RESEARCH METHODS GUIDE

Distinguishing within- from between-individual effects: How to use the within-individual centring method for quadratic patterns

Rémi Fay[1,2] 🆔 | Julien Martin[3] 🆔 | Floriane Plard[1,4] 🆔

[1]Swiss Ornithological Institute, Sempach, Switzerland

[2]Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway

[3]Department of Biological Sciences, University of Ottawa, Ottawa, Canada

[4]Department of Aquaculture and Fish Biology, Hólar University, Háeyri, Iceland

Correspondence
Rémi Fay
Email: fay.remi@gmail.com

Handling Editor: Roberto Salguero-Gómez

Abstract

1. Any average pattern observed at the population level (cross-sectional analysis) may confound two different types of processes: some processes that occur among individuals and others that occur within individuals. Separating within- from among-individual processes is critical for our understanding of ecological and evolutionary dynamics.

2. The within-individual centring method allows distinguishing within- from among-individual processes and this method has been widely used in ecology to investigate both linear and quadratic patterns. Here we show that two alternative equations could be used for the investigation of quadratic within-individual patterns. We explain the different assumptions and constraints of both equations. Reviewing the literature, we found that mainly one of these two equations has been used in studies investigating quadratic patterns. Yet this equation might not be the most appropriate in all circumstances leading to bias and imprecision.

3. We show that these two alternative equations make different assumptions about the shape of the within-individual pattern. One equation assumes that the within-individual effect is related to an absolute process whereas the other assumes the effect arises from an individual relative process. The choice of using one equation instead of the other should depend upon the biological process investigated.

4. Using simulations, we showed that a mismatch between the assumptions made by the equation used to analyse the data and the biological process investigated might led to flawed inference affecting output of model selection and accuracy of estimates. We stress that the equation used should be chosen carefully. We provide step by step guidelines for choosing an equation when studying quadratic pattern with the within-individual centring approach. We encourage the use of the within-individual centring method, promoting its relevant application for non-linear relationships.

1 | INTRODUCTION

Ecological data are fundamentally hierarchical data and separating processes occurring at different levels such as individuals, populations, species or communities is a recurrent challenge in this field (Bolnick et al., 2011; Hamel et al., 2018; Vindenes & Langangen, 2015). As processes driving trait variation may vary across levels, patterns observed at one level could differ from those occurring at other levels. In case of data collected over several individuals, the patterns observed at the population level do not necessarily reflect the changes occurring at the individual level, and vice versa, because both within- and among-individual effects drive trait variation (Vaupel et al., 1979; Vaupel & Yashin, 1985). A famous example is the investigation of the relationships between a response variable, let say breeding success, and individual age (Nussey et al., 2008; van de Pol & Verhulst, 2006). Variation in breeding success may arise from within-individual variation, for example, ageing effect due to accumulating experience in early life or/ and due to senescence in late life, as well as from among-individual variation, by the selective disappearance of frail individuals (Aubry et al., 2009; Bouwhuis et al., 2009; Hämäläinen et al., 2014). Indeed, in heterogeneous populations, frail individuals are expected to die first, leaving robust individuals to be over-represented at old ages (Pigeon et al., 2017; Vedder & Bouwhuis, 2018). This change in the composition of the population over time generates positive age-related relationships at the population level. Thus, the average age-related trajectory observed at the population level (cross-sectional analysis) could differ from the average trajectory observed at the individual level, because independent processes, arising within and among individuals, co-occur (Vaupel & Yashin, 1985).

This problem of distinguishing between within-individual and among-individual patterns is not specific to age-related patterns. It may concern any type of traits, that is, behavioural, morphometric, physiological, environmental, each time several observations per individual are obtained and aggregated (e.g. Dammhahn et al., 2017; Morrongiello et al., 2019; Pick et al., 2016; Siracusa et al., 2019). Analysing such aggregated data, we need to distinguish the within-individual effect, that is, how variation in the explanatory variable X within an individual affects the response variable Y, from the among-individual effect, that is, how the average value of X for a given individual is related to the average value of Y. In this manuscript we focus on the individual level, but similar distinctions could be relevant when comparing patterns within and among clusters of data at other scales, for example, groups of individuals, species, plots embedded within replicate patches.

From an eco-evolutionary perspective, it is crucial to separate within- from among-individual variation in traits (Bolnick et al., 2011). For example, the investigation of life-history trade-offs, which are

within-individual effects, requires accounting for among-individual variation in resource acquisition (van Noordwijk & de Jong, 1986). Similarly, our understanding of the age-related change in demographic rates has been limited for a long time because most studies considered only the population-level pattern (Forslund & Pärt, 1995; Nussey et al., 2008). Patterns within and among individuals provide different information which are of interest to understand the behaviour, physiology or demography of biological systems. A within-individual effect may provide information on processes involving individual plasticity such as phenotypic change, learning or the occurrence of trade-off. On the other hand, among-individual effects depict consistent inter-individual differences and thus may suggest among-individual differences in fitness, personality or pace of life, for instance (Dingemanse & Dochtermann, 2013; Reid et al., 2010; Wilson & Nussey, 2010). Ignoring the distinction between within- and among-individual effects may thus lead to flawed inferences confounding processes occurring at different levels (Kendall et al., 2011; Stover et al., 2012; van de Pol & Verhulst, 2006; Vaupel & Yashin, 1985).

A statistical method to distinguish within- from among-individual effects is the within-individual centring method (Hofmann & Gavin, 1998; Kreft et al., 1995; Snijders & Bosker, 1999). This powerful method is simple to use and gained popularity rapidly following its introduction to ecologists (van de Pol & Wright, 2009). To date, this work has been cited 479 times and the approach has been used in 346 published studies (research among the papers citing van de Pol & Wright, 2009 using Web of Science in January 2021). While this method has been presented in the context of behavioural studies for linear patterns (van de Pol & Wright, 2009), it has been used in various ecological fields including demography, ageing, eco-physiology, and parasitology, and has been extended to investigate quadratic patterns (Figure 1).

In this paper we clarify the use of the within-individual centring method for quadratic patterns providing guidelines for its application. After briefly presenting the within-individual centring approach in case of linear patterns, we present two alternative equations which can be used for the investigation of quadratic within-individual effects. We explain the different assumptions made by each equation about the shape of the within-individual pattern and suggest when each equation should be used. We stress that previous studies investigating quadratic patterns failed to directly recognize and discuss these assumptions and, in some cases, may not have used the most appropriate equation. Based on simulations, we assess the consequences of using an equation mismatching the data generation process on model selection and quality of estimates (bias and precision). We also address the specific case of quadratic age patterns, which, to date, correspond to the most frequent individual-centred variable used when investigating quadratic individual patterns.

All explanatory variables

Age as the explanatory variable
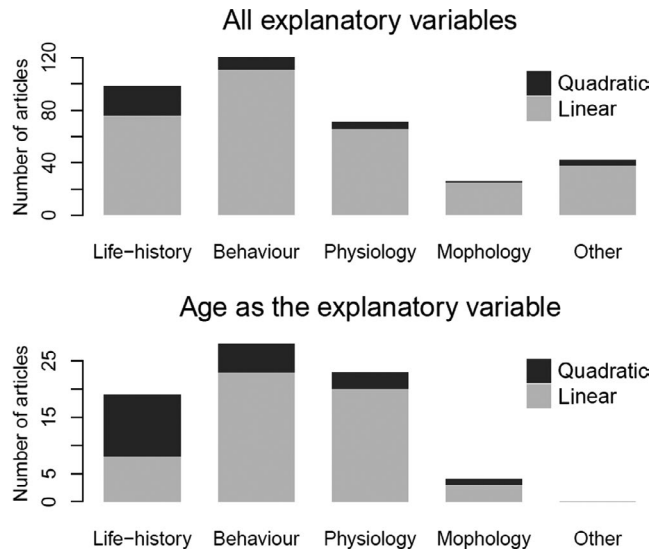
FIGURE 1 Number of articles citing the van de Pol and Wright's (2009) work and using the within-individual centring method including all explanatory variables (upper panel) and age in particular with linear (light grey) and quadratic (dark grey) models. The articles have been classified according to the type of response variable(s) used (life-history, behavioural, physiological, morphological and other traits)

2 | THE WITHIN-INDIVIDUAL CENTRING METHOD

In ecology, linear mixed models are nowadays standard tools to analyse data aggregated at different scales. However, such models may fail to estimate the true within-individual changes because the individual random effect may not capture the total among-individual effects (van de Pol & Wright, 2009). The within-individual centring method is an extension of a simple mixed model that allows distinguishing explicitly within- from among-individual variation using additional fixed effects (Hofmann & Gavin, 1998; Kreft et al., 1995; Snijders & Bosker, 1999; van de Pol & Wright, 2009). We present only briefly this method and refer readers to van de Pol and Wright (2009) for further information.

