RESEARCH METHODS GUIDE

Conceptualizing and quantifying body condition using structural equation modelling: A user guide

Magali Frauendorf1,2 | Andrew M. Allen1,2,3 | Simon Verhulst4 | Eelke Jongejans1,2,3 | Bruno J. Ens2,5 | Henk-Jan van der Kolk1,2 | Hans de Kroon2,3 | Jeroen Nienhuis5 | Martijn van de Pol1,2

1Department of Animal Ecology, Netherlands Institute of Ecology, Wageningen, The Netherlands
2Centre for Avian Population Studies, Wageningen, The Netherlands
3Department of Animal Ecology and Physiology & Experimental Plant Ecology, Radboud University, Nijmegen, The Netherlands
4Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands
5Sovon—Dutch Centre for Field Ornithology, Nijmegen, The Netherlands

Correspondence
Magali Frauendorf
Email: magalifrauendorf@gmail.com

Funding information
Applied and Engineering Sciences domain of the Netherlands Organisation for Scientific Research (NWO-STW), Grant/Award Number: 14638; NAM gas exploration; Birdlife Netherlands; Royal Netherlands Air Force; Deltas

Handling Editor: Jean-Michel Gaillard

Abstract
1. Body condition is an important concept in behaviour, evolution and conservation, commonly used as a proxy of an individual's performance, for example in the assessment of environmental impacts. Although body condition potentially encompasses a wide range of health state dimensions (nutritional, immune or hormonal status), in practice most studies operationalize body condition using a single (univariate) measure, such as fat storage. One reason for excluding additional axes of variation may be that multivariate descriptors of body condition impose statistical and analytical challenges.

2. Structural equation modelling (SEM) is used in many fields to study questions relating multidimensional concepts, and we here explain how SEM is a useful analytical tool to describe the multivariate nature of body condition. In this 'Research Methods Guide' paper, we show how SEM can be used to resolve different challenges in analysing the multivariate nature of body condition, such as (a) variable reduction and conceptualization, (b) specifying the relationship of condition to performance metrics, (c) comparing competing causal hypothesis and (d) including many pathways in a single model to avoid stepwise modelling approaches. We illustrated the use of SEM on a real-world case study and provided R-code of worked examples as a learning tool.

3. We compared the predictive power of SEM with conventional statistical approaches that integrate multiple variables into one condition variable: multiple regression and principal component analyses. We show that model performance on our dataset is higher when using SEM and led to more accurate and precise estimates compared to conventional approaches.

4. We encourage researchers to consider SEM as a flexible framework to describe the multivariate nature of body condition and thus understand how it affects biological processes, thereby improving the value of body condition proxies for predicting organismal performance. Finally, we highlight that it can be useful for...
other multidimensional ecological concepts as well, such as immunocompetence, oxidative stress and environmental conditions.

**Keywords**

body condition index, composite variable, fitness component, latent variable, multiple regression, multiple-indicator multiple-cause model, path analysis, principal component analysis

---

**1 INTRODUCTION**

Body condition is (by definition) a key determinant of an individual’s fitness (Wilder et al., 2016). Understanding how best to quantify body condition therefore has implications in behavioural (Kirk & Gosler, 1994), evolutionary (Aubry et al., 2013) and conservation (Stirling & Derocher, 2012) studies, but how to analyse and estimate body condition has remained a subject of debate for decades (Wilder et al., 2016). Notably, there is currently no consensus on a strict definition of body condition, although there is conceptual agreement that the term describes the degree to which an animal’s health state influences its performance (e.g. reproduction; Brown, 1996; Peig & Green, 2009). The concept of body condition can thus encompass a wide range of morphological and physiological metrics that describe the nutritional, immune and/or hormonal state of an individual. Which of such metrics are most relevant will depend on the study system and aims, and therefore also on the performance variable of interest (e.g. survival, reproduction, activity; Table 1).

Despite the multidimensional nature of body condition, most ecological studies operationalize body condition by quantifying a single (univariate) measurable variable (Wilder et al., 2016). In studies in which the energy stores are thought to be important for the individuals performance, often it is approximated by the mass of an individual relative to its size, as this is assumed to reflect internal fat content (Durell et al., 2001; Janssen et al., 2011; Schulte-Hostedde et al., 2005; see van der Meer and Piersma (1994) for the distinction between energy reserves and energy stores). Well-known examples include the body mass index in humans (BMI, mass/length²; Nuttall, 2015), or the residuals of body mass regressed to one or more variables of structural size (e.g. tarsus or head length; Jakob et al., 1996), which are widely used in avian studies (Peig & Green, 2009). However, two individuals with similar BMI may still differ in various other aspects of their health, quality or vigour (Stevenson & Woods, 2006). The BMI in humans, for instance, reflects different body composition in men and women (e.g. women have more fat and men more muscles; Jackson et al., 2002) and therefore appears to be a poor predictor of morbidity and mortality, as people with similar BMI can have different metabolic health (Roberson et al., 2014). In fact, it has now become clear that different metrics of body condition are also related to aspects other than the energetic state of an individual (Table 1). Such non-energetic aspects can be an additional or even more appropriate predictor of fitness (Cox & Calsbeek, 2015; McGuire et al., 2018; with fitness measuring the genetic contribution to future generations). For example, an animal can have high lipid content but have low reproductive success if it is deficient in other nutrients (e.g. protein, vitamins, minerals) that it needs to reproduce (Nie et al., 2014).

These examples indicate that operationalizing body condition via univariate descriptors of body energy stores is unsatisfactory for two reasons. First, the underlying assumption that there is a (monotonic) positive association between energy stores and organismal performance has often been falsified. This suggests that it is important to also consider other measures of the health state beyond energy stores, such as physiological measures (Table 1; note that here our point is to emphasize that a variety of body condition measures can be and are used in ecological and (veterinary) health research and that no ‘perfect’ metric exists, i.e. universally suitable for all studies, although we appreciate that in specific studies some condition metrics can be more relevant than other metrics for biological reasons). Second, a univariate approach to body condition is unlikely to fully capture what is essentially a multivariate concept, at the expense of lower power to predict variation in performance. Indeed, efforts to combine several variables into one compositional measure have been able to better predict fitness variation than each of these condition variables separately (Sousa et al., 2007; Verhulst et al., 2004). A solution to these two problems may be to not operationalize condition as one or two major axes of variation among traits that are a priori assumed to be related to fitness (e.g. univariate metrics such as fat content), but rather operationalize condition as the axis of variation (i.e. a specific composite body condition metric) that best explains variance in individual fitness or performance. Such a view is quite similar to how Wilson and Nussey (2010) conceptualized ‘individual quality’, though body condition is always linked to a specific aspect of individual quality (i.e. health state) and is a much more dynamic concept as it fluctuates during lifetime while individual quality is a less flexible characteristic (e.g. social dominance).

Adopting a multivariate description of body condition and identifying which condition index best explains variation in performance or fitness impose statistical and analytical challenges that have thus far not been fully addressed. Studies that aim to describe the multivariate nature of body condition have used a variety of statistical methods (e.g. multiple regression, principal component analyses), which have limited flexibility and implicitly make specific assumptions that are not always recognized. Principal component analyses (PCA) identify axes of body condition based on the correlation patterns among condition variables (Abdi & Williams, 2010) while...
**TABLE 1** An overview highlighting the diversity of studies investigating body condition with different aspects of an individual’s health state and performance measures. Note that this overview is not based on an exhaustive search, as it only aims to illustrate the diversity in health status metrics considered for different performance metrics.

