

# Improving structured population models with more realistic representations of non-normal growth

Megan L. Peterson<sup>1</sup>  | William Morris<sup>2</sup> | Cristina Linares<sup>3</sup> | Daniel Doak<sup>1</sup>

<sup>1</sup>Environmental Studies Program, University of Colorado Boulder, Boulder, Colorado

<sup>2</sup>Biology Department, Duke University, Durham, North Carolina

<sup>3</sup>Department of Evolutionary Biology, Ecology and Environmental Sciences, Institut de Recerca de la Biodiversitat (IRBIO), University of Barcelona, Barcelona, Spain

## Correspondence

Megan L. Peterson

Email: [megan.peterson@colorado.edu](mailto:megan.peterson@colorado.edu)

## Funding information

National Science Foundation, Grant/Award Number: 1242355, 1242558, 1340024, 1753954 and 1753980

Handling Editor: Res Altwegg

## Abstract

1. Structured population models are among the most widely used tools in ecology and evolution. Integral projection models (IPMs) use continuous representations of how survival, reproduction and growth change as functions of state variables such as size, requiring fewer parameters to be estimated than projection matrix models (PPMs). Yet, almost all published IPMs make an important assumption that size-dependent growth transitions are or can be transformed to be normally distributed. In fact, many organisms exhibit highly skewed size transitions. Small individuals can grow more than they can shrink, and large individuals may often shrink more dramatically than they can grow. Yet, the implications of such skew for inference from IPMs has not been explored, nor have general methods been developed to incorporate skewed size transitions into IPMs, or deal with other aspects of real growth rates, including bounds on possible growth or shrinkage.
2. Here, we develop a flexible approach to modelling skewed growth data using a modified beta regression model. We propose that sizes first be converted to a (0,1) interval by estimating size-dependent minimum and maximum sizes through quantile regression. Transformed data can then be modelled using beta regression with widely available statistical tools. We demonstrate the utility of this approach using demographic data for a long-lived plant, gorgonians and an epiphytic lichen. Specifically, we compare inferences of population parameters from discrete PPMs to those from IPMs that either assume normality or incorporate skew using beta regression or, alternatively, a skewed normal model.
3. The beta and skewed normal distributions accurately capture the mean, variance and skew of real growth distributions. Incorporating skewed growth into IPMs decreases population growth and estimated life span relative to IPMs that assume normally distributed growth, and more closely approximate the parameters of PPMs that do not assume a particular growth distribution. A bounded distribution, such as the beta, also avoids the eviction problem caused by predicting some growth outside the modelled size range.
4. Incorporating biologically relevant skew in growth data has important consequences for inference from IPMs. The approaches we outline here are flexible and easy to implement with existing statistical tools.

## KEYWORDS

asymmetry, beta regression, growth, integral projection model, matrix model, quantile regression, skew, skewed normal

## 1 | INTRODUCTION

In most organisms, rates of survival, growth and reproduction are dependent on individual size. Size-structured population models, which account for size effects on individual performance, are widely used in analyses of population dynamics, conservation biology and life-history patterns. A critical component of size-structured models is the set of growth transitions governing how individuals change in size over time.