The standard random effect model can be described by the following regression equation:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}. \tag{1}$$

where β_0 is the intercept, β_1 is the slope, x_{ij} is the value of a trait x at measurement i from individual j, u_j is the deviation from the intercept for individual j assumed to be drawn from a normal distribution with a mean of zero and among-subject variance σ_u^2, and e_{ij} is the residual error for each measurement assumed to be drawn from a normal distribution with a mean of zero and variance σ_e^2.

Starting from this standard mixed model, the within-individual centring approach decomposes the effect of x in two terms: the within-individual and the among-individual effects. The among-individual effect is fitted as the average x value for each individual

($\bar{x}_j$). The within-individual effect is fitted as the deviation from the individual mean for a given observation ($x_{ij} - \bar{x}_j$). The previous equation becomes:

$$y_{ij} = \beta_0 + \beta_W \left(x_{ij} - \bar{x}_j \right) + \beta_B \bar{x}_j + u_j + e_{ij}, \tag{2}$$

where β_W is the slope of the within-individual effect and β_B is the slope of the among-individual effect. Thus, Equation 1 corresponds to a specific case of Equation 2 when $\beta_W = \beta_B$ assuming that among- and within-individual effects are identical, or when the among-individual effect is null. van de Pol and Wright (2009) showed that a standard mixed model does not allow estimation of within-individual effects as reliably as the within-individual centring model. Note that Equation 2 does not estimate individual variation in the within-individual effect. Although individual variation is estimated for the intercept, the within-individual effect (β_W) is assumed to be the same for all the individuals making this equation more appropriate to describe average within-individual patterns. When the focus is on individual variation in within-individual effect, Equation 2 needs to be modified to include random slope for the within-individual effect (see Dingemanse & Dochtermann, 2013; Nussey et al., 2007; van de Pol & Wright, 2009; Westneat et al., 2020).

3 | THE WITHIN-INDIVIDUAL CENTRING METHOD TO INVESTIGATE QUADRATIC PATTERNS

The within-individual centring method has been regularly used in ecology to investigate quadratic changes in traits (11.5% of all the articles citing van de Pol & Wright, 2009 and applying this method, Figure 1). Quadratic terms could be added to both the within- and the among-individual effects independently since they can arise from different processes. Among the published articles using a quadratic relationship, all introduced it for the within-individual effect but only some of those used it simultaneously for the among-individual effect. The relevance of including quadratic terms to the within-individual effect, the among-individual effect, or both, depends on the system studied and the question investigated. Here, for the sake of clarity, we first present how quadratic terms can be added to the within- and among-individual effects separately. Then we give the general equation including a quadratic term for both within- and among-individual effects.

3.1 | Quadratic within-individual effect

To use an individual centring approach including a quadratic within-individual effect, the individual centring should be applied on both linear and quadratic terms. We suggest the following equation to estimate within-individual quadratic effect:

$$y_{ij} = \beta_0 + \beta_{W1} \left(x_{ij} - \bar{x}_j \right) + \beta_B \bar{x}_j + \beta_{W2} \left(x_{ij}^2 - \bar{x}_j^2 \right) + u_j + e_{ij}. \tag{3}$$

An alternative equation which has been used by most of the studies investigating quadratic individual trajectories is:

$$y_{ij} = \beta_0 + \beta_{W1}\left(x_{ij} - \overline{x}_j\right) + \beta_B\overline{x}_j + \beta_{W2}\left(x_{ij} - \overline{x}_j\right)^2 + u_j + e_{ij}. \quad (4)$$

In Equation 3, the quadratic within-individual effect (β_{W2}) is included as the deviation of the squared x value from the squared average x value of an individual, that is, $x_{ij}^2 - \overline{x}_j^2$ (see Appendix S1 for full derivation and explanation of the equation), whereas in Equation 4, the quadratic within-individual effect (β_{W2}) is included as the square of the centred variable, that is, $\left(x_{ij} - \overline{x}_j\right)^2$. It should be noted that Equations 3 and 4 are not equivalent and generate different individual patterns (Figure 2).

To better understand the difference between these two equations, we can develop and rearrange the quadratic term in Equation 4 as:

$$y_{ij} = \beta_0 + \beta_{W1}\left(x_{ij} - \overline{x}_j\right) + \beta_B\overline{x}_j + \beta_{W2}\left(x_{ij}^2 + \overline{x}_j^2\right) - 2\beta_{W2}x_{ij}\overline{x}_j + u_j + e_{ij}. \quad (4b)$$

From Equation 4b, we can see that a difference between Equations 3 and 4 is the presence of the term $-2\beta_{W2}x_{ij}\overline{x}_j$ in Equation 4 only. This term reveals a latent interaction between x_{ij}, the value of a trait x at measurement i from individual j, and $\overline{x}_j$, the average x value sampled for individual j. By doing so, Equation 4 assumes that the quadratic within-individual effect (β_{W2}) is not related to the absolute value of X but depends on the average $\overline{x}_j$ value sampled for each individual. Consequently, the within-individual effect depends on the relative value of X (deviation from the subject mean). Thus, Equation 4 assumes that the within-individual response to X vary among individuals according to $\overline{x}_j$. For instance, looking at the Figure 2b, we see that the

within-individual response may increase or decrease when X values range between −3 and 0 according to the average X value of the individuals. In contrast, because there is no interaction between x_{ij} and $\overline{x}_j$ in Equation 3, this equation assumes that the within-individual effect of X on Y depends on the absolute value of X. Looking at the Figure 2a, we see that the individual trajectories are parallel meaning that the within-individual response vary with the same rate for all individuals for a given range of X values. In other words, Equation 3 assumes that there is a general quadratic pattern over the entire range of X which is shared by all the individuals.

As a consequence, the two equations differ in where the quadratic pattern is expected to reach its maximum/minimum within each individual. Equation 3 assumes that the maximum/minimum is identical for all individuals on the absolute scale of X values, independently of the ranges of each individual trajectory. Equation 4 assumes that the maximum/minimum is different among individuals but is reached at the same relative place within the range of the explanatory variable sampled for each individual (mean average measured for each individual, for instance). See Appendix S2 for a mathematical demonstration of this difference.