| Performance \ Health status | Nutritional state | Immune state | Hormonal state |
|-----------------------------|-------------------|--------------|----------------|
| Reproduction                | Reproductive output (eggs, fledglings) | Abdominal profile index<sup>1</sup> | Principal component of packed cell volume, haemoglobin, scaled mass, muscle score, fat score<sup>2</sup> | Corticosterone<sup>26, 27, 28, 29, 30</sup> |
| Milk production             | Body condition scoring<sup>3, 4, 5</sup> | Abdominal profile index<sup>1</sup> | | |
| Lay date                    | Residual mass<sup>6, 7</sup> | | | |
| Paternal investment (sexual selection, male attraction) | Fulton index (weight/length)<sup>7</sup> | Lymphocyte<sup>6</sup> | | |
| Survival                    | Subcutaneous and abdominal fat<sup>8</sup> | Principal component of haematocrit, buffy coat, residual mass<sup>11</sup> | | |
| Survival (adult, juveniles) | Principal component of haematocrit, buffy coat, residual mass<sup>11</sup> | Parrot load<sup>12</sup> | Testosterone<sup>33</sup> |
| Growth                      | Ratio (mass/length)<sup>13</sup> | | Corticosterone<sup>24, 34, 35, 36</sup> |
| Adult growth                | Muscle index<sup>15</sup> | | Testosterone<sup>37</sup> |
| Foraging                    | Visible fat scores<sup>17</sup> | | Corticosterone<sup>35</sup> |
| Foraging efficiency/strategy| Ratio (mass/length)<sup>17</sup> | | | |
| Foraging condition          | Residual mass<sup>39</sup> | | Corticosterone<sup>38, 39, 40, 41, 42</sup> |
| Subcutaneous fat<sup>18</sup> | Plasma parameters (e.g. protein, urea, cholesterol, carbohydrates)<sup>20</sup> | Nutrients (e.g. calcium, phosphorus, nitrogen)<sup>21</sup> | | |
| Migration                   | Subcutaneous or abdominal fat scoring<sup>1, 17, 22</sup> | Scaled mass index<sup>23</sup> | | |
| Response to environment     | Response to ecotoxicologically polluted area | | | |

Notes: <sup>1</sup>Béty et al. (2003), <sup>2</sup>Berry and Wilder (2013), <sup>3</sup>Russell et al. (1969), <sup>4</sup>Edmonson et al. (1989), <sup>5</sup>Roche et al. (2009), <sup>6</sup>Gleeson et al. (2005), <sup>7</sup>Neff and Cargnelli (2004), <sup>8</sup>Geman et al. (2006), <sup>9</sup>Nuttall (2015), <sup>10</sup>MacCracken and Stebbings (2012), <sup>11</sup>Verhulst et al. (2004), <sup>12</sup>Souchay et al. (2013), <sup>13</sup>Beintema (1994), <sup>14</sup>Williams et al. (2008), <sup>15</sup>Brown and Murphy (2004), <sup>16</sup>Metón et al. (1999), <sup>17</sup>Swanson et al. (1999), <sup>18</sup>Bearhop et al. (2004), <sup>19</sup>Durell et al. (2001), <sup>20</sup>Alonso-Alvarez et al. (2002), <sup>21</sup>Nie et al. (2014), <sup>22</sup>Wiersma and Piersma (1995), <sup>23</sup>Tête et al. (2013), <sup>24</sup>Milenkaya et al. (2015), <sup>25</sup>Simons et al. (2012), <sup>26</sup>Angelier et al. (2010), <sup>27</sup>Ouyang et al. (2011), <sup>28</sup>Kouwenberg et al. (2013), <sup>29</sup>Perez et al. (2016), <sup>30</sup>Montreuil-Spencer et al. (2019), <sup>31</sup>Schoech et al. (2004), <sup>32</sup>Saino and Møller (1995), <sup>33</sup>Dufty (1989), <sup>34</sup>Rivers et al. (2012), <sup>35</sup>Kitaysky et al. (2003), <sup>36</sup>Sheriff et al. (2009), <sup>37</sup>Sheriff et al. (2009), <sup>38</sup>von Engelhardt et al. (2006), <sup>39</sup>Jimeno, Briga, et al. (2017), <sup>40</sup>Jimeno et al. (2018), <sup>41</sup>Jimeno et al. (2017), <sup>42</sup>Marra and Holberton (1998), <sup>43</sup>Kitaysky et al. (1999).
the weight of a variable in an estimate of condition instead ideally reflects how much it contributes to explaining variation in performance relative to other variables. Multiple regression (MR) models have low flexibility in that they cannot handle variables that are simultaneously response and explanatory variables, which is resolved using stepwise modelling approaches that can potentially cause bias in the parameter estimates (Darlington & Smulders, 2001; Freckleton, 2002). Thus, there is a need for an alternative analytical framework to integrate multiple variables in one condition estimate, with their relative weight depending on their power in explaining variation in fitness, or, more usually, fitness proxies such as survival or reproductive traits (e.g. lay date, egg volume, fledgling number).

To this end, we introduce structural equation modelling (SEM) and show how it can be used to overcome the above-mentioned challenges in quantifying a body condition index. SEM is a flexible conceptual framework, where statistical and mathematical tools, along with analytical principles are used to learn about a system (Grace, 2006; Shipley, 2016). SEM builds on path analysis, which is a statistical regression analysis that allows one to investigate patterns of effect within a system of variables by examining the impact of a set of predictor variables on multiple, possibly interrelated, response variables. Moreover, SEM extends path analysis by including unmeasured variables. For example, ‘latent variables’ or ‘composite variables’ can be used to summarize multivariate concepts, like body condition, that are not directly measurable themselves. SEM can be used to evaluate competing multivariate models of body condition, and thus learn more about the study system, and quantify the relative importance of different pathways. Although SEM originates from the social sciences (Grace et al., 2010), the capacity for evaluating multivariate hypotheses has also attracted the interest of ecologists, with some aspects of SEM (e.g. latent variables, path analysis) being widely used in specific subfields of ecology (Grace, 2006; Shipley, 2016). Belovsky et al. (2004) suggested that the ecological sciences can be advanced by a better integration of theory and empirical evidence and by studying multiple causes simultaneously, both of which can be addressed with SEM.

We emphasize that the SEM techniques we describe are not novel from a statistical perspective. Instead, our aim is to demonstrate in a comprehensive and educational way its usefulness for the quantification of body condition indices as well as for testing biological hypotheses concerning condition through worked empirical examples on a shorebird species. Data are from a large-scale study in which we measured a variety of morphological and physiological variables on Eurasian oystercatchers *Haematopus ostralegus*. We provide R-code of all analyses with annotations to facilitate its usage and as a learning tool for those new to SEM.

2 | INTRODUCING THE CASE STUDY USED TO EXPLAIN THE CONCEPTUAL FRAMEWORK

2.1 | The case study system

The global Eurasian oystercatcher population has declined rapidly which can be caused either by reduced reproduction or by reduced survival, and thus our aim is to identify the impact of potential environmental drivers in winter affecting survival. However, survival probabilities of individuals are not directly observable, as we have to infer them from the stochastic realizations of individuals dying or not, leading to information loss (i.e. even more problematic when there is imperfect detection and we have to rely on capture–mark–recapture analysis for inference, which also requires many years of data collection). An additional challenge in our long-lived study species is that mortality events are rare, with an average annual adult survival rate of ~0.9 (Allen et al., 2019), meaning that there is little variation in this performance metric in a single year. Both challenges limit the opportunity to detect associations between environmental drivers and survival probabilities. Creating a body condition index that is closely related to survival gives the opportunity to use a more informative fitness proxy in the future based on data that are less
time-consuming to collect in the field (compared to the long-term survival data).

Therefore, the overall aim of the case study is to use health status measurements on captured birds to derive an integrative body condition index that maximizes the explained variation in a fitness component (survival) and helps to identify how environmental drivers (competitor density) affect body condition (Figure 1).

To quantify body condition, 774 oystercatchers were caught with nets at high-tide roosts and colour-banded in the Netherlands during 19 catching events spanning two winters (2016–2017: 196 birds; 2017–2018: 578 birds; Figure S1; Table S1). Several morphological and physiological variables were measured that could potentially reflect the immune, hormonal and nutritional state of an individual. In addition, we measured (carotenoid based) bill colour that may serve as a visual status signal of an individual’s body condition to conspecifics (Figure 1). We considered these variables based on previous studies that have shown that they can be associated with either body mass or survival (Text S1). Some variables that we are not primarily interested in may influence multiple condition variables (Norte et al., 2009), and thereby could explain some of the intercorrelation between various condition variables. Accounting for such confounding variables may clarify to what extent these condition variables are different axes of body condition (Figure 1). By considering effects of handling time (proportion of time between capture and measuring) on condition variables, we aim to statistically account for time-dependent mass loss and possible effects on other physiological variables that we are not interested in. Similarly, we want to consider confounding effects of individual characteristics such as age, sex and bill shape. The age of the birds was classified in 1st, 2nd, 3rd calendar year and adults (>3rd calendar year). We focused only on sub-adults (2nd and 3rd calendar year) and adults (>3rd calendar year) in the analysis because the number of sampled juveniles (1st calendar year) was small (n = 28; Table S1). Adult oystercatchers cannot be accurately aged and therefore we do not consider any possible effect of senescence on body condition or survival. A long-term study population in the Netherlands, in which there are many known-age individuals (e.g. marked at birth), also did not reveal any evidence that survival or body condition varies systematically with age among adults. Bill tip height was used as a proxy for the type of individual’s feeding specialization ranging from worm specialists (pointed bill characterized by a low bill tip height) to shellfish specialists (blunt bill characterized by a high bill tip height; van de Pol et al., 2009).

