

# Testing hypotheses in evolutionary ecology with imperfect detection: capture–recapture structural equation modeling

SARAH CUBAYNES,<sup>1,2,4</sup> CLAIRE DOUTRELANT,<sup>1</sup> ARNAUD GRÉGOIRE,<sup>1</sup> PHILIPPE PERRET,<sup>1</sup> BRUNO FAIVRE,<sup>3</sup>  
AND OLIVIER GIMENEZ<sup>1</sup>

<sup>1</sup>Centre d'Ecologie Evolutive et Fonctionnelle UMR 5175, 1919 Route de Mende, 34293 Montpellier, Cedex 5, France

<sup>2</sup>Institut de Mathématiques et Modélisation de Montpellier, UNR 5149, Place Eugène Bataillon, 34095 Montpellier, Cedex 5, France

<sup>3</sup>Biogéosciences, Université de Bourgogne UMR 5561, 6 Boulevard Gabriel, 21000 Dijon, France

**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 the imperfect detection of individuals to provide unbiased demographic parameter estimates. To cope with these issues, we developed 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 (1) estimate model parameters, (2) implement a model selection procedure to evaluate competing hypotheses about causal mechanisms, and (3) 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 (*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.

**Key words:** 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 population 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 structural equation model 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 the testing of competing causal assumptions and refuting of an unlikely hypothesis, based on correlational evidence. Hence, the term “causal model” should be understood to be 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).

Manuscript received 12 February 2011; revised 26 August 2011; accepted 14 September 2011. Corresponding Editor: J. R. Sauer.

<sup>4</sup> E-mail: sarah.cubaynes@gmail.com

The flexibility of SEM to represent complex scenarios involving several observed and latent variables has led 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  $<1$  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 the 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 provide information about 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 used an 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 Markov chain Monte Carlo (MCMC) sampling procedure, which provides powerful computer-intensive methods for handling complex models and are becoming increasingly

popular in SEM (Congdon 2006, Lee 2007, 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 using this model over the total number of iterations. We also reported the relative importance of a particular factor by calculating the number of iterations using a model containing the corresponding parameter over the total number of iterations (Appendix A). In addition, we calculated posterior predictive checks to evaluate the fit of the model to the data (Appendix B). The simulations were performed using JAGS, a program in R (Ihaka and Gentleman 1996), using the package *rjags* (Plummer 2003; JAGS program available online).<sup>5</sup>

#### *Level 1: modeling survival using CR data*

To account for the issue of detectability  $<1$ , 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  $X$  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 distrib-

<sup>5</sup> <http://mcmc-jags.sourceforge.net/>

uted 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:

$$\mathbf{x} = \mathbf{\Theta} \xi + \delta \quad (1)$$

which related the observed variables  $\mathbf{x}$  to a vector of latent exogenous variables  $\xi$  through the matrix  $\mathbf{\Theta}$  of regression parameters; we also had

$$\mathbf{y} = \mathbf{\Lambda} \eta + \varepsilon \quad (2)$$

which related the observed variables  $\mathbf{y}$  to a vector of latent endogenous variables  $\eta$  through the matrix  $\mathbf{\Lambda}$  of regression parameters. The measurement errors  $\delta$  and  $\varepsilon$  were 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 blackbirds (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  $\mathbf{y}$ . In the second example on Blue Tits (see *Applications*), 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  $\mathbf{x}$ .

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

$$\eta = \mathbf{B} \eta + \mathbf{\Gamma} \xi + \zeta \quad (3)$$

where the matrix  $\mathbf{B}$  captures the relationships among the latent endogenous variables, while the matrix  $\mathbf{\Gamma}$  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 latter 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 data set 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 that we called “overall size” (Fig. 1A). Morphological traits were different ways of measuring body size that 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., better-surviving individuals 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).