3.2 | Quadratic among-individual effect

To use an individual centring approach including a quadratic among-individual effect, Equation 2 becomes:

$$y_{ij} = \beta_0 + \beta_W\left(x_{ij} - \overline{x}_j\right) + \beta_{B1}\overline{x}_j + \beta_{B2}\overline{x}_j^2 + u_j + e_{ij}. \quad (5)$$
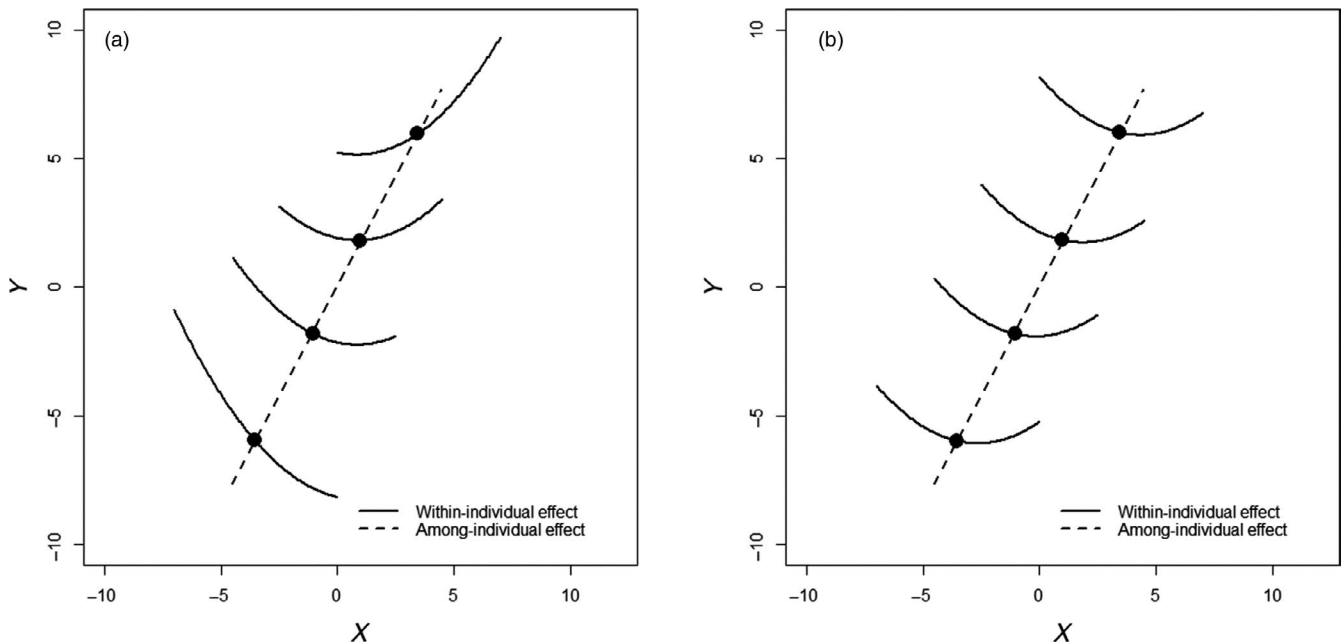
FIGURE 2 Individual trajectories illustrating quadratic within-individual effects simulated with Equation 3 (a) and Equation 4 (b). In both graphs, the four individual life histories have been simulated using the Equations 3 and 4, respectively, with the same following parameters: $\beta_0 = 0$, $\beta_{W1} = -0.2$, $\beta_{W2} = 0.12$, $\beta_B = 1$, $\sigma_u^2 = 1$ and $\sigma_e^2 = 0$. Black circles show $\overline{x}_j$, the average X values for each individual. The Equation 3 (a) assumes that all individuals show the same pattern over the range of the explanatory variable. In contrast, Equation 4 (b) assumes that the within-individual changes of a given individual depend on the range of the explanatory variable sampled for this individual

The quadratic among-individual effect (β_{B2}) is fitted by the square of the average x value for each individual ($\bar{x}_j^2$) (Figure 3). Because including a quadratic among-individual effect is not a source of confusion, the rest of the article focuses on the quadratic within-individual effect.

3.3 | Quadratic within-and among-individual effects

To use an individual centring approach including both a quadratic within-individual effect and a quadratic among-individual effect, the two alternative equations are:

$$y_{ij} = \beta_0 + \beta_{W1}\left(x_{ij} - \bar{x}_j\right) + \beta_{B1}\bar{x}_j + \beta_{W2}\left(x_{ij}^2 - \bar{x}_j^2\right) + \beta_{B2}\bar{x}_j^2 + u_j + e_{ij}, \quad (3b)$$

and

$$y_{ij} = \beta_0 + \beta_{W1}\left(x_{ij} - \bar{x}_j\right) + \beta_{B1}\bar{x}_j + \beta_{W2}\left(x_{ij} - \bar{x}_j\right)^2 + \beta_{B2}\bar{x}_j^2 + u_j + e_{ij}. \quad (4c)$$

A key difference between the two equations is that when within- and among-individual effects are equal, that is, $\beta_{W1} = \beta_{B1}$ and $\beta_{W2} = \beta_{B2}$, only Equation 3b is equivalent to a model not using within-individual centring. Indeed, assuming identical within- and among-individual effect, we may simplify Equations 3b and 4c as follows:

$$y_{ij} = \beta_0 + \beta_{W1}x_{ij} + \beta_{W2}x_{ij}^2 + u_j + e_{ij}, \quad (3c)$$
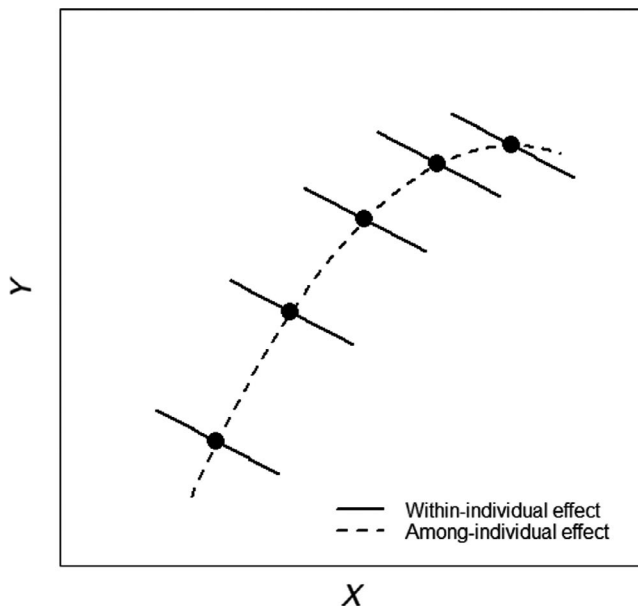
FIGURE 3 Individual trajectories illustrating a quadratic among-individual effect. Here, Y decreases linearly in relation to X within each individual but Y increases quadratically among individuals in relation to the average $\bar{x}_j$ values. Individual trajectories are simulated with Equation 5 with the following parameters: $\beta_0 = 0$, $\beta_W = -0.5$, $\beta_{B1} = 1.2$, $\beta_{B2} = -0.08$, $\sigma_u^2 = 1$ and $\sigma_e^2 = 0$. Black circles show $\bar{x}_j$

and

$$y_{ij} = \beta_0 + \beta_{W1}x_{ij} + \beta_{W2}x_{ij}^2 + 2\beta_{W2}x_{ij}\bar{x}_j + u_j + e_{ij}. \quad (4d)$$

Thus, while Equation 3c is a generalization of a classical quadratic regression, Equation 4d is not.

4 | THEORETICAL EXAMPLES OF BIOLOGICAL PROCESSES FITTING EACH EQUATION

Equation 3 assumes that the within-individual effect is generated by an absolute process. As a consequence, all the individuals show a common quadratic pattern over the range of the explanatory variable. Equation 3 can thus be used when researchers expect a functional relationship shared by all the individuals regardless of the mean value of the explanatory variable sampled for each individual (Figures 2a and 4a). Taking the example of the flying speed performance in an insect, it is reasonable to expect similar within-individual responses to variation in temperature. Because insects are ectotherms, flying performance of all individuals is expected to increase with temperature until a certain point where extreme temperature may have a negative effect. Here, we may expect a general quadratic pattern over the entire range of temperature shared by all the individuals (Figure 4a). The effect of temperature on flying speed should depend on the absolute value of the temperature, not on the deviation from the average temperature sampled for each individual. In this case, using Equation 3 should be suitable to estimate the quadratic within-individual effect. More generally, Equation 3 could be used any time the within-individual response depends on the absolute value of the explanatory variable.