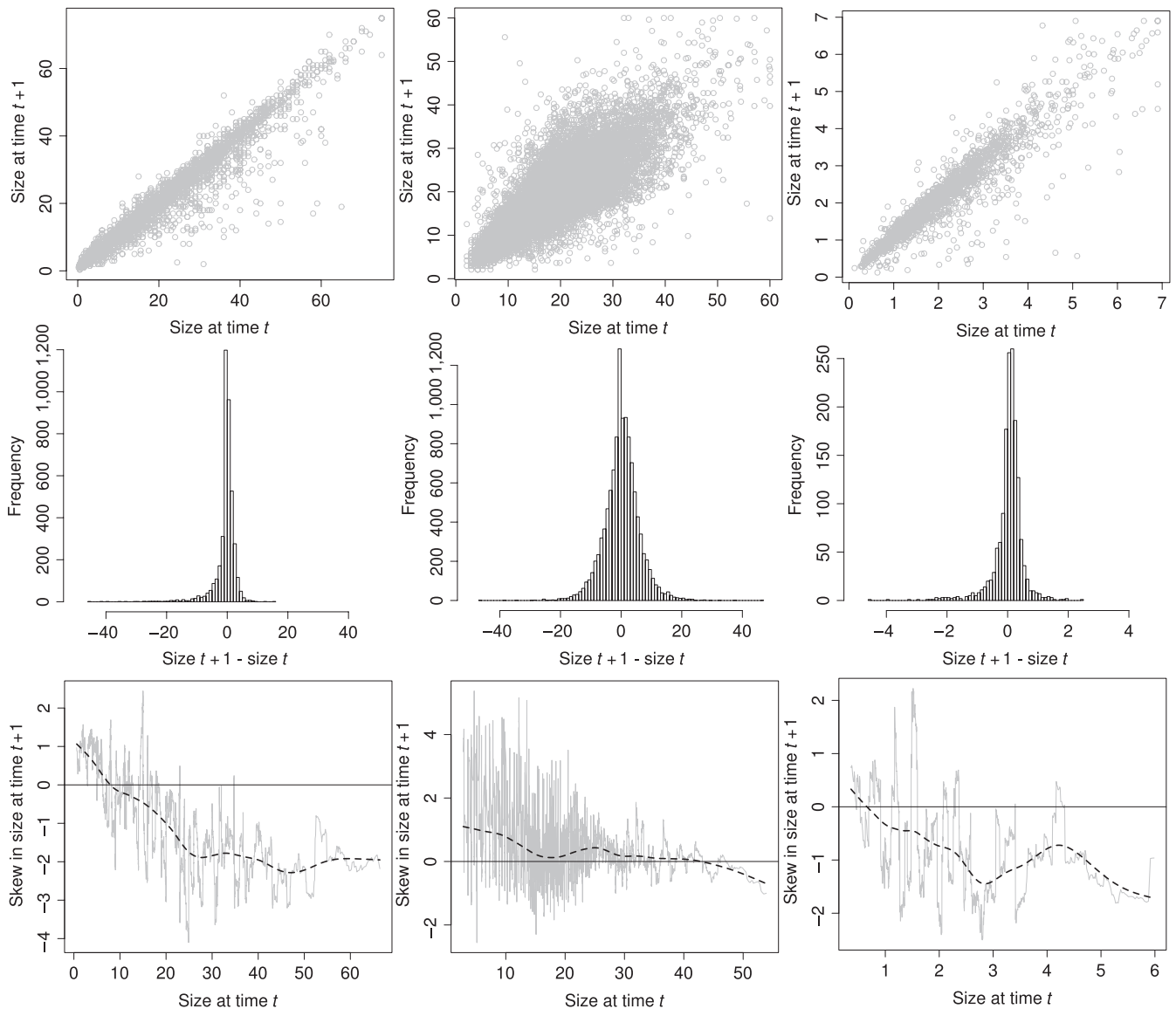
When parameterizing classic population projection matrix models (PPMs; Caswell, 2001), the probabilities of individuals in a given size class shrinking or growing to join other size classes over a time step (often a year) has often been estimated directly from the observed frequencies of size transitions in the data. However, developers of PPMs have increasingly employed continuous estimation of size-based vital rates to estimate demographic rates (Doak & Morris, 2010; Gross, Morris, Wolosin, & Doak, 2005). In addition, PPMs have progressively been supplanted by integral projection models (IPMs), a class of stage-structured models that represent survival, reproduction and growth as continuous functions of state variables, frequently including size (Easterling, Ellner, & Dixon, 2000; Ellner & Rees, 2006). IPMs can be constructed using a diverse array of continuous functions, and several authors have emphasized the need when building IPMs to critically compare multiple flexible models and distributions, including splines or generalized additive models (GAMs), that can describe a wide range of relationships between growth, survival or reproduction and size (Dahlgren, García, & Ehrlén, 2011; Ellner, Childs, & Rees, 2016; Rees, Childs, & Ellner, 2014). In practice, however, almost every study of which we are aware has assumed a normally distributed growth process on the scale at which size is modelled, with the probabilities of transitioning by a given amount above or below the average size next year being the same (but see Montero-Serra et al., 2017; Needham, Merow, Chang-Yang, Caswell, & McMahon, 2018; Shriver, Cutler, & Doak, 2012). This nearly ubiquitous assumption appears to be due to the lack of clear guidance about alternative approaches, rather than any biological reason to assume this would be the case. Yet, the consequences for inference of population dynamics of how growth has been represented in most IPMs (assuming normality) versus PPMs (using the observed distribution), or some other alternative, has not been investigated.

Many species will violate the assumption of normally distributed size transitions. While there are multiple aspects of growth distributions that could be substantially non-normal, even with appropriate transformations of the state variable, two stand out. First are the bounds on possible shrinkage or growth. Unbounded distributions, such as the normal, will predict non-zero probabilities of reaching sizes that exceed any biologically plausible outcome. For example,

small individuals cannot have negative sizes on an absolute (e.g. not log-transformed) scale and also are often unable to reach larger sizes for a species in a single transition, while large individuals cannot grow to sizes outside some biological upper bound and may also be unable to shrink to typical small sizes for a species without dying. The problem of unbounded distributions predicting sizes outside the size limits of an IPM (i.e. 'eviction') is well-recognized (Williams, Miller, & Ellner, 2012) and usually dealt with by truncating the growth distribution; however, this approach does not deal with unrealistic predictions within the overall size envelope for a species.

A second violation of normality is the inherent skew in distributions of possible size transitions, expressed as probabilities of being different sizes after a time step. For example, organisms, particularly large ones, often shrink much more than they are able to grow in a single time step, resulting in negatively skewed growth rates (Figure 1). Shrinkage due to die-back, breakage or starvation has been documented in a wide range of organisms (Linares, Doak, Coma, Díaz, & Zabala, 2007; Montero-Serra et al., 2017; Wikelski & Thom, 2000) and can have profound effects on population dynamics (Salguero-Gómez & Casper, 2010). Previous studies have either included such shrinkage within the estimation of normal growth distributions or modelled shrinkage as a distinct process. For example, Shriver et al. (2012) estimated a probability of extreme shrinkage and separate size distributions for 'normally growing or shrinking' versus 'extreme shrinkage' individuals of the lichen *Vulpicida pinastri*, an approach also taken by Montero-Serra et al. (2017) to model die-back of a slow-growing precious red coral (*Corallium rubrum*). Conversely, some organisms such as trees may be incapable of meaningful shrinkage, resulting in positively skewed growth rates. Recently, Needham et al. (2018) used a mixture of two gamma distributions to model the positively skewed growth of slow-growing trees along with the more normally distributed growth of fast-growing trees. However, these mixed-distribution solutions require a somewhat ad hoc approach in distinguishing normal growth from skewed growth that is not likely to be generally applicable.