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  $\delta_1, \delta_2 \sim \mathcal{N}(0, \sigma_1^2)$  as  $x_1$  and  $x_2$  were measured with the same tool, and  $\delta_3, \delta_4 \sim \mathcal{N}(0, \sigma_2^2)$  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  $\Lambda = 1$  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  $\xi \sim \mathcal{N}(0, \sigma_\xi^2)$ , where  $\gamma$  is a regression parameter capturing the effect of the exogenous variables on the endogenous variables. For identifiability issues, we specified  $\zeta \sim \mathcal{N}(0, 1)$  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  $\xi$ ; 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

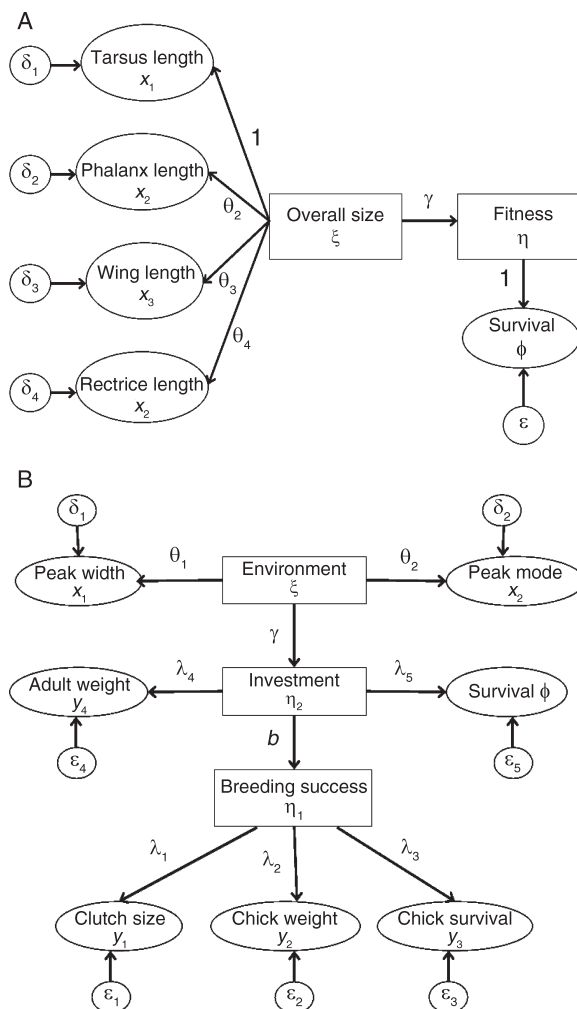


FIG. 1. Path diagram for (A) the Blackbird and (B) the Blue Tit case studies. Rectangles are for latent variables, while ellipses are for observed variables and arrows indicate causality. The  $\delta$ 's are measurement errors associated with observed variables connected to exogenous variables, and the  $\varepsilon$ 's are measurement errors associated with endogenous variables. The  $\theta$ 's are regression parameters relating observed variables to exogenous variables, and the  $\lambda$ 's are regression parameters relating observed variables to endogenous variables. The parameter  $\gamma$  is a regression parameter capturing the effect of the exogenous variables on the endogenous variables, while  $b$  is a regression parameter capturing the relationships among the latent endogenous variables.

probabilities,  $U(0, 1)$ , and normal distributions with large variance for the regression parameters,  $\mathcal{N}(0, 1000)$ . 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 5000 iterations that resulted in acceptable mixing



TABLE 1. Parameter estimates for the structural equation model applied to the Blackbird and Blue Tit capture–recapture data.

Parameter	Median	CI
<b>Blackbird</b>		
$\gamma_1$	−0.54	[−1.49, 0.31]
$\gamma_2$	0.56	[−0.09, 1.59]
$\theta_2$	1.03	[0.80, 1.32]
$\theta_3$	0.37	[0.07, 0.75]
$\theta_4$	0.63	[0.34, 0.99]
$\sigma_1$	0.58	[0.49, 0.71]
$\sigma_2$	0.91	[0.79, 1.03]
$\sigma_3$	0.80	[0.62, 1.02]
$p$	0.70	[0.55, 0.85]
<b>Blue Tit</b>		
$\gamma$	1.32	[0.74, 2.10]
$b$	2.37	[1.11, 5.25]
$\theta_1$	11.73	[7.62, 14.60]
$\theta_2$	11.40	[7.27, 14.26]
$\sigma_{\delta 1}$	4.50	[0.23, 13.09]
$\sigma_{\delta 2}$	5.46	[1.16, 13.46]
$\lambda_1$	1.10	[0.48, 2.16]
$\lambda_2$	0.03	[0.01, 0.06]
$\lambda_3$	0.15	[0.06, 0.30]
$\lambda_4$	4.08	[2.70, 6.02]
$\lambda_5$	0.20	[−0.03, 0.48]
$\sigma_{e1}$	1.08	[0.10, 2.24]
$\sigma_{e2}$	0.13	[0.09, 0.20]
$\sigma_{e3}$	0.29	[0.17, 0.48]
$\sigma_{e4}$	0.84	[0.06, 2.62]
$\sigma_{e5}$	5.46	[1.16, 13.46]
$p$	0.66	[0.45, 0.83]