We relate the body condition index to survival in the year following capture. The survival estimates (mean [SD], min, max: 0.87 [0.06], 0.73, 0.92) used in the case study were generated from a multi-state live and dead recoveries model and therefore are on an area level. See Text S1 and Table S2 for details on data collection and capture-mark-recapture modelling.

### 2.2 | The conceptual framework applied to the case study

In the following three sections, we describe step by step how we quantify body condition with SEM in the oystercatcher population. We first describe how to summarize the multidimensional nature of body condition using a variable reduction approach (box 1 in Figure 1). We then describe how to quantify body condition based on a fitness trait (box 2 in Figure 1). The last section combines the first two steps but also adds confounding variables to the model (box 3 in Figure 1) to illustrate the high flexibility of SEM. For comparison, we start all sections by analysing the data with a conventional statistical approach before demonstrating the SEM approach. In the last section, we compare the model performance between SEM and a conventional approach.

### 3 | SECTION 1: SUMMARIZING THE MULTIDIMENSIONAL NATURE OF BODY CONDITION: VARIABLE REDUCTION

In this section, we will illustrate how the multivariate concept of body condition can be summarized and compared to alternative models by considering two variables of health status, namely ‘energy stores’ and ‘bill colour’ (box 1 in Figure 1).

### 3.1 | Conventional method: Principal component analysis (PCA)

A conventional method that combines several measures into one compositional measure is PCA, a multivariate statistical technique that is used in most scientific disciplines (Abdi & Williams, 2010). PCA has also been used to create condition indices (e.g. Bearhop et al., 2004; Milenkaya et al., 2015; Verhulst et al., 2004). PCA summarizes multiple variables using the shared correlation structure to derive the principal components. If the first principal component (PC1) explains the majority of the variation, the other principal components can be ignored without much information loss, and variable reduction is achieved through using PC1 as a single new condition variable. The use of PCA is suitable for the (adult) bill colour and energy stores examples because the different (adult) bill colour and energy stores metrics are strongly linearly interrelated (Figure S2; Table S3), and such intercorrelation is a requirement for PCA.

Concerning the energy stores, we considered two alternative variables in our analysis that reflect some of the diversity of measures used in the ecological literature (Peig & Green, 2009): (a) the residuals of mass regressed against structural size measures: head length, tarsus length and wing length and (b) the ratio of mass to size (tarsus length). Both variables are assumed to be positively related to fitness (Stevenson & Woods, 2006) and are highly intercorrelated...
Luminance (Table S3; Figure S6). The PC1 of bill colour explained 88.8% and the intercorrelation of the variables between adults and subadults (components of a colour; Quesada & Senar, 2006) for traits into a single variable.

For bill colour, we ran a PCA with the variables hue, chroma and luminance (components of a colour; Quesada & Senar, 2006) for both age classes separately, as we clearly see differences between the intercorrelation of the variables between adults and subadults (Table S3; Figure S6). The PC1 of bill colour explained 88.8% and 69.4% of the variation for adults and subadults, respectively, and loading was approximately equal for each of the three components, except for hue in subadults: hue = 0.99, chroma = −0.90, luminance = 0.95, n = 599; subadults: hue = −0.49, chroma = 0.99, luminance = 0.92, n = 163; see Figures S4–S6 and Tables S5 and S6). For bill colour, we therefore also can conclude that the PCA performs well in summarizing all three correlated colour traits into a single variable.

### 3.2 SEM method: Confirmatory factor analysis using latent variables

A closely related statistical approach to PCA that is available in the SEM framework is a confirmatory factor analysis (CFA). CFA includes two types of variables: observed variables and latent variables which are graphically represented, respectively, by boxes and circles (Figure 2). The ‘confirmatory’ part speaks to the knowledge one has of the theory behind the relationship between the latent and their observed variables. Latent variables are hypothetical or theoretical variables (constructs) that we assume to exist but for which we have no direct measurements, and are therefore referred to as ‘unmeasured variables’ (Grace et al., 2010). Latent variables are used in models to represent an underlying process while observed variables serve as indicators of the effects or manifestations of the latent factors (Grace & Bollen, 2008). Latent variables play an important role in SEM because they represent a bridge between observed data and theoretical generalization (Grace & Bollen, 2008), and a part of the concept of body condition is a prime example of a latent theoretical concept, as these are latent variables made up of typically correlated condition measures (e.g. ratio and residual body mass or BMI).

CFA provides a conceptual advantage: we often choose a single indicator as a surrogate for a latent concept (e.g. hue representing the more general concept ‘colour’). Including more indicators (e.g. chroma, luminance) helps to generalize this phenomenon by testing that the result is not impacted by the choice of any single indicator. This emphasizes that CFA/SEM is a logical choice in the context of the multivariate nature of body condition. The purpose of CFA is to statistically test the ability of a hypothesized (factor) model to reproduce a set of sampled data (usually through a variance-covariance matrix; Nusair & Hua, 2010). In a CFA, the estimated values for all records can be extracted and used for instance for plotting in a similar way as is done with PCA.

The use of latent variables (as a variable reduction tool) is common in social sciences and psychology (Bollen, 2002; Grace et al., 2010) and there is a growing interest in ecological research also. Latent variables are, for instance, used to quantify species performance in plant research (Grace et al., 2016; Travis & Grace, 2010; Visser et al., 2018), or morphometrics in birds (Araya-Ajoy et al., 2019; Cubaynes et al., 2012; Pugesek & Tomer, 1996) and mammals (MacKay et al., 2017; Nespola et al., 2003), as well as to quantify environmental conditions (Guan et al., 2016; Souchay et al., 2018).

![Figure 2](image)

**Figure 2** Path diagrams of the confirmatory factor analysis (CFA) for the latent variables: (a) energy stores (determined by mass corrected for size [equivalent to residuals on the mass axis] and by the ratio of mass/tarsus length) and (b) bill colour (determined by hue, chroma and luminance) for the two age classes. Boxes indicate measured variables; circles represent latent variables. Arrows indicate positive significant (solid line), negative significant (dashed line) or non-significant (dotted line) path strengths with values of standardized coefficients shown next to each line. Standardized coefficients are estimates expressed in equivalent units to make it possible to compare different path strengths in the model. Standardized coefficients (scale standardization) are estimated by multiplying the slope of the ratio of the standard deviation of the predictor and response variable (Lefcheck, 2019). Latent variables are described by arrows going away from the latent variable indicating that they are measured by the observed variables. (c) and (d) show the equations belonging to models (a) and (b), respectively. α, λ and ε illustrate the intercept, slope coefficients and residual error, respectively. See Codes S1 and S2 for detailed model results and specifications.
We created the variable 'energy stores' as a latent variable construct in the SEM framework using CFA (Figure 2a). The variable energy stores was defined by both (a) the ratio of mass to tarsus length and (b) the residual of mass regressed to three structural size measurements (head length, tarsus length and wing length). Model results (Code S1) indicate that the ratio and residual variables are positively related to the latent variable energy stores (Figure 2a).

Similarly, we can create a variable 'bill colour' as a latent variable in the SEM framework (Figure 2b; Code S2), which on a more abstract level can be seen as a common factor of several observed variables (hue, chroma and luminance) so that its associated biological consequence can be viewed to occur at the level of a bill colour factor, rather than at the level of each observed variable. Hue values indicate a scale from orange to yellow, whereas luminance and chroma give information on the lightness and saturation of the colour (see Figure S7 for colour component interpretation). Since we expect that the relationship among the colour variables differs between the two age classes, we conduct a multi-group CFA. The structural equation model (Figure 2b, Code S2) shows that in adults, the hue and luminance are positively intercorrelated, whereas chroma is negatively intercorrelated. In accordance with the PCA results, the CFA for sub-adults shows low intercorrelation of hue with chroma and luminance but hue contributes less to the estimate of 'bill colour' in this group.

To conclude, CFA is a technique within SEM that is conceptually related to the variable reduction technique PCA, and in our case study the variables from the latent variable and PC1 are highly correlated for the energy store \( r = 0.98 \) and the bill colour model (adults: \( r = 0.95 \); sub-adults: \( r = 0.92 \)), showing linear relationships over the entire range (Figure S8). Note that at this point 'bill colour' and 'energy stores' are latent variables which are not yet related in any way to condition. This will follow in the next two sections.