Here, we propose a flexible approach to modelling skewed and bounded growth in continuous population models using a modified implementation of the beta-distributed regression model. The beta distribution is a flexible, continuous distribution that can be symmetric or skewed (Figure S1), but in its usual form requires data to be bounded between 0 and 1, although so-called four-parameter beta distributions can be used to convert variables with other distributional limits to (0,1) for analysis (Wang, 2007). Here, we describe a workflow to convert size data to a (0,1) interval based on size-dependent minimum and maximum thresholds, and then the use of beta regression methods to estimate the effects of size and other independent variables on the mean, variance and skew of the distribution. We first describe



**FIGURE 1** Many species exhibit skewed size distributions. We show three examples from extensive demographic studies: (Left) a long-lived gorgonian coral (*Paramuricea clavata*), (Centre) an arctic/alpine plant (alpine bistort, *Polygonum viviparum*), and (Right) an epiphytic boreal lichen (*Vulpicida pinastri*). Top panels show the relationship between size at time  $t + 1$  and size at time  $t$ . Middle panels show the distribution of growth increments (size  $t + 1 -$  size  $t$ ), centered on zero. Bottom panels show the skew in size at time  $t + 1$  (grey lines), computed for a sliding window of 50 data points, against the mean size at time  $t$ . Solid black lines indicate a symmetric distribution with 0 skew; dashed black lines show a smoothing spline fit to the sliding skew estimates. See Supporting Methods for details of the demographic datasets

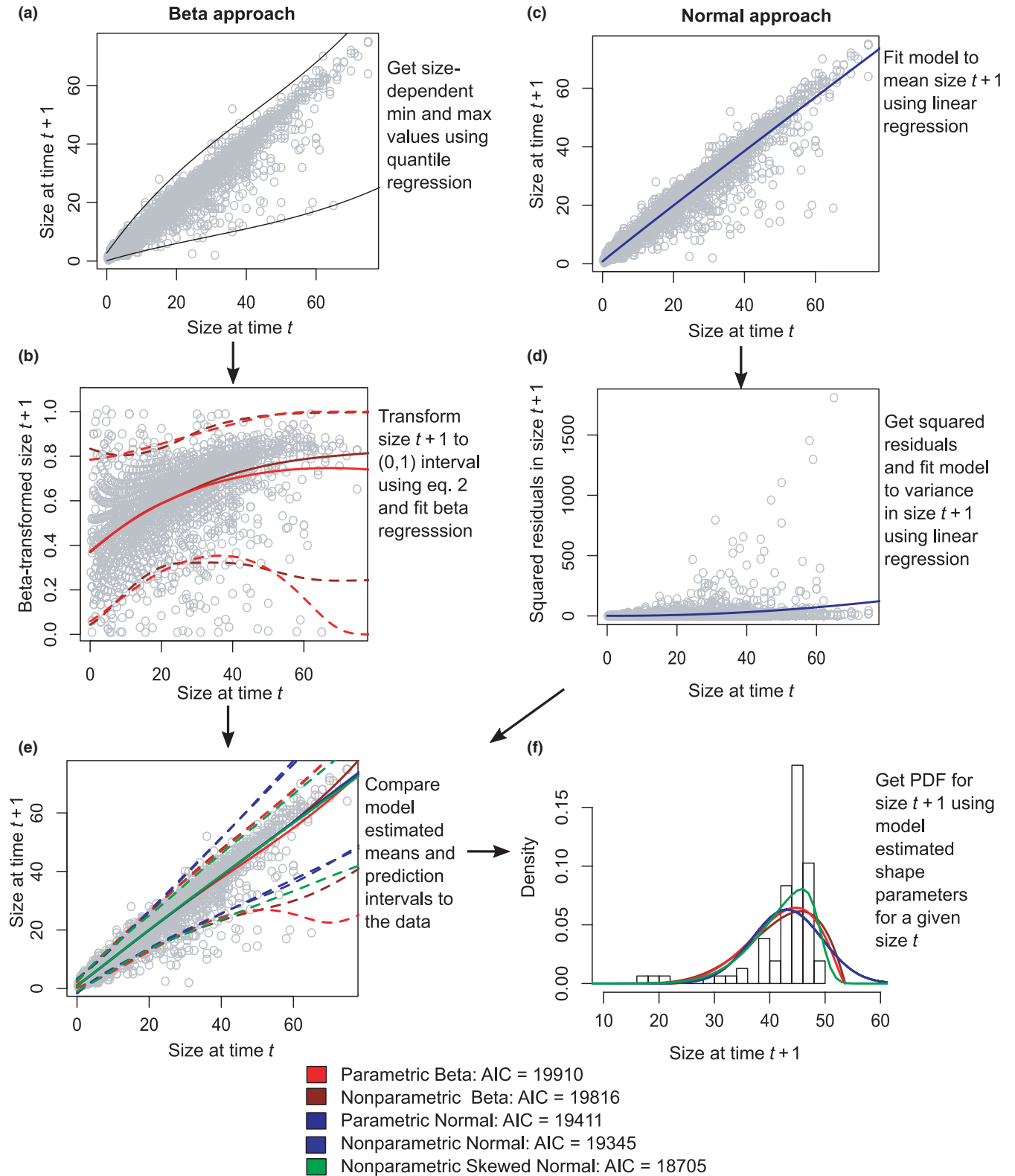
the basic approach using growth data for the red gorgonian coral (*Paramuricea clavata*, hereafter 'coral'), which grows slowly but can sometimes shrink dramatically due to breakage, mainly caused by human divers, and warming-induced mortality (Figure 1) (Linares & Doak, 2010; Linares et al., 2007). We then contrast results of our method (hereafter 'beta approach') with those from the standard method assuming normally distributed growth rates (hereafter 'normal approach') as well as an alternative approach based on the skewed normal distribution, which is able to model skewed, but still unbounded, growth and shrinkage. Finally, we compare the impact of assuming normally distributed growth versus relaxing this assumption for three population outputs – population growth