Notes: Values are posterior medians and 95% credible intervals (CI). See Fig. 1 for notation. Parameter  $p$  is the detection probability, and the  $\sigma$ 's are standard deviations of the error terms distributions.

and convergence. Model code is provided in the Supplement. 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 molt, 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 nonzero vs. 0.54 for quadratic term only nonzero. These results revealed a quadratic effect of the factor size on survival, which suggested a disruptive selection, favoring the smallest and largest 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 C.

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 B).

### 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 1.

We used data on 755 breeding individuals that were banded, released, and recaptured in spring during the breeding seasons between 1984 and 2000, in Pario, 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 mass 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 mass 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 Tit 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  $\delta_j \sim \mathcal{N}(0, \sigma_{\delta_j}^2)$ ,  $j = 1, 2$ .

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 mass, 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  $\mathbf{y} = (y_1, y_2, y_3, y_4, \phi)'$ ,  $\boldsymbol{\eta} = (\eta_1, \eta_2)'$ , with regression coefficients

$$\Lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 & \lambda_5 \end{bmatrix}$$

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

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

with the 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  $\mathcal{N}(0, 1)$ .

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.

The 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 that resulted in acceptable mixing and convergence. Model code is provided in the Supplement. The three variables clutch size, chick mass, 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 mass and nest survival, suggesting that variation of parental investment in reproduction affected mainly clutch size rather than chick mass 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 a link between adult survival and reproduction. Individual heterogeneity in strategies of resources acquisition and allocation may also mask the trade-off at the 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 C. 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 mass, or chick survival, and adult mass were coherent with observations. Both facts suggested an adequate fit of the model (Appendix B).

## 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 data set, 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.

#### ACKNOWLEDGMENTS

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. We also thank the Blue Tit team and notably Jacques Blondel, Don Thomas, Anne Charmentier, and Marcel Lambrechts for the long-term field data.

#### LITERATURE CITED

- Austin, E. 2005. Latent variable models: an introduction to factor, path and structural analysis. *British Journal of Mathematical and Statistical Psychology* 58:190–191.
- Blondel, J., D. W. Thomas, A. Charmentier, P. Perret, P. Bourgault, and L. M. M. 2006. A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics. *BioScience* 56:661–674.
- Bollen, K. 1989. Structural equations with latent variables. Wiley, New York, New York, USA.
- Buoro, M., E. Prévost, and O. Gimenez. 2010. Investigating evolutionary trade-offs in wild populations of Atlantic salmon (*Salmo salar*): incorporating detection probabilities and individual heterogeneity. *Evolution* 64:2629–2642.
- Choquet, R., L. Rouan, and R. Pradel. 2009. Program E-SURGE: a software application for fitting multievent models. Pages 845–865 in D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Environmental and Ecological Statistics Series. Volume 3. Springer, New York, New York, USA.
- Congdon, P. 2006. Bayesian statistical modelling. Second edition. Wiley, New York, New York, USA.
- Faivre, B., M. Prévost, M. Théry, J. Secondia, B. Patrisa, and F. Cézilly. 2001. Breeding strategies and morphological characters in an urban population of blackbirds, *Turdus merula*. *Animal Behaviour* 61:969–974.
- Gelman, A. 1996. Inference and monitoring convergence. Page 131–143 in W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. Markov chain Monte Carlo in practice. Chapman and Hall, London, UK.
- Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1:515–533.
- Gimenez, O., T. Anker-Nilssen, and V. Grosbois. 2011. Exploring causal pathways in demographic parameter variation: path analysis of mark-recapture data. *Methods in Ecology and Evolution*. <http://dx.doi.org/10.1111/j.2041-210X.2011.00150.x>
- Gimenez, O., A. Grégoire, and T. Lenormand. 2009a. Estimating and visualizing fitness surfaces using mark-recapture data. *Evolution* 63:3097–3105.
- Gimenez, O., B. Morgan, and S. Brooks. 2009b. Weak identifiability in models for mark-recapture–recovery data. Pages 1055–1067 in D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Environmental and Ecological Statistics Series. Volume 3. Springer, New York, New York, USA.
- Gimenez, O., V. Rossi, R. Choquet, C. Dehais, B. Doris, H. Varella, J.-P. Vila, and R. Pradel. 2007. State-space modelling of data on marked individuals. *Ecological Modelling* 206:431–438.
- Gimenez, O., A. Viallefont, A. Charmentier, R. Pradel, E. Cam, C. Brown, M. Anderson, M. B. Brown, R. Covas, and J.-M. Gaillard. 2008. The risk of flawed inference in evolutionary studies when detectability is less than one. *American Naturalist* 172:441–448.
- Grace, J. 2006. Structural equation modeling and natural systems. Cambridge University Press, Cambridge, UK.
- Grace, J., T. Anderson, H. Olf, and S. Scheiner. 2010. On the specification of structural equation models for ecological systems. *Ecological Monographs* 80:67–87.
- Ihaka, R., and R. Gentleman. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5:299–314.
- Kaplan, D. 2000. Structural equation modeling: foundations and extensions. Sage Publications, Newbury Park, California, USA.