4 | SECTION 2: QUANTIFY BODY CONDITION BASED ON A FITNESS TRAIT ...

MR estimates the relationships between a response variable (e.g. fitness component) and one or more predictor variables (condition variables; Nusair & Hua, 2010). A key characteristic of MR analysis is that it aims to show how variation in one variable can be explained by others. A limitation of MR is that they cannot handle variables being both response and predictor at the same time and therefore are limited to the examination of single process at a time. The commonly used work around is to take a multi-step approach of performing several MR analyses sequentially (e.g. Skarpaas et al., 2011). In our hypothesized conceptual model (box 2 and grey shaded area in Figure 1), body mass is seen as a predictor variable when used to explain variation in the body condition index, but simultaneously as a response variable when regressed to different structural size measurements (e.g. tarsus length; Figure 4a).

We aimed to quantify a body condition index that explains variation in survival and is based on six measured condition variables characterizing the nutritional state, immune state and hormonal state (box 2 in Figure 1; Figure 4a). With an MR approach, we would answer this question by conducting the following four sequential steps. We first need to regress mass on three structural sizes (tarsus length, head length and wing length) to account for individual variation in structural size that we are not interested in (step 1). Next, we extract the residuals of this model, and regress them together with the five other condition variables (buffy coat) on the survival variable (step 2). Then, we extract the predicted values of the survival model as a body condition index (step 3). Only now, in a final fourth step, can we tackle our main question by regressing the body condition index to the environmental variable 'competitor density' to determine the effect of density on 'condition', and thereby answer how competition (e.g. for food sources) affects survival through the condition of an individual. Conducting multiple steps using the residuals of a model as input for another model (also called residual analysis) can lead to biased parameter estimates, especially if independent variables are not totally uncorrelated (Darlington & Smulders, 2001; Freckleton, 2002).
4.2 | ... with SEM method—Composite variable and MIMIC

The hypothesized conceptual model (box 2 in Figure 1) can be fitted in a single model within the SEM framework using ‘composite variables’. Whereas ‘regular’ latent variables represent an unmeasured variable that leads to correlation between the observed variables (Figure 3a), composite variables represent the collective effect of a group of observed or latent variables (Figure 3b; Grace & Bollen, 2008). Composite variables can be seen as a special case of a latent variable and are represented by a latent variable with zero error, which signifies that it is completely specified by its causes (Grace, 2006). Thus, if condition variables are strongly correlated, we advise to use latent variables to integrate them, whereas in composite variables, the observed contributing condition variables do not necessarily need to be correlated, which is an important flexibility of its application (Nusair & Hua, 2010).

In contrast to the use of latent variables in ecology, composite variables have been long recognized as a potentially important element of SEM, but have received very limited use in ecology among others because of a lack of theoretical consideration (Grace & Bollen, 2008; Hopcraft et al., 2012; Pugesek & Tomer, 1996). Composite variables come in two types. One is a fixed-weight composite, in which the loadings (\( \gamma \) in Figure 3d) from the causes are specified a priori (Grace, 2006). An example is the ‘importance value’ in plant ecology, which is defined as the unweighted sum of the relative density, relative abundance and relative frequency (usually for a species within a community). A second type of composite is the unknown-weight composite (Grace & Bollen, 2008). The unknown weight composite is related to a multiple regression predictor, where some weighted combinations of causal influences maximize variance explanation in one or more response variables, and weights are thus estimated from the data. The concept of body condition (box 2 in Figure 1) is a prime example of an ‘unknown-weight composite’ as first, the six condition variables (Figure 4a) show low intercorrelation (Figure S2) and second, we do not know a priori how important each condition measure is for our fitness proxy survival. We created a condition index making use of a composite variable defined as a weighted sum of the condition variables, with the weights reflecting the dependency of survival on each of the condition variables (Figure 4a). This step is basically the same as in an MR. The values of the condition index variable for each record can be extracted from the SEM model (Figure 4a) so that the relationship between condition index and density can be plotted (Figure 4b). Furthermore, building a condition index as a composite variable can also be done by relating it to other fitness traits like reproduction (see Section 6), or the weights can be used to quantify the condition of birds captured in future studies, without the need for measuring their survival. We note that the six condition variables were at most weakly correlated (e.g. correlation between cholesterol and corticosterone was \( r = -0.02 \); Figure S2), and using a latent variable as part of a CFA approach would thus not be an alternative to the composite variable approach taken here.

In an MR, it is not possible to handle mass as a response variable and at the same time as an explanatory variable, forcing the use of other solutions such as using the residuals of mass (as illustrated in Figure 2a). An advantage of the SEM approach is that there is no need to first extract the residuals (of body mass on size) and then summarize the variables; instead, we can include the size, mass and size-corrected-mass variables in the same model (Figure 2; see also next paragraph). In addition, in the same SEM, we specify condition as a function of survival probability, and in a separate equation (but

\[
\begin{align*}
X_1 &= \alpha_1 + \lambda_1 L + \varepsilon_{X1} \\
X_2 &= \alpha_2 + \lambda_2 L + \varepsilon_{X2} \\
X_3 &= \alpha_3 + \lambda_3 L + \varepsilon_{X3}
\end{align*}
\]

\[
Y = \alpha_Y + \gamma_Y C + \varepsilon_Y
\]

\[
C = \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \varepsilon_C
\]

With \( \gamma_{1-3} \) being estimated using:

\[
Y = \alpha_z + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \varepsilon_z
\]

**Figure 3** Graphical representation of a model using a latent (a) variable and a composite (b) variable. The path diagram shows the relationships between observed variables (X’s in boxes) and (a) latent (L; in circle) or (b) unknown-weight composite (C; in hexagon) variable. Note that for the unknown-weight composite in (b), a response variable \( Y \) is required, while for defining a latent variable in (a) this is not a necessity, but could be added if useful. The equations for the latent and composite variable are shown in (c) and (d), respectively. (c) \( X \) is the result of the influence of the latent variable (L), proportional to \( \lambda \) (= loading), plus the error \( \varepsilon \) and the intercept \( \alpha \). For details on how \( L \) and \( \lambda \) are derived from data, see chapter 4 in Grace (2006); (d) composite variable \( C \) is equal to the sum of the effects of \( X \) plus the error which is fixed to 0. The response variable \( Y \) is a result of the intercept \( \alpha_Y \), the slope coefficient \( \gamma \), \( C \) and the error \( \varepsilon \).
in the same model; Figure 4c) regress density on the condition index (see Code S3 for model specifications and R-code). This has the additional advantage that an overall model fit can be assessed, to quantify to what extent the hypothesized model fits the sample data (Musil et al., 1998) or to compare it with potential alternative models, which will be discussed in detail in Section 3.

Since larger individuals may have a higher body weight, we want to account for the size when including mass as a condition measure in the model. This is often done by taking the residuals of mass regressed on structural size variables (Jakob et al., 1996; Peig & Green, 2010). SEM avoids using residuals (as in Figure 2a) when considering mass corrected for size as a condition variable using a technique called the multiple-indicator, multiple-cause model (MIMIC). MIMIC incorporates covariates of interest which are modelled under a latent variable framework (Chang et al., 2020). Specifically, in our case study, we create a new latent variable ‘Size-corrected-mass’ (Figure 4) affecting only one observed variable ‘mass’ with a default factor loading of 1. The structural size covariates (tarsus length, head length and wing length) affect the latent variable ‘Size-corrected-mass’, rather than the observed variable mass. By incorporating the size variables as covariates in this way in SEM, the latent construct ‘Size-corrected-mass’ allows for the model to estimate the relationship of mass on condition while accounting for the effect of size-variables on mass, similar to taking the residuals of mass regressed on size when using a stepwise MR approach. In addition, there may be relationships between the size variables and other condition variables (e.g. cholesterol) that could be considered in the model.

So far, we only considered linear relationships, but often relationships are curvilinear in biology. To illustrate how to model nonlinear relationships, we run the model from Figure 4 with an additional variable ‘squared mass’ (Code S4). By doing so, we can investigate a possible nonlinear relationship of mass on condition. However, in this case study, squared mass does not show a significant effect on condition ($p = 0.07$, Code S4), and we henceforth ignore it for reasons of simplicity.