rate, life span and reactivity – using data from three diverse and well-studied species: the red gorgonian coral, an epiphytic lichen and a perennial arctic/alpine plant.

## 2 | BACKGROUND AND METHODS

### 2.1 | The normal approach

The standard approach to modelling growth in an IPM is to fit two separate models to predict the mean and variance of the size distribution at the end of a time step as a function of starting size and any other influencing variables (Easterling et al., 2000). In this process,



the first model is generally a regression of size at time  $t + 1$  on size at time  $t$  (and other independent variables, such as climate, habitat or herbivory) to estimate the mean ending size as a function of starting size (Figure 2c). The second model takes the squared residuals from the first model as point estimates of the variance in size  $t + 1$ , again as a function of starting size plus other variables (Figure 2d). There are

a variety of approaches to flexibly model how the mean and/or variance may change with size, including simultaneously estimating both parameters by maximum likelihood (reviewed in Ellner et al., 2016; Rees et al., 2014), but the most common approach is to use separate linear regressions for both parameters (e.g. Metcalf, McMahon, Salguero-Gómez, Jongejans, & Rees, 2013). Many implementations of

**FIGURE 2** Comparison of the workflows using the beta versus normal approach to model skewed growth data in red gorgonian coral (*Paramuricea clavata*). (a) Quantile regression estimates minimum and maximum size values in the next time step as functions of current size. Lines show the 0.1 and 99.9 percentiles of size. (b) The beta regression fitted median (solid lines) and 99% prediction intervals (dashed lines) fit to size  $t + 1$  transformed to a (0,1) interval using Equation 2. (c) The linear regression fit (solid line) of the mean size  $t + 1$ . (d) The linear regression fit (solid line) of the squared residuals in size  $t + 1$ . (e) Lines give the fitted medians (solid) and 99% prediction intervals (dashed) for beta (red), normal (blue) and skewed normal (green) models on the original data scale. (f) The estimated beta (red) and skewed normal (green) distributions better capture the left skewness of the observed size distribution compared to the estimated normal distributions (blue), as illustrated here by predictions for individuals with size 45 at time  $t$  against data for individuals with sizes 44–46. Note that predictions and AIC are given for both parametric and nonparametric beta and normal models, as well as a nonparametric skewed normal model. R code used for each of these steps and to generate figures is given in Appendix S1

this approach use Akaike information criterion (AIC) or similar criteria to judge the support for quadratic size effects or the influence of climate or other variables on the mean and variance of growth rates. For a given starting size, the estimated mean and variance are then used to compute the normal probability density function (pdf) for ending size, which is used to compute the probability of transitioning into a given size bin at the next time step in IPMs. Typically, this is done using the point estimate of the pdf for a given starting size multiplied by the bin width (i.e. the midpoint rule; Ellner & Rees, 2006; Merow et al., 2014; Metcalf et al., 2013); however, we have found this approach to be less stable (M. Peterson and D. Doak, pers. obs.) so here we integrate the pdf across size bins by taking the difference in the cumulative density function of each bin edge. Absolute size measurements are often log-transformed to improve normality of size and other vital rate relationships, but this approach can only address positive skew in growth whereas many species may show negative skew or shifts in skew from positive to negative with increasing size.

This method works well for size transitions that are reasonably approximated by a normal distribution. However, although the normal distribution is able to capture the mean and variance of skewed distributions, it will badly misestimate the mode, resulting in a distribution that will under- or overestimate the probability of reaching certain sizes depending on the direction of skew (Figure S2). In addition, the normal distribution, or any other unbounded distribution, will predict non-zero (albeit generally small) probabilities of growth and shrinkage to unrealistic sizes. This latter problem has been recognized when predicting sizes outside any seen in a study (Williams et al., 2012), and is usually dealt with by truncating the pdf for each starting size to be within the modelled size range, typically by either renormalizing growth probabilities within the modelled size range to sum to 1, or by assigning the probability density falling outside the modelled size range to the most extreme size class within the range. These approaches typically alter the pdfs for the smallest and largest individuals most strongly. However, unbounded distributions are also problematic when they predict an unreasonable subsequent size given the starting size of an individual (e.g. allowing a small fraction of individuals to make unrealistic jumps in size), and this problem may be more difficult to diagnose and correct. As we argue in the Introduction, both these issues are likely to be common. An additional problem that arises when regressing squared residuals against size is that the best-fit regression model may predict negative variances over some size ranges, which is dealt with in various ad hoc ways in the

literature. For example, it is common for authors to set negative estimates to small positive values or use zero-intercept or exponential models for variance estimation (Ellner & Rees, 2006).