In contrast, Equation 4 assumes that the within-individual effect is generated by a relative process since each individual has a quadratic pattern that depends on its individual mean. Continuing with the example of flying performance in an insect, an example of biological process that could fit Equation 4 is the effect of body mass on flying performance. In this case, absolute body mass might not be critical to determine flying performance, but rather the deviation from the optimum individual body mass (Figure 4b). As a consequence, patterns are centred on individual means in this example. The same change in body mass may be associated with higher or lower flying speed according to the size of the individuals. More generally, Equation 4 assumes that the effect of the explanatory variable varies among individuals. This equation could thus be used to investigate relative processes which depend on individual features or past experienced conditions. It could correspond, for instance, to an acclimation process when individuals have optimal performance based on the environment they experienced in the past.

It should be acknowledged that biological processes could be generated by a mix of both absolute and relative processes. In that case, the choice of using Equations 3 or 4 should be done determining which process best fit the data (see the recommendations in Section 8).
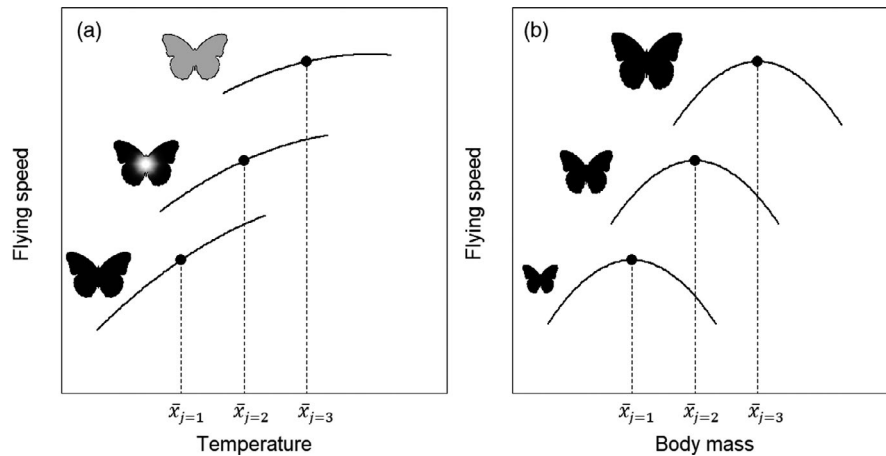
FIGURE 4 Theoretical effect of temperature (a) and body mass (b) on flying speed in a butterfly. These examples illustrate relationships following assumptions made by Equations 3 and 4 respectively. Variation in flying speed is due to both a within-individual effect, that is, effect of variation in temperature or body mass on individual performance, and an among-individual effect generated by variation in cohort (a) and size (b), for instance. In accordance with assumptions made by Equation 3, all individuals show the same relationship between flying speed and temperature. In other words, the effect of the temperature on flying speed is independent of the range of temperature encountered by each individual. In contrast, variation in flying speed according to body mass shows a strong individually specific response. Each quadratic pattern is centred on the mean body mass of each individual. In this case it is the mass deviation related to the individual optimum, not the absolute mass value, which is critical

5 | CONSEQUENCES OF MISMATCHES BETWEEN GENERATING AND ANALYTICAL EQUATIONS

Above we showed that two alternative equations could be used for the investigation of quadratic within-individual effects. Here we illustrate the consequences of a mismatch between the equation generating the quadratic within-individual patterns and the equation used to analyse the data.

5.1 | Methods

We simulated 1,000 datasets consisting of 50 individuals with 15 measurements each. First, we used a normal distribution to simulate environmental values for each individual, X_j, with a mean $\overline{X}_j$, taken randomly between −7 and 7, and a variance 1.5 to sample 15 measurements. Then, we simulated the response variable Y using either Equations 3 or 4. In both cases, we used the same parameters: $\beta_0 = 0$, $\beta_{W1} = 0.4$, $\beta_{W2} = -0.11$, $\beta_{B1} = 1$, $\beta_{B2} = 0$, $\sigma_u = 2$ and $\sigma_e^2 = 0.5$. Finally, for each simulated dataset, we fitted three models: the matching and mismatching models (Equations 3 and 4) and a simpler linear model (Equation 2).

Based on these sets of simulations, we assessed the consequences for model selection and quality of estimates of analysing the data with the mismatching equation. We used the Akaike information criterion (AIC) to compare matching, mismatching and linear models. This comparison allows us to check if a mismatch between the generating and analytical equations affects our ability to detect quadratic within-individual effects. When differences of AIC between the two models were below 2, we followed the principle of parsimony, selecting the model with the lower number of

parameters. To assess the effect of mismatches between generating and analytical equations on the quality of the estimates, we computed bias by subtracting the estimated parameters from the true parameter values used to generate the data. Results were summarized with violin plots to visualize both mean and dispersion of bias. We also computed the mean square error (MSE) which gives the accuracy of the estimates, that is, combination of bias and precision to define the performance of an estimator. Finally, we computed the 95% confidence interval coverage that is the proportion of confident intervals which include the true parameter value used to generate the data.

5.2 | Results

Quadratic models matching the equation used to simulate the data were always selected as the best of the three models (lowest AIC). Mismatching quadratic models were always rejected (ΔAIC ≥144). However, simpler linear models were selected against a mismatching quadratic model in 3% of the data simulated with Equation 3 and analysed with Equation 4 and in 38% of the data simulated with Equation 4 and analysed with Equation 3. This demonstrates that using a mismatched equation may lead to flawed inferences.