The model results show that size-corrected mass and corticosterone have a negative effect on the condition index, whereas buffy coat, haematocrit, uric acid and cholesterol show smaller (non-significant) effects on condition (Figure 4a). The condition has a path strength of 0.41 on survival, explaining 17% of the spatial variation in survival within the Netherlands (Code S3; Figure S9). Oystercatcher density shows a strong ($-0.57$) negative association with the body condition index (Figure 4b; Figure S10).

## 5 | SECTION 3: HIGH FLEXIBILITY, COMPARING COMPETING HYPOTHESES, AND MODEL PERFORMANCE OF SEM

In this section, we combine the two models from Sections 1 and 2 and include confounding variables related to measurement procedure (handling time) and individual characteristics (sex, feeding specialization and age class; box 3 in Figure 1). Next, we compare the model performance of a SEM approach and a conventional statistical approach.

### 5.1 Model results and interpretation

To construct the full model (Figure 5), we connect the latent variable ‘energy store’ (Figure 2a) with the composite variable body condition (Figure 4). In addition, by using MIMIC modelling, we correct for the effect of measurement procedure (handling time) and individual (sex, age)-related confounding variables to account for its effect on the condition variables. The correlation structure of the three colour parameters (hue, chroma and luminance) differs strongly between the two age classes. The colour parameters are only highly correlated in adults and not in sub-adults (Table S3) as illustrated in the multi-group CFA (Figure 2b). Therefore, it is not useful to use a latent variable for bill colour in this example. Instead, for simplicity, we use only hue as a measure of bill colour (Figure 5), which we expect to have more variation giving information on the colour in contrast to luminance and chroma, which reflects lightness (Figure S7). For completeness, we illustrate alternative model structures to handle the different correlation structures between age classes in Figure S11.

Whereas MR assumes independence of error terms, which may bias the estimates (James et al., 2013), it is possible to specify non-independence of error terms in SEM. For instance, in our example, we expect the error term of haematocrit and cholesterol to be correlated because cholesterol is quantified as a concentration in blood, which means that it is related to the proportion of haematocrit in the blood (Code S5). Similarly, in the example of modelling nonlinearity in mass using squared mass as additional variable, we can specify that the error terms of mass and its squared version are correlated.

The full model (Figure 5) shows that the latent variable energy stores is a result of the positive influence of mass (corrected for structural size variables and handling time) as well as the ratio of mass divided by tarsus length. We also find that individuals with a higher bill tip height (blunter shaped bill) show a more yellowish bill (standardized estimate = 0.22, $p < 0.001$; Code S5). In addition, males have a more orange than yellowish bill compared to females (standardized estimate = $-0.25$, $p < 0.001$; Code S5). Most importantly, we find that energy stores and corticosterone have a negative effect on the condition, whereas the other condition variables (buffy coat, haematocrit, cholesterol, uric acid and bill colour) show weaker (non-significant) effects on the condition index.

These aforementioned results emphasize the point made in the introduction that we often assume positive relationships with performance or fitness (e.g. high energy stores benefit survival), but that those assumptions do not always hold true. We stress that the negative contribution of the latent construct ‘energy stores’ to body condition results from a negative correlation between mass (corrected for size) and survival (Figure S12), and is thus not a result of the (SEM) modelling procedure itself. A possible explanation for the negative relationship of energy stores with condition may be that birds with higher condition are mainly in areas with more food and therefore
do not need to store that many resources, as the food source is more reliably available (Cuthill et al., 2000). Another possible reason may be that the number of predatory birds (e.g. peregrine falcon \textit{Falco peregrinus}) in the oystercatcher wintering grounds increased during recent decades (Sovon Vogelonderzoek Nederland, 2018) which may increase the survival of lighter, more agile individuals by allowing them to escape predation. Lastly, it is important to keep in mind that we here analyse site-level variation in mortality rather than individual-level mortality, and between and within site (population) patterns need not be similar. Thus, the negative relationship between energy stores and condition does not exclude that energy stores in principle are important for starvation just that under conditions of high food stocks or predator abundance encountered in our study, energy stores may potentially be unrelated to starvation and instead reflect other aspects, such as foraging efficiency or agility.

We a priori hypothesized a positive relationship of body mass on the condition index, and therefore called the latent variable ‘energy stores’. We can learn from the outcome of this model that our data are not consistent with the theoretical ideas that led us to model the concept of energy stores in the way we did. However, our latent variable ‘energy stores’ could be consistent with interpreting this latent variable as the concept ‘mass-related agility’ of an individual, but this hypothesis would require further testing. We note that the concept ‘mass-related agility’ could still be thought of as a condition variable, as it reflects an aspect of health status (responsiveness to danger) and predicts performance (the fitness proxy survival).

Our results also confirmed the expected negative relationship between corticosterone and the condition index, meaning that birds at sites with higher survival had lower corticosterone levels. Corticosterone fulfills its main functions by mobilizing stored resources to facilitate upregulating metabolism for coping with increased energetic challenges (Jimeno et al., 2018). Thus, birds coping with harsh environmental conditions may show elevated corticosterone secretion (Angelier et al., 2010; Jimeno et al., 2017), and our results are consistent with this idea.

We included bill colour as a potential variable affecting individuals’ condition, as we hypothesized a priori that such a carotenoid-based trait is likely to be sensitive to nutritional state and may thus serve as a visual status signal of the individual condition to conspecifics meaning that individuals with a more reddish bill have higher condition. However, on the other hand, if bill colour would be an ‘honest’ signal for sexual selection, it may also be costly and could...
result in a negative relationship with survival. The pathway of bill colour on condition shows no significant effect on the condition index and survival. We can learn from the model that bill colour is not a useful condition variable that predicts variation in survival performance in this population.

Finally, we found that the body condition index shows a strong (~0.55) negative association with oystercatcher density, whereas body condition positively affected annual survival probability (0.45 standardized path coefficient; Figure 5). This result suggests that that competition may affect the survival through body condition, and that this mediating effect of condition is quite strong: a one standard deviation decrease oystercatcher density at a site is associated with a quarter standard deviation increase in annual survival (~0.55 \times 0.45 = −0.25; rules of path tracing; e.g. Grace, 2006).

5.2 | Competing hypothesis and model fit comparison

Researchers may want to compare competing multivariate descriptions of a conceptual model (based on a priori hypotheses), but formal ways to compare such models and identify whether they are sufficient descriptions in the first place have rarely been considered. In SEM, we obtain one overall model fit and can compare model fits between models, which makes it possible to learn about a system. Various metrics of model fit exist for SEM, but we focus here on the comparative fit index (CFI; Bentler, 1990) that is mostly reported in ecological studies (see Table S7 for other metrics). CFI is a chi-square estimate using a maximum-likelihood solution, where values ≥0.90 and ≥0.95 indicate an acceptable and reasonably good fit, respectively (Hooper et al., 2008; Hu & Bentler, 1999; Lefcheck, 2019).

We constructed an alternative model to compare with the original one (Figure 5) to explore whether we can improve the model. In the original model, by design, all effects of body size go indirectly through condition, whereas by adding a direct pathway of a latent construct ‘body size’ on survival, we can learn the relative role of body size and body condition (which is often understood as being independent; Schulte-Hostedde et al., 2005). Both models show good model fit with a CFI of 0.94 (Codes S6 and S7). The alternative model reveals no direct effect of body size on survival, but only an indirect effect through mass and condition (Figure S13). Comparing both models with an ANOVA confirms that the alternative model

---

**FIGURE 5** Full SEM model including latent (circles), composite (hexagon) and measured (squares) variables as well as confounding, predictor and response variables. Values indicate the standardized path strength (see Figure 2 for explanation). Solid and dashed lines indicate a positive and negative significant (p < 0.05) path and are indicated by an asterisk next to the standardized path strength, respectively. Dotted lines indicate non-significant pathways (p > 0.05). For simplicity, we only documented the path strength of the condition, predictor and response variables. The MIMIC part of the condition variables (buffy coat, haematocrit, cholesterol, uric acid and corticosterone) contains latent variables of the observed condition variables (indicated by an apostrophe) that are corrected for confounding variables (e.g. sex, age). ‘Bill colour’ and ‘size-corrected mass’ are also MIMIC parts (corrected for confounding variables) within the model, but have another biological meaningful name (rather than the name of the measured variable) and therefore are not noted without apostrophe.
(with an additional direct effect of body size on survival) does not show a significant improvement ($p = 0.4$; Table S8) compared to the original model. We have thus learned that effects of body size primarily affect survival through its effect on body condition. This model comparison also illustrates that if no direct arrow connects two variables, independence between variables is assumed (e.g. between size or sex and survival; Figure 5). The decision on which pathways to include in a model or not, should be based on prior biological knowledge and the aim of the study, and thus requires careful thinking about hypotheses.