## 2.2 | The beta approach

To address these problems, we developed a modified beta regression approach that provides a highly flexible alternative for modelling growth rates. The beta distribution describes continuous data on a (0,1) interval, and has been most commonly used in the analysis of proportions. Note that some implementations of beta regression are for beta-binomial data, with successes and failures as discrete events and the probability of a success being chosen from a continuous distribution, but the beta distribution itself is continuous. It is straightforward to convert any continuous variable  $y$  to a (0,1) interval when there is a minimum and maximum value, as:

$$y' = (y - y_{\min}) / (y_{\max} - y_{\min}), \quad (1)$$

where  $y'$  is the transformed data. Previous implementations of beta regression with transformed data have used a constant minimum and maximum value to describe processes that represent a proportion of some whole, such as test scores or concentrations of environmental contaminants (Gray & Alava, 2018; Kim & Wolt, 2011; Smithson & Verkuilen, 2006).

Setting a constant minimum and maximum value for possible sizes irrespective of starting size could be used to model growth, but does not work well in most cases, largely because it enforces unrealistic patterns of skew into the model fit (M. Peterson, pers. obs.). Instead, to apply this method to size data, we suggest allowing the minimum and maximum values to vary as functions of starting size ( $y_t$ ), transforming each value of size at time  $t + 1$  ( $y_{t+1}$ ) thus:

$$y'_{t+1} = (y_{t+1} - y_{\min|y_t}) / (y_{\max|y_t} - y_{\min|y_t}), \quad (2)$$

where  $y_{\max|y_t}$  and  $y_{\min|y_t}$  are the maximum and minimum sizes at time  $t + 1$  conditional on starting size  $y_t$ . This transformation serves two purposes. First, for most organisms, it is more biologically realistic to allow minimum and maximum potential sizes to be size-dependent. Second, the beta distribution for the transformed sizes is only symmetric when the mean is equal to  $(y_{\max|y_t} - y_{\min|y_t}) / 2$ , or 0.5 on the (0,1) interval; it is positively skewed with smaller mean values and negatively skewed with larger mean values (Figure S1). Using a

constant maximum and minimum value, as with Equation 1, would enforce a flip from positive to negative skew at  $(y_{\max} - y_{\min})/2$ , whereas allowing the maximum and minimum values to change with starting size greatly increases the flexibility to model variable skew, including symmetry, in the predicted distributions of sizes.

Ideally, we would know the theoretical bounds on the sizes that an individual of a given starting size could achieve in one time step. However, most often, these bounds will have to be estimated from the data. Here, we use quantile regression to estimate the 0.1 and 99.9 quantiles as functions of starting size (Figure 2a), which we use as the minimum and maximum values when transforming each observation. We found that using the 0.1 and 99.9 quantiles produced better estimates than using more extreme values (e.g. the 0 and 1 quantiles), which are highly subject to the influence of outliers. However, this approach will leave a small fraction of individuals outside the (0,1) interval, which can then be excluded as outliers, adjusted to values just within this range (e.g. 0.01 or 0.99), or accommodated by adjusting the minimum and maximum values slightly to create bounds that are wider than any observed values. In many cases, outliers in growth and shrinkage are likely to be the result of measurement errors, so tighter quantiles could represent more reasonable estimates of the minimum and maximum boundaries. Here, we found that cubic quantile regressions for these bounds performed well, but there are also even more flexible spline-based methods available (Koenker, 2009) if the minimum and maximum values vary as complex functions of starting size, as well as methods to incorporate random effects into quantile models (Geraci, 2014). Using size-dependent minimum and maximum values produces a transformed dataset in the (0,1) interval with mean values that are not strongly tied to starting size (Figure 2b).

Once the size data have been suitably transformed, existing methods can be used to fit models with beta-distributed errors. In particular, Ferrari and Cribari-Neto (2004) suggested a maximum-likelihood model based on an alternative parameterization of the beta distribution that uses mean and precision ( $\phi$ ) instead of the alpha and beta shape parameters. This model was extended by Simas, Barreto-Souza, and Rocha (2010) to allow both the mean and precision parameters to vary with different sets of independent variables. In these models, the precision parameter is:

$$\phi = \left( \mu(1 - \mu) - \sigma^2 \right) / \sigma^2, \quad (3)$$

where  $\mu$  is the mean and  $\sigma^2$  is the variance. Since  $\mu(1 - \mu)$  is the maximum variance,  $\phi$  is a measure of the reduction in variance below the theoretical maximum, given the mean. Importantly for its implementation, the mean and precision parameters are estimated simultaneously using maximum likelihood. These methods can be implemented using the BETAREG package (Cribari-Neto & Zeileis, 2010) or the GAMLSS package (Stasinopoulos & Rigby, 2007) in R (R Core Development Team, 2015) and models can be compared with AIC. Once the best-supported model is found, it is straightforward to convert the estimated mean and precision for any given starting size into the alpha and beta shape parameters to compute the beta pdf, as:

$$\alpha_{y'_{t+1}|y_t} = \phi_{y'_{t+1}|y_t} \mu_{y'_{t+1}|y_t} \quad (4)$$

$$\beta_{y'_{t+1}|y_t} = \phi_{y'_{t+1}|y_t} \left( 1 - \mu_{y'_{t+1}|y_t} \right), \quad (5)$$

where  $\mu_{y'_{t+1}|y_t}$  is the mean and  $\phi_{y'_{t+1}|y_t}$  is the precision of the fitted beta distribution for  $y'_{t+1}$  conditional on starting size  $y_t$ . The beta pdf is used just like the normal pdf to get the probability of transitioning into any given size bin in the IPM. It is also easy to back-transform the model predictions onto the original data scale to use them in a demographic model or to visualize the model fit (Figure 2e), as:

$$\mu_{y_{t+1}|y_t} = \left( \mu_{y'_{t+1}|y_t} (y_{\max|y_t} - y_{\min|y_t}) \right) + y_{\min|y_t} \quad (6)$$

$$\sigma^2_{y_{t+1}|y_t} = \sigma^2_{y'_{t+1}|y_t} (y_{\max|y_t} - y_{\min|y_t})^2, \quad (7)$$

where  $\mu_{y'_{t+1}|y_t}$  and  $\sigma^2_{y'_{t+1}|y_t}$  are the mean and variance of  $y'_{t+1}$ , conditional on starting size  $y_t$ .

### 2.3 | The skewed normal approach

An alternative approach to modelling skewed growth is to use the skewed normal distribution (Azzalini, 1985). This is a simpler solution to represent skew, but does not address the problem of modelling appropriately bounded growth and shrinkage rates. The skewed normal is a three-parameter distribution that is equivalent to the normal distribution when the skewness parameter is 0.

### 2.4 | Illustration of different methods

We compared the fit of each approach using previously published data on the growth of corals (*Paramuricea clavata*) (Linares et al., 2007). For this test, we used data on 4,110 growth events (Appendix S1). This slow-growing octocoral occasionally suffers severe breakage due to diving activity and warming-induced 'partial mortality' (death of part but not all of a colony), leading to highly skewed growth rates (Figure 1).

We used three approaches to model growth of this species: the beta approach, the normal approach and the skewed normal approach. We modelled the size-dependent mean and precision of the beta distribution, the mean and variance of the normal distribution, and the location, scale and skewness parameters of the skewed normal distribution. For each of the three approaches, we fit sets of models that considered each of the parameters as constant, linear, quadratic or nonparametric cubic spline functions of size and used AIC to identify the best-supported models. We fit parametric beta models using the BETAREG package (Cribari-Neto & Zeileis, 2010), parametric normal models using separate linear regressions for the mean and variance with the lm function, and all other models (nonparametric beta, nonparametric normal and all skewed normal models) using the GAMLSS package (Stasinopoulos & Rigby, 2007), which fits the functions for all parameters simultaneously using maximum penalized likelihood. For the beta approach, we also used quantile regression to estimate the 0.1 and

99.9 quantiles of size as linear, quadratic or cubic functions of starting size using the `QUANTREG` package (Koenker, 2009). We then used these values, from the best-supported model based on AIC, as the minimum and maximum values, respectively, to transform size  $t + 1$  to a (0,1) interval, setting any values  $\leq 0$  to 0.01 and any values  $\geq 1$  to 0.99.

Finally, we compared the fit of each approach to the data with AIC and visualized the predicted means and 99% prediction intervals on the original scale. We calculated AIC values using the total number of parameters in all models for a given approach (e.g. the quantile and beta models for the beta approach, the mean and variance models for the parametric normal approach or the mean, variance and skew functions for the skewed normal approach). We also compared the predicted pdfs of each approach to the observed distribution of size  $t + 1$  for individuals across a range of starting sizes. Example code for each of these steps in `R` v.3.4.3 is given in Appendix S1.

## 2.5 | Comparison of each method's fit to real growth distributions

We next tested the relative performance of the beta, normal and skewed normal approaches using previously published demographic data for three biologically diverse species with varying degrees of skewness in growth rates (Figure 1). Alpine bistort (*Polygonum viviparum*, hereafter 'bistort') is a long-lived arctic/alpine perennial plant for which demographic data were collected annually from 2001 to 2011 in four populations at Niwot Ridge, CO (Doak & Morris, 2010). Individual plants can sometimes shrink dramatically due to herbivory or frost damage. As described above, the red gorgonian (*Paramuricea clavata*) is a slow-growing and long-lived arborescent octocoral. Demographic data for individual colonies were collected annually at three sites for 3–5 years each from 1999 to 2004 (Linares & Doak, 2010; Linares et al., 2007). *Vulpicida pinastri* (hereafter 'lichen') is a short-lived epiphytic lichen that was studied on bush alder (*Alnus* sp.) stems. Data for this species were collected annually from 2004 to 2009 in Kennicott Valley, AK. Extreme shrinkage in this species occurs due to mechanical damage from falling branches or mammal activity (Shriver et al., 2012).

We tested the ability of each approach to accurately capture the mean, variance and skewness of these real growth datasets. In fitting the models to each dataset, we followed the steps outlined above, considering both parametric and nonparametric beta, normal and skewed normal models and using AIC to identify the best-supported models for each approach. For each of the three species, we fit models to 200 bootstrap replicates obtained by randomly resampling the data with replacement to generate a distribution of estimates. We compared the estimated size-dependent means, variances and skews predicted from the best-supported models to the moments of each dataset, estimated for a sliding window of 50 data points and fit with a smoothing spline. Code for these analyses with gorgonian coral is given in Appendix S2.

## 2.6 | Consequences for population model inferences

Failing to account for skewness in growth may have substantial consequences for the inferences drawn from population models.