Regarding the performance of the estimators, no bias was observed when the equation used to generate the data matched the one used to analyse the data (Equation 3 bias $\beta_{w1} = -0.0005[-0.027:0.024]$ and $\beta_{w2} = -0.0001[-0.003:0.003]$, Equation 4 bias $\beta_{w1} = -0.0009[-0.027:0.025]$ and $\beta_{w2} = -0.0001[-0.012:0.013]$) (Figure 5) and the 95% confidence interval coverage reached their expected value, that is, 0.95. However, it should be noted that the performance of Equation 4 depends on
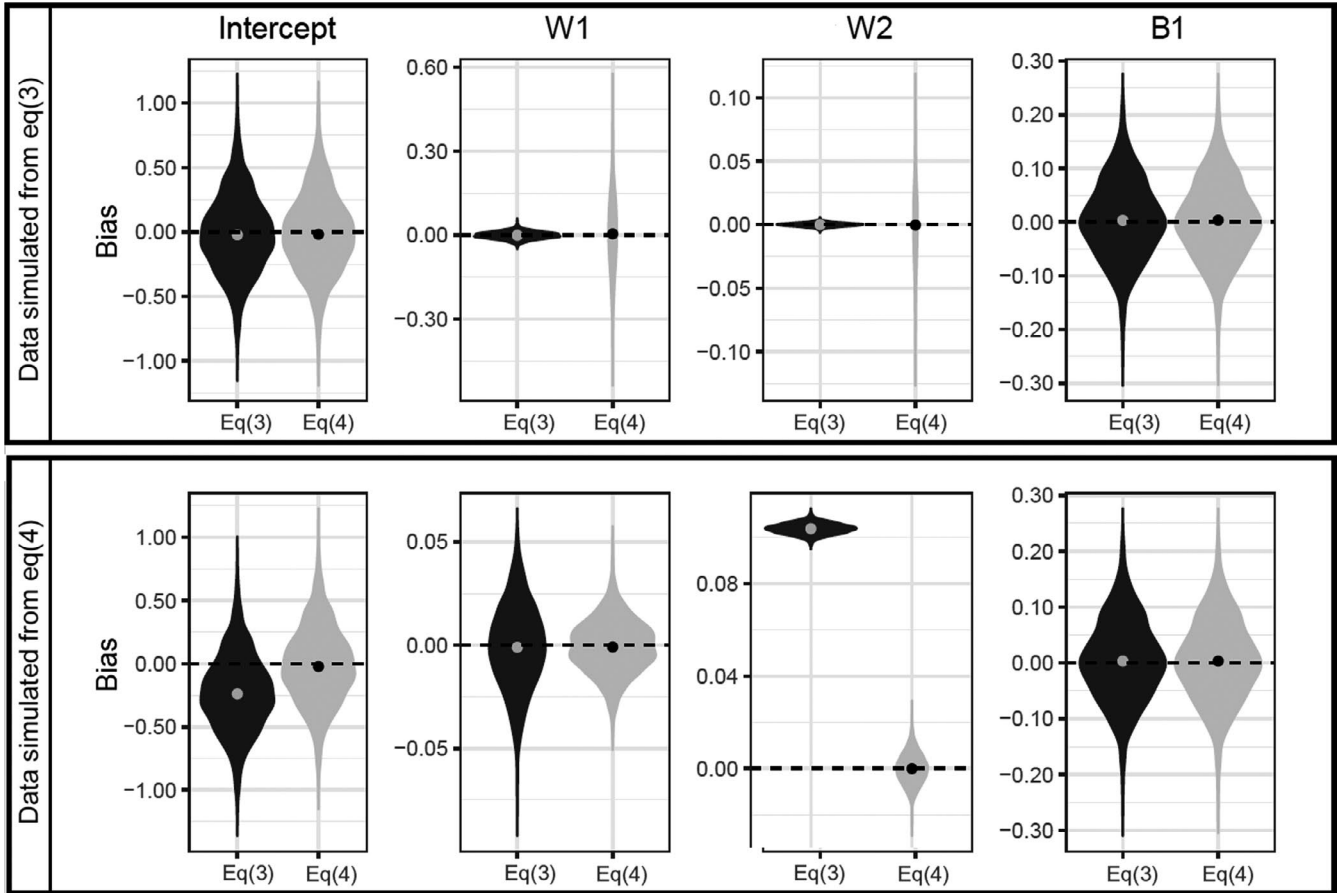
FIGURE 5 Bias in the estimates of the within- and among-individual effects from 1,000 datasets simulated and analysed by the two alternative quadratic equations of the within-centring method (data analysed with Equation 3 in black and Equation 4 in grey). Parameters include the intercept (beta0), the linear (W1) and the quadratic (W2) within-individual effect of *X*, and the linear among-individual effect (B1)
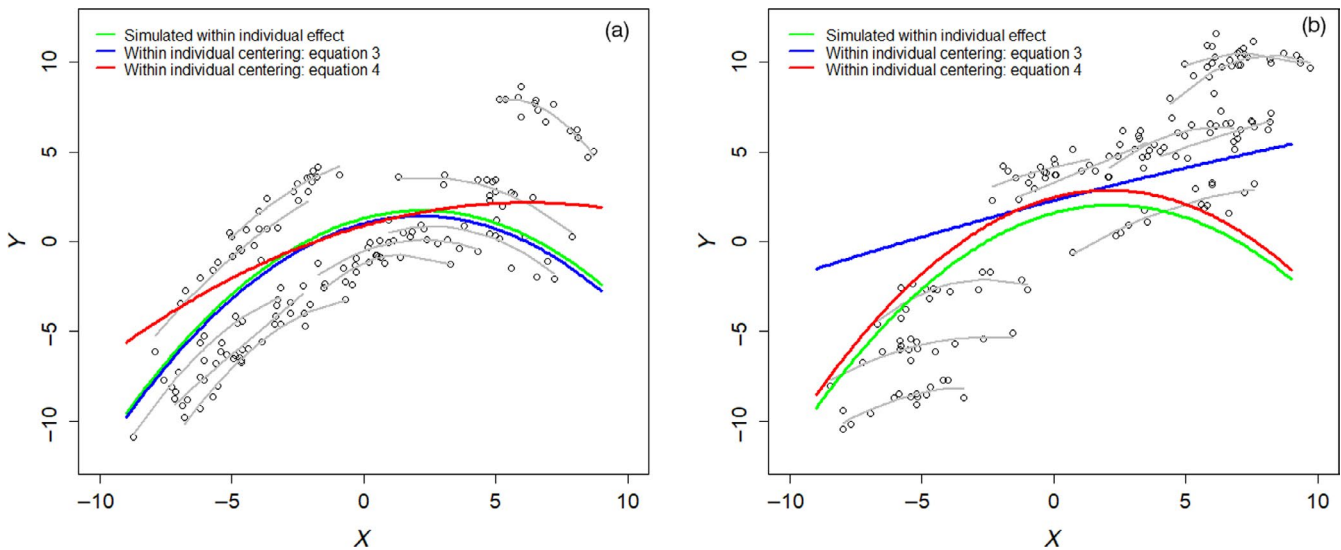

FIGURE 6 Estimation of the quadratic within-individual effect for data simulated with Equation 3 (data squared and then centred with square of individual mean; panel a) and 4 (data individually centred then squared; panel; b). Each dot represents one observation per individual and the grey lines depict observed within-individual trajectories. For illustration, we show the row observations for only 10 individuals. The green line shows the simulated within-individual effect and the blue and red lines show prediction of the within-individual effect estimated using the within-individual centring method according to Equations 3 and 4 respectively [Colour figure can be viewed at wileyonlinelibrary.com]

how accurately the average explanatory variable sampled for each individual corresponds to the true/biological average explanatory variable. When these quantities differ, for instance due to sampling bias, estimates of the quadratic within-individual pattern could be biased (Appendix S3).

Mismatch between the generating and analytical equations affects both bias and precision of the estimates. When data were simulated with Equation 3 but analysed with Equation 4, estimates were unbiased on average but very imprecise. In particular, the linear and quadratic within-individual estimates showed poor performance in this mismatching case (Figure 5; Figure S1) with a weak 95% confidence interval coverage (0.40 and 0.76 respectively). Thus, using Equation 4 to study a process generated by Equation 3 provide unbiased estimate in average but each single study might strongly mis-estimated the within-individual effect. When data were simulated with Equation 4 but analysed with Equation 3, the intercept was negatively biased and the quadratic within-individual effect was positively biased. Ninety-five per cent confidence interval coverage was 0.89, 0.85 and 0 for intercept, linear and quadratic within-individual effect respectively. Although the quality of the estimates is affected differently depending on type of mismatch, using a mismatched equation may generally lead to strong mis-estimation of the within-individual pattern (Figure 6). Nevertheless, the among-individual effect was estimated without bias and with a consistent precision and coverage regardless of the combination of generating and analytical equations used (Figure 5).

6 | THE SPECIFIC CASE OF AGE TRAJECTORIES WITH SELECTIVE DISAPPEARANCE: WHICH EQUATION SHOULD BE USED?

The importance in distinguishing within-individual from among-individual effects in age-related pattern has been recognized for a long time (Forslund & Pärt, 1995; Nussey et al., 2008; Reid et al., 2003). Using the within-individual centring method for age-related pattern, the age effect is split into two terms: the average individual age whose associated slope estimates the among-individual effect (e.g. the effect due to selective disappearance of certain phenotypes), and the centred age (i.e. the difference between the age at time of measurement and the average age over all an individual's measurements), whose slope estimates the within-individual effect. We found that 27% of the studies using the within-individual centring method to investigate age-related patterns included a quadratic within-individual effect to account for the decrease in performance in late life due to senescence (Figure 1). Among them, all studies describing the method with enough detail used Equation 4, but none used Equation 3. Although Equation 4 has been applied systematically, this equation might not always be the most appropriate one to use especially when the focus is on the average within-individual age patterns.