- Kirk, K. M., S. P. Blomberg, D. L. Duffy, A. C. Heath, I. P. F. Owens, and N. G. Martin. 2001. Natural selection and quantitative genetics of life-history traits in western women: a twin study. *Evolution* 55:423–435.
- Kuo, L., and B. Mallick. 1998. Variable selection for regression models. *Sankhya Series B* 60:65–81.
- Lack, D. 1954. The natural regulation of animal numbers. Oxford University Press, Oxford, UK.
- Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- Lebreton, J.-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs* 62:67–118.
- Lee, S.-Y. 2007. Structural equation modeling: a Bayesian approach. Wiley, New York, New York, USA.
- Lindström, J. 1999. Early development and fitness in birds and mammals. *Trends in Research in Ecology and Evolution* 14:343–348.
- Mitchell, R. J. 1992. Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. *Functional Ecology* 6:123–129.
- Palomo, J., D. Dunson, and K. Bollen. 2007. Bayesian structural equation modeling. Pages 163–188 in S.-Y. Lee, editor. *Handbook of latent variable and related models*. Elsevier Publishing, Amsterdam, The Netherlands.
- Pearl, J. 2000. *Causality: models, reasoning, and inference*. Cambridge University Press, Cambridge, UK.
- Pettifor, R. 1993. Brood-manipulation experiments. II. A cost of reproduction in blue tits (*Parus caeruleus*)? *Journal of Animal Ecology* 62:145–159.
- Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20–22, Vienna, Austria.
- Pollock, K. H. 2002. The use of auxiliary variables in capture-recapture modelling: an overview. *Journal of Applied Statistics* 29:85–102.
- Pugsek, B., and A. Tomer. 1996. The Bumpus house sparrow data: a reanalysis using structural equation models. *Evolutionary Ecology* 10:387–404.
- Pugsek, B., A. Tomer, and A. von Eye. 2003. *Structural equation modeling: applications in ecological and evolutionary biology*. Cambridge University Press, Cambridge, UK.
- Royle, J. 2008. Modeling individual effects in the Cormack–Jolly–Seber model: a state-space formulation. *Biometrics* 64:364–370.
- Schofield, M. R., and R. J. Barker. 2008. A unified capture–recapture framework. *Journal of Agricultural, Biological and Environmental Statistics* 13:458–477.
- Shipley, B. 2002. *Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference*. Second edition. Cambridge University Press, Cambridge, UK.
- Shipley, B. 2009. Confirmatory path analysis in a generalized multilevel context. *Ecology* 90:363–368.
- Stearns, S. 1992. *The evolution of life histories*. Oxford University Press, Oxford, UK.
- van Noordwijk, A., and G. de Jong. 1986. Acquisition and allocation of resources: their influence on variation in life history tactics. *American Naturalist* 128:137–142.
- Zandt, H. 1994. A comparison of three sampling techniques to estimate the population size of caterpillars in trees. *Oecologia* 97:399–406.

## SUPPLEMENTAL MATERIAL

### Appendix A

Model selection (*Ecological Archives* E093-024-A1).

### Appendix B

Posterior predictive checks (*Ecological Archives* E093-024-A2).

### Appendix C

Posterior distributions of the CR-SEM parameters (conditional on the covariates being in the model) (*Ecological Archives* E093-024-A3).

### Supplement

BUGS code to implement the CR-SEM approach (*Ecological Archives* E093-024-S1).