When large changes in size tend to be in a downward direction, then the assumption of normally distributed size transitions traditionally used in IPMs will overestimate the probability of high growth (Figure S2). Since survival and reproductive rates often increase with size, these biased estimates of high growth could inflate estimates of the population growth rate by allowing some individuals to reach large sizes more quickly than is biologically realistic. Similarly, if extreme shrinkage is more common than extreme growth (even if still rare), then the normal distribution will underestimate its probability (Figure S2), potentially inflating predictions of the average life span.

We explored the potential consequences of failing to account for skewness in growth data using previously published demographic data for the same three species as above. For each of these species, we constructed IPMs using either the beta, normal or skewed normal approaches outlined above, and compared their outputs to those from a high-dimension PPM that uses the observed size transitions directly (i.e. without fitting a statistical model to growth, see Supporting Methods for details). While estimates of population behaviour from the PPM are not necessarily correct, they do represent a model-free estimate that is not subject to distributional assumptions. We fit PPMs with 50 size classes and estimated transition rates based on their observed frequencies in the dataset, pooling across years, populations and plots to ensure sufficient sample sizes in each class. Population growth rate, life span and damping ratio values converge for PPMs with 20 or more classes (Figure S3), suggesting that a  $50 \times 50$  PPM is of sufficiently high dimension to yield accurate results. IPMs were fit to the same pooled datasets using either the beta, normal or skewed normal approaches to modelling growth, and keeping all other size-dependent vital rates constant. We renormalized all growth predictions to prevent eviction. IPMs were discretized using the midpoint rule (Morris & Doak, 2002) into matrices with  $>100$  size classes ( $N = 118\text{--}125$ ) chosen to evenly divide the 50 size classes used for PPMs. We compared three metrics of population dynamics inferred from each IPM or PPM: the population growth rate ( $\lambda$ ), the life span (defined here as the time steps until a newborn has  $<1\%$  probability of still being alive) and the reactivity (Neubert & Caswell, 1997; Stott, Townley, & Hodgson, 2011). Reactivity is inherently tied to the definition of size classes or bins, so to compare values from the IPMs to the PPM we used a population vector giving the proportional abundance of each of the size classes in the IPM represented by the largest size class in the PPM (see Supporting Methods). We obtained distributions of each metric by refitting all models across 200 bootstrap replicates obtained by randomly resampling the data for each species with replacement. Code for each of these steps is given in Appendices S2–S4.

## 2.7 | Dissecting the fit of each growth model to different segments of the growth distribution

If extreme shrinkage or other skewed growth transitions are sufficiently rare (e.g. Figure 1), standard statistical measures of model

fit, such as AIC, may support a normal distribution over a beta or skewed normal distribution if the normal is a better fit to most of the data, especially since the beta approach requires fitting functions for the minimum and maximum bounds as well as the mean and variance in growth. However, failing to capture rare but biologically important transitions, such as extreme shrinkage, could bias estimates of population growth even if it has relatively little effect on the likelihood of the individual growth models. To further explore this issue, we first compared the AIC of the beta, normal and skewed normal growth models described previously, when fit to the growth data for each of the three species. We also compared the log likelihood of each data point for the three approaches to determine where in the growth distribution each approach provided a better fit. Finally, we examined the impacts for population models of better or worse fits to different segments of the growth distribution. For each transition from a size bin at time  $t$  to a size bin at time  $t + 1$ , we calculated the difference in its probability as predicted by each of the three growth models and its sensitivity (i.e. effect on population growth). We compared these values for each observation to its log likelihood under each of the three growth models, to ask whether differences in fit were associated with large changes in predicted growth probabilities and/or transitions with high sensitivity.

### 3 | RESULTS

#### 3.1 | Illustration of different methods with growth of red gorgonian coral

Methods that allow skew, including the beta approach and the skewed normal, are better able to capture the asymmetry in growth of gorgonian corals relative to the symmetric prediction intervals given by linear regressions of the mean and normal variance in growth (Figure 2e). In particular, the assumption of symmetry causes the normal variance model to overestimate the probability of high growth and underestimate the probability of extreme shrinkage. This is most obvious when comparing the predicted pdf for a given starting size (Figure 2f). The parametric and nonparametric normal models produced very similar predictions, whereas the nonparametric beta model differed from and outperformed the parametric beta model. AIC strongly supported the skewed normal model over either the normal or beta approaches.

#### 3.2 | Comparison of each method's fit to real growth distributions

All three approaches closely approximated the mean of the observed growth distributions, although the normal approach tended to underestimate the mean relative to either the beta or skewed normal models (Figure 3). Estimates of the variance were more variable between approaches. For corals, the beta and skewed normal most closely approximated the variance, although all approaches underestimated the variance at large sizes. For bistort, the beta approach

tended to overestimate the variance. For lichen, the variance was best captured by the skewed normal at smaller sizes and by the beta approach at larger sizes, while the normal approach always underestimated the variance. The methods also differed strongly in their estimates of the skew. The beta approach most closely captured the high positive skew at small sizes and the low negative skew at large sizes for all three species, although the skewed normal more closely matched the skew for the bulk of the data points at intermediate sizes for coral and lichen. The normal approach assumes 0 skew.