When we compare the individual-trajectories predicted by the two equations, it appears that they differ substantially (Figure 7). Equation 3 assumes that the within-individual effect depends on
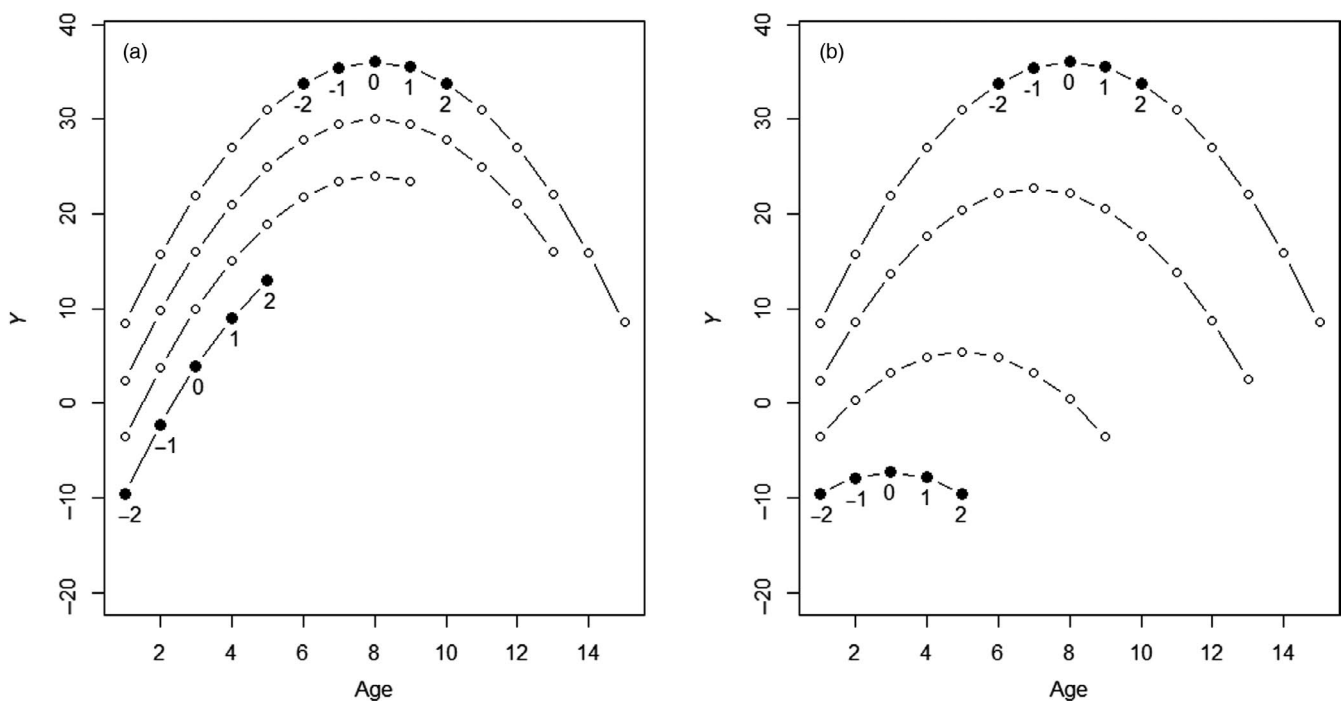
FIGURE 7 Age-related individual trajectories simulated according to Equation 3 (a) and Equation 4 (b). The five individual life histories are similar between the graphs in terms of the intercept and age range. Equation 4 assumes that the age pattern of an individual dying at 5 years old is similar to the age pattern between 6 and 10 years old of an individual dying at 15 years old (black dot with associated age-related changes)

the absolute age. All the individuals follow the same general pattern, that is, individual trajectories are parallel, among-individual differences being limited to variation in average performance (individual random intercept). Consequently, Equation 3 assumes a fixed age of onset and rate of senescence estimating a general senescence pattern (Figure 7a). In contrast, Equation 4 assumes that the within-individual effect depends on the relative age (age deviation relative to the individual mean). For instance, Equation 4 assumes that the age trajectory of individuals dying at 5 years old ($\overline{x}_j = \overline{age}_j = 3$) is similar to the age pattern between 6 and 10 years old of an individual dying at 15 years old ($\overline{x}_j = \overline{age}_j = 8$), because in both cases the age-related changes are the same, ranging from −2 to 2 (Figure 7b). Consequently, Equation 4 assumes a fixed rate but a variable age at onset of senescence that is proportional to longevity.

Although both equations allow modelling some individual variation in the age trajectories, neither of them allows a realistic description of among-individual variation in ageing pattern. Equations 3 and 4 represent two extreme biological processes differing by whether age is treated absolutely (chronological age) or relatively (biological age). In nature, senescence patterns are probably a mix of both chronological and biological ageing, although chronological ageing is likely dominating (Hammers et al., 2012; Hayward et al., 2015). Most importantly, the proper investigation of individual variation in individual ageing trajectories requires more flexible models including random slopes, although enough data should be available to correctly estimate them (see Dingemanse & Dochtermann, 2013; Nussey et al., 2007; Westneat et al., 2020).

Most of studies using the within-individual centring for age-related variable focused on the average within-individual change, not on the individual variation in the individual trajectories (e.g. Evans et al., 2011; Reichert et al., 2020; Vedder et al., 2014; Verhulst et al., 2014). Average within-individual change predicts the performances of a known-age individual taken randomly from the population. In that context, Equation 3 should be used because the research objective is framed in terms of absolute age. Furthermore, age-related patterns are generally dominated by the effect of absolute age, instead of relative age, suggesting that Equation 3 is likely more suitable for the investigation of age trajectories (Hammers et al., 2012; Hayward et al., 2015). However, Equation 4 considering a relative age process might still be suitable in specific cases and we stress that the choice of the equation used depend on the exact research question and the expected data generating process (see the recommendations in Section 8).

In the context of age patterns, the among-individual effect may be generated by both selective appearance and disappearance effects (van de Pol & Verhulst, 2006). Furthermore, the stochastic nature of mortality weakens the relationship between age of death and the average individual performance. These processes make the estimation of the within- and the among-individual effects more complex than what we simulated previously. Thus, we run additional simulations to investigate the performance of

Equations 3 and 4 in estimating correctly the within and among individual patterns in the particular case of age, when these processes are acting together.

7 | SIMULATION OF QUADRATIC AGE TRAJECTORIES WITH SELECTIVE DISAPPEARANCE

Here we assess the robustness of Equations 3 and 4 in distinguishing within- from among-individual effect for age-related pattern in the presence of selective disappearance, the most frequent process generating among individual patterns in age-related traits (Nussey et al., 2008; van de Pol & Verhulst, 2006). We generated age-related data simulating selective disappearance instead of using Equations 3 or 4 as in Section 5. None of the equations thus perfectly match the generating process. Given that a strong predominance of the effect of absolute age versus relative age effect was reported in previous studies (Hammers et al., 2012; Hayward et al., 2015) we limited our simulations to an absolute age process with selective disappearance.

7.1 | Methods

We simulated individual age-related trajectories using the following equation:

$$y_{ij} = \beta_0 + \beta_{w1}age_{ij} + \beta_{w2}age_{ij}^2 + u_j + e_{0ij},$$

Where y_{ij} is the response of individual i at occasion j, $\beta_0 = 0$ is the intercept, $\beta_{w1} = 0.8$ and $\beta_{w2} = -0.06$ are the linear and quadratic within-individual effects respectively, u_j is deviation from the intercept for individual j assumed to be drawn from a normal distribution with zero mean and among-subject variance $\sigma_u^2 = 2$, and e_{ij} is the residual error assumed to be drawn from a normal distribution with zero mean and variance $\sigma_e^2 = 1$. We assumed that the within-individual age effect is generated by an absolute process since the response variable (Y) depends directly of age. To simulate an among-individual effect, we generated the selective disappearance of frail individuals simulating age of death as a function of the individual-specific deviation in the response variable (u_j). This differential survival among individuals in relation to their age-independent performance corresponds to what has been reported by empirical studies as individual quality (McCleery et al., 2008; Pigeon et al., 2017; Vedder & Bouwhuis, 2018). We simulated individual survival using a Bernoulli distribution with an annual individual survival probability of ilogit($u_j/2 + 1.5$). Individuals with negative deviation, that is those with lower performances, also have lower survival prospects (see Appendix S4 for the code used). In contrast with data simulated with Equations 3 and 4, here the longevity is correlated with performance only in a probabilistic framework. An individual with low performance, and thus poor survival prospect, may still by chance reach a high longevity due to stochastic realization of survival. Using the procedure described above, we simulated 1,000

datasets with perfect detection, each consisting of 50 individuals. We analysed these datasets assessing the relationships between the response variable and age with both Equations 3 and 4 and computed bias and MSE as explained in Section 7.1. Finally, because detection is often imperfect, we performed the same simulation analysis but with individual recapture probability set to 0.5 and to 0.2 instead of 1.

