recaptured in spring during the breeding seasons between 1984 and 2000, in Pirio, on the island of Corsica (France).

We examined the influence of parental investment in reproduction (‘investment’ and ‘breeding success’) on adult annual survival (‘survival’), while accounting for an effect of environmental conditions (‘environment’). The SEM was constructed based on the assumption that 1) environmental conditions affect parental investment in reproduction which in turn influences breeding success, and 2) adult investment in reproduction may influence their survival (Fig. 1B).

The SEM involved the three following latent variables. Investment in reproduction was assessed via ‘adult weight’ that was recorded 15 days after hatching for all individuals that were captured and averaged over the year. Second, breeding success was assessed via ‘clutch size’, ‘chick weight’ at 15 days and ‘chick survival’ at the nest until day 15 which were recorded in all nest boxes (Blondel et al., 2006) and averaged over the year. Third, as blue tits reproduction is mainly dependent on caterpillar abundance, environment was assessed via annual intensity (‘peak mode’) and length (‘peak width’) of the peak of caterpillar abundance (see Zandt, 1994, for details about the method). Years of strong environmental constraints, with lower food
availability, corresponded to low values of this latent variable.

Using the formulation above, we wrote down the measurement and structural equations of the model. The first measurement equation, Eq. (1), related the observed indicators of food abundance \( \mathbf{x} = (x_1, x_2)' \), namely peak width and peak mode, to the exogenous latent variable \( \xi \) representing environmental breeding conditions. Hence, we had \( \Theta = (\theta_1, \theta_2)' \) the regression parameters relating the two indicators to the latent variable environment and \( \delta = (\delta_1, \delta_2)' \) the associated measurement errors. We assumed \( \delta_j \sim N(0, \sigma_{\delta_j}^2) \), \( j = 1, 2 \).

Then, the SEM involved two endogenous latent variables. Breeding success, \( \eta_1 \), had three previously listed indicators \( y_1, y_2 \) and \( y_3 \), while parental investment, \( \eta_2 \), had one indicator, \( y_4 \), namely adult weight, and was also related to survival \( \phi \). This way, survival was introduced as an indicator of an endogenous latent variable and parameter \( \lambda_5 \), connecting parental investment and survival, captured the cost of reproduction on survival. Hence, the measurement equation relating the observed variable to the endogenous latent variable, Eq. (2), involved

\[
\mathbf{y} = (y_1, y_2, y_3, y_4, \phi)' , \ \eta = (\eta_1, \eta_2)',
\]

with regression coefficients

\[
\mathbf{A} = \begin{bmatrix}
\lambda_1 & \lambda_2 & \lambda_3 & 0 & 0 \\
0 & 0 & 0 & \lambda_4 & \lambda_5
\end{bmatrix}
\]

and associated measurement errors \( \mathbf{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5)' \). We assumed \( \delta_j \sim N(0, \sigma_{\delta_j}^2) \), \( j = 1, 2 \) and \( \varepsilon_k \sim N(0, \sigma_{\varepsilon_k}^2) \), \( k = 1, \ldots, 5 \). In addition, for identifiability issues we specified \( \xi \sim N(0, 1) \).

Third, the structural equation specified the relationship among the three latent variables, i.e. the hypothesized assumption that environmental conditions influenced parental investment which in turn influenced breeding success. We specified the matrix of coefficients of the exogenous latent variables in the structural relationships

\[
\mathbf{B} = \begin{bmatrix}
0 & b \\
0 & 0
\end{bmatrix}
\]

with parameter \( b \) representing the influence of parental investment on breeding success. Then, we
specified the matrix of coefficients of the endogenous latent variables in the structural
relationship $\Gamma = (0, \gamma)'$, with parameter $\Gamma$ linking the latent variables environment and
investment representing the impact of environmental breeding conditions on parental investment
in reproduction. Finally, the two components of the equation errors (random disturbances) in the
structural relationship between the latent variables $\zeta$ were assumed $N(0,1)$.
Model selection was performed on the $\theta$'s, the $\lambda$'s, $\gamma$ and $b$. Regarding the detection process, a
preliminary analysis using program E–SURGE (Choquet et al., 2009) suggested that the
recapture probability was time-varying.
Same priors were used as in the previous example. We used two MCMC–chains of 300,000
iterations, a burn–in of 50,000 iterations which resulted in acceptable mixing and convergence.
Model code is provided in Appendix 3. The three variables clutch size, chick weight and chick
survival were all positively related to the latent variable breeding success (Table 1; relative
importance of $\lambda_1$ is 1, 0.68 for $\lambda_2$ and 1 for $\lambda_3$). The two observed variables width and intensity
of the peak of caterpillar were also very often selected in the model (relative importance 1 for
both $\theta_1$ and $\theta_2$) and both positively related to the latent variable environment. The posterior
distribution of $\lambda_5$ which captures the relationship between survival and investment was centered
on positive values, but not selected by the model selection procedure (relative importance 0.08),
indicating no cost of reproduction on survival. The coefficient $\gamma$ of the regression between
environment and investment was also positive (relative importance 1), showing a positive effect
of food abundance during the rearing period on parental investment in reproduction, and thus
indirectly on the three indicators of breeding success as this latent variable was positively related
via $b$ to investment (relative importance 1). The effect was higher for clutch size than chick
weight and nest survival, suggesting that variation of parental investment in reproduction
affected mainly clutch size rather than chick weight and survival.
Hence, we did not find any evidence for a cost of reproduction on adult survival at the population
level. These results favored the hypothesis that individuals tend to optimize their clutch size
(Pettifor, 1993) and that clutch size is driven by the amount of food as predicted by Lack (1954).
This optimization could explain the absence of link between adult survival and reproduction.
Individual heterogeneity in strategies of resources acquisition and allocation may also mask the trade-off at populational level (van Noordwijk and de Jong, 1986). We envisage exploring the incidence of individual heterogeneity in a future study conducted at the individual level.