5.3 Model performance of SEM versus Conventional approach

In the conventional approach, we can imitate latent variables using PCA and composite variable using residuals from MR (Figure 6a vs. 6b). However, using conventional methods (MR and PCA) to create a condition index that predicts survival requires a large number of steps, especially when multiple condition variables are measured (Figure 6b). We expected the SEM to perform better than the conventional approach, as in SEM we can estimate the joint likelihood of all the relations simultaneously (Figure 6a), contrasting with the four steps in the conventional approach (Figure 6b). The conventional approach involves using the residuals from a previous model in input in the actual model, which may cause bias in parameter estimation (Darlington & Smulders, 2001; Freckleton, 2002).

We compared how well the body condition index could predict survival using either the SEM or the conventional method (Figure 6). Model performance was quantified using a measure of prediction error (root mean square error; lower values imply less deviations between model predictions and observed survival estimates) and a measure of explanatory power ($R^2$; higher values indicating more variation explained). We applied 10-fold cross-validation to the original dataset, with 80% of the dataset being used as training dataset to which the SEM (Figure 6a) and conventional model (Figure 6b) were fitted and 20% being used as test dataset to assess model performance of both methods. This process was repeated three times, meaning that in total 30 different training and test sets were used to estimate the performance (see Codes S8 and S9). The results of the cross validation show that the SEM approach has higher model performance which is indicated by a 1% reduction of the prediction error (RMSE) and an increase of the explained variation ($R^2$) by 9% compared to the conventional (MR/PCA) approach (SEM: RMSE = 0.438; $R^2 = 0.177$; MR&PCA approach: RMSE = 0.442; $R^2 = 0.163$; Figure 6c).

**FIGURE 6** Comparison of two analysis approaches using (a) structural equation modelling (SEM) and (b) conventional methods: multiple regression (MR) and principal component analysis (PCA). Arrows point from predictor to response variable within the model. (a) SEM to quantify body condition in oystercatcher that is positively related to survival including latent and composite variables in one step, (b) MR&PCA approach to quantify body condition which needs 4 steps (compared to 1 step in SEM). (c) shows the root mean square error (RMSE) and $R^2$ for both models as a result from cross validation. Note that the variable ‘competitor density’ (from Figure 5) was not included in the SEM model to compare the model performance of both approaches (a and b) since the aim was to create a proxy of body condition that best explains the variation in survival and therefore predicts the survival well with a train and test dataset (cross validation). See Code S8 and Table S9 for model specifications and R code as well as the details of each step of the MR&PCA approach, respectively.
In addition, we compared the bias and precision of estimators from both approaches. To this end, we simulated 1,000 datasets with 1,000 observations each, based on the covariance patterns found in the case study population using the 'simulateData' function in lavaan (Rosseel, 2012). Next, we estimated the effect of each condition variable on condition index using either a SEM (Figure S14) or conventional stepwise MR&PCA (Figure 6b; Figure S15) approach. For each condition variable of both approaches, we calculated the mean bias ($\bar{\epsilon} = \frac{1}{n_{\text{sim}}} \sum_{i=1}^{n_{\text{sim}}} (\hat{\theta}_i - \theta)$) and the average empirical standard error ($\bar{E} = \sqrt{\frac{1}{n_{\text{sim}}} \sum_{i=1}^{n_{\text{sim}}} (\hat{\theta}_i - \bar{\theta})^2}$) of the estimate, with $\theta$ being the parameter value used to simulate the data (‘true value’), $\hat{\theta}$ being the estimated parameter from the simulated $i$th repetition, $\bar{\theta}$ being the mean of the estimated parameters and $n_{\text{sim}}$ being the number of simulated datasets.

Simulation results indicate that the SEM method is practically unbiased while the conventional method can have substantial bias and not consistently in the same direction (Figure 7). For condition variables having a ‘true’ value not close to 0, we calculated the relative bias (Morris et al., 2019) which ranged from −85% to 50% and 0.8% to 3.5% for the conventional approach and SEM, respectively (Figure S16). The bias of the conventional estimators is largest in the variables having the strongest (significant) effect on the condition index (Figure 7; Figure S17), ranging from −0.16 to +0.21 absolute bias (Figure S18). Overall, the estimator of the SEM approach was also more precise (with an empirical standard error around 0.056) than the conventional approach estimator (with an empirical standard error around 0.087), when considering estimators that were unbiased in both methods (e.g. uric acid; Figures S19 and S20).

### 6 | DISCUSSION

We showed how to use structural equation modelling (SEM) to quantify body condition, comparing it with the conventional approaches of multiple regression and/or principal component analysis. We illustrated how latent variables can be used to summarize the multidimensional nature of body condition, yielding variable reduction. Next, we showed that composite variables make it possible to define the body condition index based on associations between observed condition variables and a fitness trait, even if the different condition variables contributing to body condition are uncorrelated (potentially reflecting different axes of condition). This is advantageous, because the direction of a relationship between a condition measure (e.g. body mass) and fitness is often assumed, but rarely known with certainty. Finally, we showed how it is possible in SEM to combine latent and composite variables in one model, which more realistically reflects complex natural situations, as it allows for summarizing both strongly correlated and less correlated measures of condition. A major advantage of SEM is that, in contrast to conventional methods, it can handle variables that are response and explanatory variables at the same time. As a consequence, SEM allows all the relationships to be estimated in a single model, resulting in one estimate for the model fit and generating the possibility to compare competing models, whereas in conventional methods several steps are required (e.g. extracting residuals of a model and use them as input for another model) that provide additional challenges (e.g. biased parameter estimates and carrying over the uncertainty from previous models).

In addition, we showed how we can correct for the effect of confounding variables in SEM using MIMIC (multiple-indicator, multiple-cause) modelling. Overall, the high flexibility and integration of SEM make it a powerful tool that might increase model predictability and result in unbiased and precise estimates, as illustrated by our worked examples on a case study, and therefore may also provide novel insights in ecological processes.

While we used SEM to quantify body condition, there are numerous other contexts in which it can be used to describe complex concepts. For example, in determining an immune response or experienced stress level of an individual, because usually multiple immunological, physiological or hormonal metrics are measured, which may also reflect different aspects of the immune or stress response, and it is a challenge to derive an integrated immune estimate from these data. Similarly, climatic conditions are highly multivariate, and while most studies have, for example, detailed climatic data on various weather variables, studies tend to focus on a single weather variable such as temperature.

### 6.1 | From a biological point of view: Body condition related to different performance metrics

A condition index based on predicting survival does not necessarily mean that a high body condition of an individual indicates high fitness, because fitness also depends on the association between condition and other fitness components, such as reproductive success. Quantifying condition based on survival data may result in a different condition index than when it is based on the reproductive success of an individual, or a fitness metric that integrates survival and reproduction, such as reproductive value. This difference highlights that definitions of condition can be context-dependent and may vary during the life cycle, which may be particularly relevant when investigating species that alternate strategies (e.g. albatrosses prioritizing reproduction in one year and survival in another year; Froy et al., 2013). Our SEM approach provides a framework for identifying how different condition measures define a successful survivor or a successful breeder and how they may differ from each other.

### 6.2 | From a statistical point of view: Additional techniques, challenges and limitations

We want to emphasize that in order to compare the two approaches, the data used to analyse the data and generating the data had the same model structure. However, in reality, this will rarely be the case. Different research questions require different statistical
frameworks (model structures) and a thorough understanding of the definitions and components of the frameworks as well as of the study system is crucial for making biologically meaningful inferences (Benthem et al., 2017). Therefore, identifying a proper model structure (that fits the data well) and that addresses a biological question properly is by no means trivial and probably one of the largest challenges in the art of modelling.

We considered confounding variables (e.g. sex and age) to illustrate how it can be modelled in SEM in general, but there may be many more confounding variables, and most will not have been measured (e.g. social dominance on foraging grounds). The presence of unmeasured confounders is a general challenge in observational studies (VanderWeele & Arah, 2011), but sensitivity analysis and bias-modelling techniques can help handling uncontrolled confounding variables (Lin et al., 1998; McCandless et al., 2007).