7.2 | Results

Results showed that Equation 3, but not Equation 4, is suitable for the investigation of quadratic age patterns when the within-individual effect is generated by the absolute age and that the among-individual effect is generated by selective disappearance. Only Equation 3 provided unbiased estimates of both linear and quadratic effects (for recapture rate = 1, bias β_{w1} = −0.818[−0.860:−0.763] and β_{w2} = −0.0001[−0.006:0.006]) (Figure S2). When using Equation 4, the linear within-individual age effect was strongly bias but the quadratic effect was not (for recapture rate = 1, bias β_{w1} = −0.818[−0.860:−0.763] and bias β_{w2} = −0.0002[−0.008:0.008]; Figure S2). Using Equation 4 instead of Equation 3 for describing average individual age-related trajectory led to flawed inferences with a strong overestimation of senescence (earlier onset and higher rate, Figure S3). Since real data often show incomplete histories of individual trajectories, that is, detection probability is typically lower than 1, we reassessed the performance of these models considering incomplete individual capture histories. Results show that missing information within the lifetime of an individual does not change the respective performance of the models previously considered, but increases uncertainty for all parameters. Even when the detection probability is very low, that is, 0.2, leading to very patchy capture histories for each individual, the within-individual centring estimates from Equation 3 are accurate whereas those from Equation 4 are consistently biased (Figure S2 and S4).

Given that the simulations are based on an absolute quadratic age effect, it is not surprising that Equation 3 performed better than Equation 4. However, previous simulations (Section 5) indicated that Equation 4 provided unbiased but imprecise estimates when fitted on data generated by Equation 3. Here, Equation 4 provided biased estimates for the linear age effect. Even if these data have not been generated from Equation 3, they are satisfactorily modelled by Equation 3 indicating that this equation is robust to the presence of selective disappearance and to the stochasticity generated by the survival process.

8 | CONCLUSION AND RECOMMENDATIONS

The within-individual centring method should be used with caution for nonlinear relationships. We showed that two different equations allow estimating within-individual quadratic relationships and that each equation makes different assumptions about the shape of the individual pattern. The choice of using one equation instead of the other should depend upon the biological process investigated. We provide the following guidelines:

Step 1: Determine if the research interest is on the among-individual variation in the within-individual effect (e.g. individual variation in reaction norm) or on the average within-individual effect (e.g. the average reaction norm). In the first case, more sophisticate models including random slopes for the within-individual effect are required (see Dingemanse & Dochtermann, 2013; Nussey et al., 2007; Westneat et al., 2020). In the second case, go to the next step.

Step 2: Make an educated guess whether the within-individual effect is generated by an absolute or relative process (see examples in Section 4). In case of absolute process, use Equation 3. In case of relative process, use Equation 4. When the knowledge of the study system does not allow to address this issue, see step 3.

Step 3: When it is impossible to decide whether to use Equations 3 or 4, comparing the fit of both equations within a standard model selection framework would determine which equation fit the data best. However, as always in statistic, a fit of model to the data does not necessarily mean the analysis model is the generating process.

Step 4: If Equation 4 is used, the estimates from the model will be unbiased only if the average explanatory variable sampled for each individual match the average condition in the generating process (see Appendix S3). This could be difficult to assess in practice and thus estimates obtained using Equation 4 should be interpreted and discussed cautiously in full knowledge of the situation.

Step 5: When reporting results, be transparent about the decision to fit one equation instead of the other. Make explicit the assumptions and constraints involved, and interpret the results accordingly.

Note that in the case of age pattern, an alternatively method to disentangled within from among individual effect is the use of age at first and last reproduction as fixed covariates in addition to the age to model explicitly selective appearance and disappearance respectively (van de Pol & Verhulst, 2006). This method, which is essentially a reparameterization of Equation 3b (see Appendix S5), allows estimating average within-individual pattern distinguishing the selective appearance and disappearance. Furthermore, this method facilitates the implementation of threshold models (e.g. Fay et al., 2021), which have been claimed to be more suitable than polynomial models for the investigation of ageing patterns (Berman et al., 2009; Froy et al., 2017; Murgatroyd et al., 2018; Rodríguez-Muñoz et al., 2019).

Using the within-individual centring approach with polynomial terms requires large sample sizes at the individual level to be able to properly assess the linear and quadratic patterns both at the within- and among-individual levels. Even if the simulations showed no biases when matching equations were used, it should be noted that we used datasets with numerous repeated measures (15 observations per individual). The biases and reliability of within-individual centring with quadratic patterns have not yet been properly assessed and we urge researchers to consider their sample size and data structure before using this type of approach. In any case, when

using the within-individual centring for nonlinear relationships, the equation used should be chosen carefully in order to run meaningful analyses. In line with Westneat et al. (2020), we stress that it is critical to have a clear understanding of the equation used to ensure that the assumptions made by a statistical model (using within-individual centring or not) match the biological process investigated. We encourage the use of the within-individual centring method and promote its relevant application for nonlinear relationships.

CONFLICT OF INTEREST
The authors have no conflict of interest to declare.

AUTHORS' CONTRIBUTIONS
R.F., F.P. and J.M. discussed the ideas that lead to the study, R.F. and F.P. designed and ran the simulations, J.M. develop the mathematical formalization, R.F. led the writing of the manuscript with the help of F.P. and J.M.

DATA AVAILABILITY STATEMENT
Data sharing is not applicable to this article as no data were analysed in this study.

ORCID
Rémi Fay https://orcid.org/0000-0002-7202-367X
Julien Martin https://orcid.org/0000-0001-7726-6809
Floriane Plard https://orcid.org/0000-0002-8509-723X

REFERENCES
Aubry, L. M., Koons, D. N., Monnat, J.-Y., & Cam, E. (2009). Consequences of recruitment decisions and heterogeneity on age-specific breeding success in a long-lived seabird. *Ecology*, *90*, 2491–2502. https://doi.org/10.1890/08-1475.1

Berman, M., Gaillard, J.-M., & Weimerskirch, H. (2009). Contrasted patterns of age-specific reproduction in long-lived seabirds. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 375–382. https://doi.org/10.1098/rspb.2008.0925

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H. W., Schreiber, S. J., Urban, M. C., & Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, *26*, 183–192. https://doi.org/10.1016/j.tree.2011.01.009

Bouwhuis, S., Sheldon, B. C., Verhulst, S., & Charmantier, A. (2009). Great tits growing old: Selective disappearance and the partitioning of senescence to stages within the breeding cycle. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 2769–2777. https://doi.org/10.1098/rspb.2009.0457

Dammhahn, M., Landry-Cuerrier, M., Réale, D., Garant, D., & Humphries, M. M. (2017). Individual variation in energy-saving heterothermy affects survival and reproductive success. *Functional Ecology*, *31*, 866–875. https://doi.org/10.1111/1365-2435.12797

Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour: Mixed-effect modelling approaches. *Journal of Animal Ecology*, *82*, 39–54. https://doi.org/10.1111/1365-2656.12013

Evans, S. R., Gustafsson, L., & Sheldon, B. C. (2011). Divergent patterns of age-dependence in ornamental and reproductive traits in the collared flycatcher. *Evolution*, *65*, 1623–1636. https://doi.org/10.1111/j.1558-5646.2011.01253.x

Fay, R., Ravussin, P.-A., Arrigo, D., von Rönn, J. A. C., & Schaub, M. (2021). Age-specific reproduction in female pied flycatchers: Evidence for asynchronous aging. *Oecologia*, *196*, 723–734. https://doi.org/10.1007/s00442-021-04963-2

Forslund, P., & Pärt, T. (1995). Age and reproduction in birds—Hypotheses and tests. *Trends in Ecology & Evolution*, *10*, 374–378. https://doi.org/10.1016/S0169-5347(00)89141-7

Froy, H., Lewis, S., Nussey, D. H., Wood, A. G., & Phillips, R. A. (2017). Contrasting drivers of reproductive ageing in albatrosses. *The Journal of Animal Ecology*, *86*, 1022–1032. https://doi.org/10.1111/1365-2656.12712