Detection probability was relatively high. Posterior distributions are displayed in Appendix 4. The posterior distributions of the replicated survival looked similar to the posterior distributions of the estimated survival. Besides, whatever the year considered, the replicated food peak width or mode, clutch size, chick weight or chick survival, and adult weight were coherent with observations. Both facts suggested an adequate fit of the model (Appendix 2).

**DISCUSSION**

By combining SEM and CR models, we used observational data to test complex scenarios involving demographic parameters that were estimated in presence of imperfect detection of individuals.

We emphasize that CR–SEM is a relevant option when manipulative experiments cannot be conducted, but does not provide evidence of causality. For a given dataset, there could be several competing models that are not falsified. So, adequate fit of a model consistent with a causal hypothesis does not rule out an equally good fit by another model consistent with a different causal assumption. Thus, if a model is corroborated by the data, this does not mean that it has been proven true. For example, omitting variables involved in causal processes may lead to flawed inference via biased parameter estimates and inaccurate estimation of variability around these estimates (Pearl, 2000). Rather than being strictly causal, CR–SEM allows testing hypotheses of causality within a multivariate system based on correlational evidence.

CR–SEM allows modeling quantities of interest that were not measured directly via their expression under the form of latent variables and to analyze the relationships among these latent variables (blue tit example). The possibility of explicitly estimating latent variables has also interesting potential (Austin, 2005). Another advantage of SEM lies in the possibilities to consider indirect effects that may help to account for complex interactions between environment
and life histories (blue tit example), in contrast with standard multiple regression methods. Two particular cases are worth discussing. First, when a single latent variable is involved (blackbird example), the approach is similar to a principal component analysis which is often used in the analysis of selection gradients (Gimenez et al., 2009a). Here the added value of using the CR–SEM approach is to account for uncertainty in parameter estimation and the possibility to test for nonlinearities. Second, when the model does not involve unmeasured (latent) variables, CR–SEM reduces to confirmatory path analysis (Shipley, 2002), which is currently extended to deal with CR data (Gimenez et al., 2011). Finally, we acknowledge that several constraints have to be considered to ensure model identifiability. Necessary conditions exist on the number of paths and latent variables one can consider given the number of observed variables (Kaplan, 2000), which are however not sufficient. Here, we used a Bayesian approach to compare prior to posterior distributions and visually evaluate whether the information contained in the data led to considerable updating of the prior distributions (e.g. Gimenez et al., 2009b).

In conclusion, we provide an integrated framework (CR–SEM) to evaluate causal assumptions about mechanisms underlying individual and / or environmental variability observed in demographic parameters in natural populations. We hope that this approach will help in studying ecological and evolutionary processes occurring in wild populations.

Acknowledgements

We are grateful to all volunteers for contributing to the field work. This research was also supported by the French ANR (ANR–08–JCJC–0028, 08–JCJC–0041 and 09–JCJC–0050) and OSU OREME.

References


Short description of Ecological Archives material

Ecological Archive 1: Model selection
Ecological Archive 2: Posterior predictive checks
Ecological Archive 3: BUGS code to implement the CRSEM approach
Ecological Archive 4: Posterior distributions of the CRSEM parameters (conditional on the covariates being in the model)
Table 1: Parameter estimates for the structural equation model applied to the blackbird and blue tit capture–recapture data. Posterior medians and 95% credible intervals (CI) are provided. See Fig. 1 for notation.

| Parameter | Blackbird | | Blue tit | |
|-----------|-----------|-----------|-----------|
|           | median    | CI        | median    | CI        |
| $\gamma_1$ | -0.54     | [-1.49, 0.31] | 1.32      | [0.74, 2.10] |
| $\gamma_2$ | 0.56      | [-0.09, 1.59] | 2.37      | [1.11, 5.25] |
| $\theta_2$ | 1.03      | [0.80, 1.32] | 11.73     | [7.62, 14.60] |
| $\sigma_1$ | 0.37      | [0.07, 0.75] | 11.40     | [7.27, 14.26] |
| $\theta_4$ | 0.63      | [0.34, 0.99] | 4.50      | [0.23, 13.09] |
| $\sigma_2$ | 0.58      | [0.49, 0.71] | 5.46      | [1.16, 13.46] |
| $\theta_3$ | 0.37      | [0.07, 0.75] | 0.91      | [0.79, 1.03] |
| $\sigma_3$ | 0.63      | [0.34, 0.99] | 0.91      | [0.79, 1.03] |
| $\sigma_4$ | 0.80      | [0.62, 1.02] | 0.80      | [0.62, 1.02] |
| $\sigma_5$ | 0.63      | [0.49, 0.71] | 0.63      | [0.55, 0.85] |
| $\sigma_1$ | 0.58      | [0.49, 0.71] | 0.63      | [0.55, 0.85] |
| $\sigma_2$ | 0.91      | [0.79, 1.03] | 0.91      | [0.79, 1.03] |
| $\sigma_3$ | 0.80      | [0.62, 1.02] | 0.80      | [0.62, 1.02] |
| $\sigma_4$ | 0.63      | [0.49, 0.71] | 0.63      | [0.55, 0.85] |
| $\sigma_5$ | 0.63      | [0.49, 0.71] | 0.63      | [0.55, 0.85] |
Figures

Figure 1: Path diagram for the blackbird (top panel A) and the blue tit (bottom panel B) case studies. Rectangles are for latent variables while ellipses are for observed variables.