In this paper, we used latent variables to achieve variable reduction (similar to PCA). However, latent variables can also be used to account for measurement errors of observed variables. Conventional statistical approaches like MR assume that variables are measured without error (Pugesek & Tomer, 1995) even though it is known that it is usually impossible to measure variables without error, particularly in the field (Musil et al., 1998). Ignoring measurement errors typically leads to downward bias in parameters because an error in measuring X (explanatory variable) is assigned to the error in predicting Y (response variable), implying that the true effect of X on Y is typically underestimated (i.e. regression dilution/attenuation). Pugesek and Tomer (1995) showed that SEM, by including measuring errors, estimated parameter coefficients more accurately and with less bias compared to MR. Code S10 shows an example in which we account for imperfect measurements (through latent variables) of mass regressed on body size structures (tarsus, head and wing length).

SEM statistical packages are under continuous development, but particularly frequentist statistical approaches based on maximum likelihood estimation still have limitations. Random effects in combination with latent variables are currently challenging to model with one of the most widely used R-packages, lavaan (Rosseel, 2012). In the R-package piecewisesem (Lefcheck, 2016), it is possible to model random effects when conducting SEM to identify direct and indirect pathways, but it is currently impossible to model latent variables with this package.

For categorical predictor variables, a multi-group analysis can be conducted. Multi-group analysis can also be an alternative way to account for random factors with few levels. The question that drives a multi-group analysis is whether two or more groups might differ in terms of the relationships among parameters and the whole model is run for each level (Grace, 2006). To illustrate the use of categorical variables, we used the age effect in the example of Figure 2b as a category (Code S2), which means that the model calculates the relationships (Figure 2d) for each age class separately. In addition, multi-group analysis can also be conducted to test for interactions (with a categorical variable). However, a multi-group analysis together with a binomial response variable is not supported yet in lavaan which raises practical challenges analysing 0/1 survival data using logistic regression in combination with categorical variables, such as area-specific resighting probability which were relevant in our case study. An example of R-code for modelling a binomial response variable but without multiple levels (only one area) is presented in Code S11.

Nonlinear relationships are common in ecology and there are different ways to address nonlinearity. The easiest way is
through transformation, as we did in our case study, applying logit-transformation for proportional response variables and squaring mass (Code S4). Multi-group analysis can also be used to address a certain type of nonlinear relation (Grace, 2006). Continuous nonlinear relationships are a special challenge for SEM, both for observed but especially for latent variables (Grace, 2006). A useful introduction to this topic is chapter 12 in Kline (2011).

Bayesian approaches to SEM may offer solutions to limitations of frequentist software in dealing with random effects, non-Gaussian data and multiple-level categorical predictors, because Bayesian software uses more convenient numerical algorithms. Furthermore, a Bayesian approach allows for constraining estimates using informative priors, so as to include biological knowledge. Bayesian framework has also been used to combine a mark–recapture model with SEM. Cubaynes et al. (2012) show an example of how to model a latent variable (‘overall body size’) that affects survival estimated within a mark–recapture framework (all in one model). Possible latent variables that affect survival include biological knowledge. Bayesian framework has also been used to combine a mark–recapture model with SEM. Cubaynes et al. (2012) show an example of how to model a latent variable (‘overall body size’) that affects survival estimated within a mark–recapture framework (all in one model). Possible latent variables that affect survival include biological knowledge. Bayesian framework has also been used to combine a mark–recapture model with SEM. Cubaynes et al. (2012) show an example of how to model a latent variable (‘overall body size’) that affects survival estimated within a mark–recapture framework (all in one model). Possible latent variables that affect survival include biological knowledge.

To conclude, we show that SEM is a powerful and flexible statistical tool that can lead to models of higher predictive power and with more accurate as well as precise estimates compared to conventional approaches. Therefore, we encourage researchers to consider SEM as a flexible framework to describe the multivariate nature of body condition and thus understand how it affects biological processes, thereby improving the value of body condition proxies for predicting organismal performance. We emphasize that SEM can also be a useful tool for other multidimensional ecological concepts as well, such as immunocompetence, oxidative stress and environmental conditions.

ACKNOWLEDGEMENTS

We thank everybody involved in catching birds, especially Kees Oosterbeek, Symen Deuzeman and Lauren van Kooten. Funding was provided by the Applied and Engineering Sciences domain of the Netherlands Organisation for Scientific Research (NWO-STW; grant number 14638) as well as by Royal Netherlands Air Force, Birdlife Netherlands, NAM gas exploration and Deltares. We acknowledge the feedback provided by ecologists working at these organizations during half-yearly meetings. The authors declare that they have no competing interests. We also thank Nigel G. Yoccoz and two anonymous reviewers whose comments helped improve the manuscript.

CONFLICT OF INTEREST

All authors declare no conflict of interest.

AUTHORS’ CONTRIBUTIONS

M.F. and M.v.d.P. conceptualized the research, with A.M.A., B.J.E., E.J., H.d.K., H.-J.v.d.K., J.N. and S.V. contributed to its development; M.F. and H.-J.v.d.K. managed the field data; M.F. performed the analysis and drafted the manuscript. All authors contributed to the final draft and approved the final manuscript.

DATA AVAILABILITY STATEMENT

Data files and R code to fit the models and simulations are available on GitHub and can be accessed via https://github.com/MagaliFr/QuantifyingBodyConditionWithSEM or https://doi.org/10.5281/zenodo.5153493 (Frauendorf et al., 2021).

ORCID

Magali Frauendorf https://orcid.org/0000-0003-1608-8396
Andrew M. Allen https://orcid.org/0000-0002-0119-2425
Simon Verhulst https://orcid.org/0000-0002-1143-6868
Elke Jongejans https://orcid.org/0000-0003-1148-7419
Bruno J. Ens https://orcid.org/0000-0002-4659-4807
Henk-Jan van der Kolk https://orcid.org/0000-0002-8023-379X
Hans de Kroon https://orcid.org/0000-0001-6151-3561
Martijn van de Pol https://orcid.org/0000-0003-4102-4079

REFERENCES

Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wires Computational Statistics, 2, 433–459. https://doi.org/10.1002/wics.101
Allen, A. M., Ens, B. J., van de Pol, M., van der Jeugd, H., Frauendorf, M., Oosterbeek, K., & Jongejans, E. (2019). Seasonal survival and migratory connectivity of the Eurasian Oystercatcher revealed by citizen science. The Auk, 136(1), 1–17. https://doi.org/10.1093/auk/uyk001
Alonso-Alvarez, C., Ferrer, M., & Velando, A. (2002). The plasmatropic index of body condition in Yellow-legged Gulls Larus cachinnans: A food-controlled experiment. Ibis, 144, 147–149. https://doi.org/10.1046/j.0019-1019.2001.00023.x
Angelier, F., Wingfield, J. C., Weimerskirch, H., & Chastel, O. (2010). Hormonal correlates of individual quality in a long-lived bird: A test of the ‘corticosterone–fitness hypothesis’. Biology Letters, 6, 846–849. https://doi.org/10.1098/rsbl.2010.0376
Araya-Ajoy, Y. G., & Dingemanse, N. J. (2014). Characterizing behavioural ‘characters’: An evolutionary framework. Proceedings of the Royal Society B: Biological Sciences, 281(1776), 20132645. https://doi.org/10.1098 rspb.2013.2645
Araya-Ajoy, Y. G., Ranke, P. S., Kvalnes, T., Renning, B., Holland, H., Myhre, A. M., Pärn, J., Jensen, H., Ringsby, T. H., Saether, B.-E., & Wright, J. (2019). Characterizing morphological (co)variation using structural equation models: Body size, allometric relationships and evolvability in a house sparrow metapopulation. Evolution, 73(3), 452–466. https://doi.org/10.1111/evol.13668
Atchley, W. R., & Anderson, D. (1978). Ratios and the statistical analysis of biological data. Systematic Zoology, 27(1), 71. https://doi.org/10.2307/2412816
Aubry, L. M., Rockwell, R. F., Cooch, E. G., Brook, R. W., Mulder, C. P., & Koons, D. N. (2013). Climate change, phenology, and habitat degradation: Drivers of gosling body condition and juvenile survival in lesser snow geese. Global Change Biology, 19(1), 149–160. https://doi.org/10.1111/gcb.12013
Barry, K. L., & Wilder, S. M. (2013). Macronutrient intake affects reproduction of a predatory insect. Oikos, 122(7), 1058–1064. https://doi.org/10.1111/j.1600-0706.2012.00164.x
Beentema, A. J. (1994). Condition indices for wader chicks derived from bodyweight and bill-length. Bird Study, 41(1), 68–75. https://doi.org/10.1080/00063659409477199
effect of sex, age and race on estimating percentage body fat from body mass index: The Heritage Family Study. *International Journal of Obesity, 26*(6), 789–796. https://doi.org/10.1038/sj.ijo.0802006