Hämäläinen, A., Dammhahn, M., Aujard, F., Eberle, M., Hardy, I., Kappeler, P. M., Perret, M., Schliehe-Diecks, S., & Kraus, C. (2014). Senescence or selective disappearance? Age trajectories of body mass in wild and captive populations of a small-bodied primate. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1791), 20140830. https://doi.org/10.1098/rspb.2014.0830

Hamel, S., Gaillard, J.-M., Douhard, M., Festa-Bianchet, M., Pelletier, F., & Yoccoz, N. G. (2018). Quantifying individual heterogeneity and its influence on life-history trajectories: Different methods for different questions and contexts. *Oikos*, *127*, 687–704. https://doi.org/10.1111/oik.04725

Hammers, M., Richardson, D. S., Burke, T., & Komdeur, J. (2012). Age-dependent terminal declines in reproductive output in a wild bird. *PLoS ONE*, *7*, e40413. https://doi.org/10.1371/journal.pone.0040413

Hayward, A. D., Moorad, J., Regan, C. E., Berenos, C., Pilkington, J. G., Pemberton, J. M., & Nussey, D. H. (2015). Asynchrony of senescence among phenotypic traits in a wild mammal population. *Experimental Gerontology*, *71*, 56–68. https://doi.org/10.1016/j.exger.2015.08.003

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, *24*(623), 641.

Kendall, B. E., Fox, G. A., Fujiwara, M., & Nogeire, T. M. (2011). Demographic heterogeneity, cohort selection, and population growth. *Ecology*, *92*, 1985–1993. https://doi.org/10.1890/11-0079.1

Kreft, I. G., De Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, *30*, 1–21. https://doi.org/10.1207/s15327906mbr3001_1

McCleery, R. H., Perrins, C. M., Sheldon, B. C., & Charmantier, A. (2008). Age-specific reproduction in a long-lived species: The combined effects of senescence and individual quality. *Proceedings of the Royal Society of London B: Biological Sciences*, *275*, 963–970. https://doi.org/10.1098/rspb.2007.1418

Morrongiello, J. R., Sweetman, P. C., & Thresher, R. E. (2019). Fishing constrains phenotypic responses of marine fish to climate variability. *Journal of Animal Ecology*. https://doi.org/10.1111/1365-2656.12999

Murgatroyd, M., Roos, S., Evans, R., Sansom, A., Whitfield, D. P., Sexton, D., Reid, R., Grant, J., & Amar, A. (2018). Sex-specific patterns of reproductive senescence in a long-lived reintroduced raptor. *Journal of Animal Ecology*, *87*, 1587–1599. https://doi.org/10.1111/1365-2656.12880

Nussey, D. H., Coulson, T., Festa-Bianchet, M., & Gaillard, J.-M. (2008). Measuring senescence in wild animal populations: Towards a longitudinal approach. *Functional Ecology*, 22, 393–406. https://doi.org/10.1111/j.1365-2435.2008.01408.x

Nussey, D. H., Wilson, A. J., & Brommer, J. E. (2007). The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology*, 20, 831–844. https://doi.org/10.1111/j.1420-9101.2007.01300.x

Pick, J. L., Ebneter, C., Hutter, P., & Tschirren, B. (2016). Disentangling genetic and prenatal maternal effects on offspring size and survival. *The American Naturalist*, 188, 628–639. https://doi.org/10.1086/688918

Pigeon, G., Festa-Bianchet, M., & Pelletier, F. (2017). Long-term fitness consequences of early environment in a long-lived ungulate. *Proceedings of the Royal Society B: Biological Sciences*, 284(1853), 20170222. https://doi.org/10.1098/rspb.2017.0222

Reichert, S., Berger, V., Jackson, J., Chapman, S. N., Htut, W., Mar, K. U., & Lummaa, V. (2020). Maternal age at birth shapes offspring life-history trajectory across generations in long-lived Asian elephants. *Journal of Animal Ecology*, 89, 996–1007. https://doi.org/10.1111/1365-2656.13049

Reid, J. M., Bignal, E. M., Bignal, S., McCracken, D. I., Bogdanova, M. I., & Monaghan, P. (2010). Parent age, lifespan and offspring survival: Structured variation in life history in a wild population. *Journal of Animal Ecology*, 79, 851–862. https://doi.org/10.1111/j.1365-2656.2010.01669.x

Reid, J. M., Bignal, E. M., Bignal, S., McCracken, D. I., & Monaghan, P. (2003). Age-specific reproductive performance in red-billed choughs *Pyrrhocorax pyrrhocorax*: Patterns and processes in a natural population. *Journal of Animal Ecology*, 72, 765–776. https://doi.org/10.1046/j.1365-2656.2003.00750.x

Rodríguez-Muñoz, R., Boonekamp, J. J., Liu, X. P., Skicko, I., Haugland Pedersen, S., Fisher, D. N., Hopwood, P., & Tregenza, T. (2019). Comparing individual and population measures of senescence across 10 years in a wild insect population. *Evolution*, 73, 293–302. https://doi.org/10.1111/evo.13674

Siracusa, E. R., Wilson, D. R., Studd, E. K., Boutin, S., Humphries, M. M., Dantzer, B., Lane, J. E., & McAdam, A. G. (2019). North American red squirrels mitigate costs of territory defence through social plasticity. *Animal Behaviour*, 151, 29–42. https://doi.org/10.1016/j.anbehav.2019.02.014

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.

Stover, J. P., Kendall, B. E., & Fox, G. A. (2012). Demographic heterogeneity impacts density-dependent population dynamics. *Theoretical Ecology*, 5, 297–309. https://doi.org/10.1007/s12080-011-0129-x

van de Pol, M., & Verhulst, S. (2006). Age-dependent traits: A new statistical model to separate within-and between-individual effects. *The American Naturalist*, 167, 766–773. https://doi.org/10.1086/503331

Van de Pol, M., & Wright, J. (2009). A simple method for distinguishing within-versus between-subject effects using mixed models. *Animal Behaviour*, 77, 753. https://doi.org/10.1016/j.anbehav.2008.11.006

van Noordwijk, A. J., & de Jong, G. (1986). Acquisition and allocation of resources: Their influence on variation in life history tactics. *The American Naturalist*, 128, 137–142. https://doi.org/10.1086/284547

Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 439–454. https://doi.org/10.2307/2061224

Vaupel, J. W., & Yashin, A. I. (1985). Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *The American Statistician*, 39, 176–185.

Vedder, O., & Bouwhuis, S. (2018). Heterogeneity in individual quality in birds: Overall patterns and insights from a study on common terns. *Oikos*, 127, 719–727. https://doi.org/10.1111/oik.04273

Vedder, O., Bouwhuis, S., & Sheldon, B. C. (2014). The contribution of an avian top predator to selection in prey species. *Journal of Animal Ecology*, 83, 99–106. https://doi.org/10.1111/1365-2656.12114

Verhulst, S., Geerdink, M., Salomons, H. M., & Boonekamp, J. J. (2014). Social life histories: Jackdaw dominance increases with age, terminally declines and shortens lifespan. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 20141045.

Vindenes, Y., & Langangen, Ø. (2015). Individual heterogeneity in life histories and eco-evolutionary dynamics. *Ecology Letters*, 18, 417–432. https://doi.org/10.1111/ele.12421

Westneat, D. F., Araya-Ajoy, Y. G., Allegue, H., Class, B., Dingemanse, N., Dochtermann, N. A., Garamszegi, L. Z., Martin, J. G. A., Nakagawa, S., Réale, D., & Schielzeth, H. (2020). Collision between biological process and statistical analysis revealed by mean centring. *Journal of Animal Ecology*, 89, 2813–2824. https://doi.org/10.1111/1365-2656.13360

Wilson, A. J., & Nussey, D. H. (2010). What is individual quality? An evolutionary perspective. *Trends in Ecology & Evolution*, 25(4), 207–214.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.