Jakob, E. M., Marshall, S. D., & Uetz, G. W. (1996). Estimating fitness: A comparison of body condition indices. *Oikos, 77*(1), 61–67. https://doi.org/10.2307/3545585

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning—With applications in R (Vol. 103). Springer. https://doi.org/10.1007/978-1-4614-7138-7

Janssen, M. H., Arcese, P., Kyser, T. K., Bertram, D. F., & Norris, D. R. (2012). Test of a body condition index with amphibians. *Journal of Herpetology, 46*(3), 346–350. https://doi.org/10.1670/10-292

MacKay, A. E., Forsyth, D. M., Coulson, G., & Festa-Bianchet, M. (2017). Maternal resource allocation adjusts to timing of parturition in an asynchronous breeder. *Behavioral Ecology and Sociobiology, 72*(1), 7. https://doi.org/10.1007/s00265-017-2419-9

Marra, P. P., & Holberton, R. L. (1998). Corticosterone levels as indicators of habitat quality: Effects of habitat segregation in a migratory bird during the non-breeding season. *Oecologia, 116*, 284–292. https://doi.org/10.1007/s0044420005090

McCandless, L. C., Gustafson, P., & Levy, A. (2007). Bayesian sensitivity analysis for unmeasured confounding in observational studies. *Statistics in Medicine, 26*(11), 2331–2347. https://doi.org/10.1002/sim.2711

McGuire, L. P., Kelly, L. A., Baloun, D. E., Boyle, W. A., Cheng, T. L., Clerc, J., Fuller, N. W., Gerson, A. R., Jonasson, K. A., Rogers, E. J., Sommers, A. S., & Guglielmo, C. G. (2018). Common condition indices are not more effective than body mass for estimating fat stores in insectivorous bats. *Journal of Mammalogy, 99*(5), 1065–1071. https://doi.org/10.1093/jmammal/gyy103

Metón, I., Mediavilla, D., Caseras, A., Cantó, E., Fernández, F., & Baanante, I. V. (1999). Effect of diet composition and ration size on key enzyme activities of glycolysis–gluconeogenesis, the pentose phosphate pathway and amino acid metabolism in liver of gilthead seabream (*Sparus aurata*). *British Journal of Nutrition, 82*(3), 223–232. https://doi.org/10.1079/000711499001403

Milenkaya, O., Catlin, D. H., Legge, S., & Walters, J. R. (2015). Body condition indices predict reproductive success but not survival in a sedentary, tropical bird. *PLoS ONE, 10*(8), e0136582. https://doi.org/10.1371/journal.pone.0136582

Moiron, M., Araya-Ajoy, Y. G., Mathot, K. J., Mouchet, A., & Dingemanse, N. J. (2019). Functional relations between body mass and risk-taking behavior in wild great tits. *Behavioral Ecology, 30*(3), 617–623. https://doi.org/10.1093/beheco/aryl199

Montreuil-Spencer, C., Schoenemann, K., Lendvai, Á. Z., & Bonier, F. (2019). Winter corticosterone and body condition predict breeding investment in a nonmigratory bird. *Behavioral Ecology, 30*(6), 1642–1652. https://doi.org/10.1093/beheco/arz129

Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine, 38*(11), 2074–2102. https://doi.org/10.1002/sim.8086

Musil, C. M., Jones, S. L., & Warner, C. D. (1998). Structural equation modeling and its relationship to multiple regression and factor analysis. *Research in Nursing & Health, 21*, 271–281. https://doi.org/10.1002/(SICI)1098-240X(199806)21:3<271::AID-NUR10>3.0.CO;2-G

Neff, B. D., & Cargnelli, L. M. (2004). Relationships between condition factors, parasite load and paternity in bluegill sunfish, *Lepomis macrochirus*. *Environmental Biology of Fishes, 71*, 297–304. https://doi.org/10.1007/s10641-004-1263-8

Nespolo, R. F., Arim, M., & Bozinovic, F. (2003). Body size as a latent variable in a structural equation model: Thermal acclimation and energetics of the leaf-eared mouse. *Journal of Experimental Biology, 206*(13), 2145–2157. https://doi.org/10.1242/jeb.00396

Nie, Y., Zhang, Z., Raubenheimer, D., Elser, J. J., Wei, W., Wei, F., & Kay, A. (2014). Obligate herbivory in an ancestrally carnivorous lineage: The giant panda and bamboo from the perspective of nutritional geometry. *Functional Ecology, 29*(1), 26–34. https://doi.org/10.1111/1365-2435.12302

Norte, A. C., Ramos, J. A., Sousa, J. P., & Sheldon, B. C. (2009). Variation of adult Great Tit *Parus major* body condition and blood parameters in relation to sex, age and season. *Journal of Ornithology, 150*(3), 651. https://doi.org/10.1007/s10336-009-0387-1

Nusair, K., & Hua, N. (2010). Comparative assessment of structural equation models to evaluate statistical methods. *Behavioral Ecology, 30*(3), 617–623. https://doi.org/10.1093/beheco/aryl199

Nuttall, F. Q. (2015). Body mass index: Obesity, BMI, and health: A critical review. *Nutrition Today, 50*(3), 117–128. https://doi.org/10.1097/NT.0000000000000092
composition in birds. *Physiological Zoology*, 67(2), 305–329. https://doi.org/10.1086/physzool.67.2.30163851

VanderWeele, T. J., & Arah, O. A. (2011). Unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*, 22(1), 42–52. https://doi.org/10.1097/EDE.0b013e3181f4493

Verhulst, S. (1998). Multiple breeding in the Great Tit; II. The costs of rearing a second clutch. *Functional Ecology*, 12, 132–140. https://doi.org/10.1046/j.1365-2435.1998.00165.x

Verhulst, S., Oosterbeek, K., Rutten, A. L., & Ens, B. J. (2004). Shellfish fishery severely reduces condition and survival of oystercatchers despite creation of large marine protected areas. *Ecology and Society*, 9(1), 17. https://doi.org/10.5751/ES-00636-090117

Visser, M. D., Muller-Landau, H. C., Schnitzer, S. A., Kroon, H., Jongejans, E., & Wright, S. J. (2018). A host-parasite model explains variation in liana infestation among co-occurring tree species. *Journal of Ecology*, 106(6), 2435–2445. https://doi.org/10.1111/1365-2745.12997

von Engelhardt, N., Carere, C., Dijkstra, C., & Groothuis, T. G. G. (2006). Sex-specific effects of yolk testosterone on survival, begging and growth of zebra finches. *Proceedings of the Royal Society B: Biological Sciences*, 273(1582), 65–70. https://doi.org/10.1098/rspb.2005.3274

Wiersma, P., & Piersma, T. (1995). Scoring abdominal profiles to characterize migratory cohorts of shorebirds: An example with red knots. *Journal of Field Ornithology*, 66(1), 12.

Wilder, S. M., Raubenheimer, D., & Simpson, S. J. (2016). Moving beyond body condition indices as an estimate of fitness in ecological and evolutionary studies. *Functional Ecology*, 30(1), 108–115. https://doi.org/10.1111/1365-2435.12460

Williams, C. T., Kitaysky, A. S., Kettle, A. B., & Buck, C. L. (2008). Corticosterone levels of tufted puffins vary with breeding stage, body condition index, and reproductive performance. *General and Comparative Endocrinology*, 158(1), 29–35. https://doi.org/10.1016/j.ygcen.2008.04.018

Wilson, A. J., & Nussey, D. H. (2010). What is individual quality? An evolutionary perspective. *Trends in Ecology & Evolution*, 25(4), 207–214. https://doi.org/10.1016/j.tree.2009.10.002

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Frauendorf, M., Allen, A. M., Verhulst, S., Jongejans, E., Ens, B. J., van der Kolk, H.-J., de Kroon, H., Nienhuis, J., & van de Pol, M. (2021). Conceptualizing and quantifying body condition using structural equation modelling: A user guide. *Journal of Animal Ecology*, 90, 2478–2496. [https://doi.org/10.1111/1365-2656.13578](https://doi.org/10.1111/1365-2656.13578)