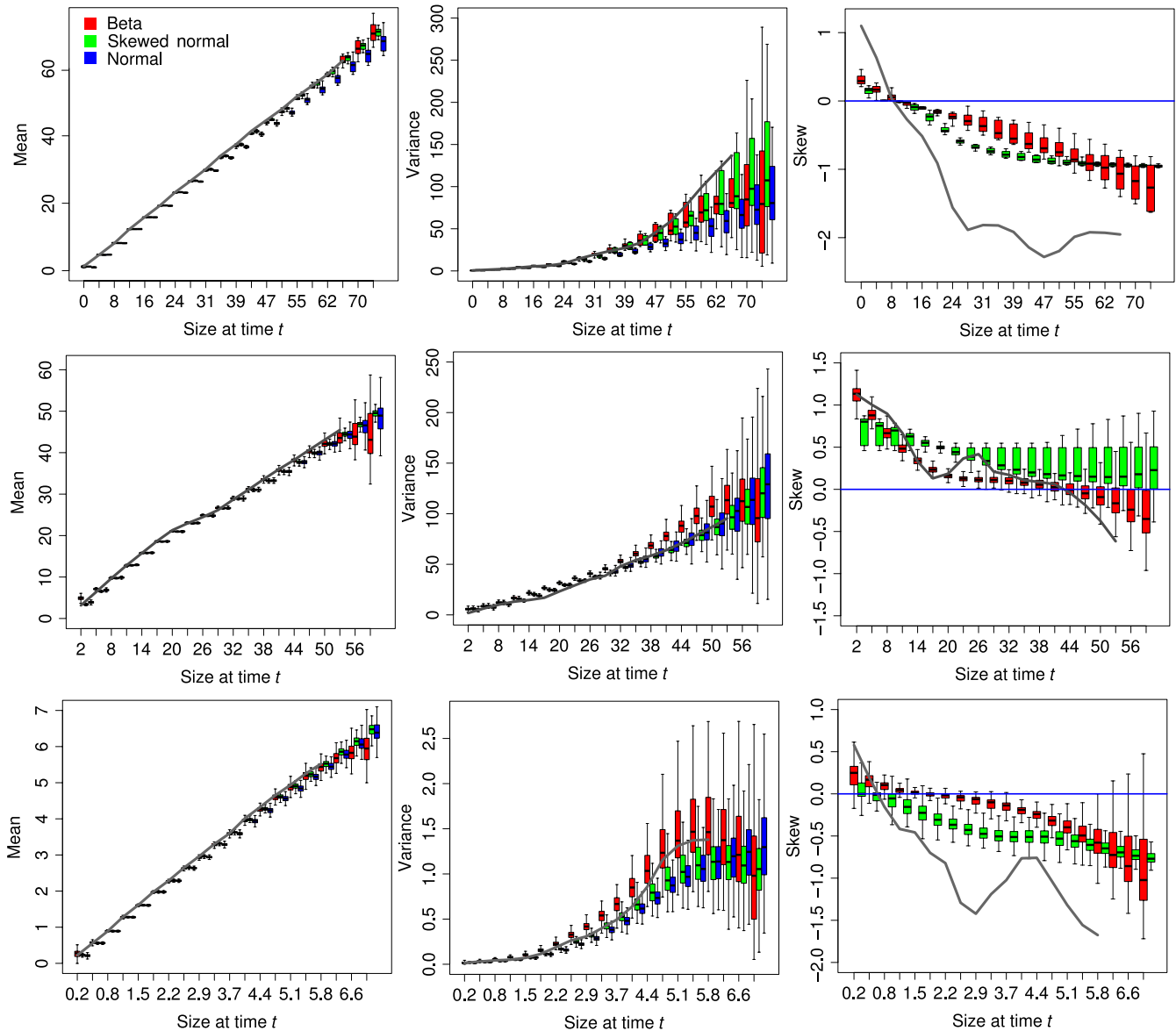
#### 3.3 | Consequences for population model inferences

We found that assuming growth was normally distributed altered inferred population dynamics and life-history patterns relative to models using either a beta or skewed normal distribution to accommodate skewed growth rates (Figure 4). For all three species, IPMs based on normally distributed growth models overestimated the population growth rate and life span relative to PPMs that do not rely on a statistical growth model but instead use the observed transition frequencies in the data. IPMs based on either beta or skewed normal models of growth yielded very similar predictions of population growth for both coral and lichen, and very similar predictions of life span for coral. However, IPMs based on the beta approach more closely approximated the results from PPMs for population growth of bistort and for life span of both bistort and lichen. There were no clear differences among approaches for the reactivity, except that IPMs tended to underestimate the reactivity compared to PPMs for both coral and lichen.

#### 3.4 | Dissecting the fit of each growth model to different segments of the growth distribution

For the three datasets examined here, we found no support for the beta approach over either the normal or skewed normal approaches as measured by AIC of growth model fits. Instead, AIC strongly supported the skewed normal model for all three species. This was largely due to both the normal and skewed normal models providing slightly better fits to the majority of data points close to the mean growth rate, as well as to extreme outliers in growth or shrinkage (Figures 5 and S4). In spite of these results for overall model fit, the beta approach provided better fits to the spread of individuals around the mean, better approximating the bands of moderate negative or positive skew in growth. It is not surprising that unbounded distributions such as the normal and skewed normal would best predict the extreme size outliers relative to a bounded distribution such as the beta. However, despite the strong log-likelihood support for the normal and skewed normal models at these data points, the probability of these size transitions tended to be similarly predicted across all models, suggesting little effect on model inference. Conversely, the data points for which the beta approach had the strongest log-likelihood support tended to be for size transitions with high sensitivity and/or substantial differences in the predictions between the two approaches. These results indicate that for species with realistic levels of skewness in growth rates, the assumption of normality in growth can have wide-ranging impacts on the inference





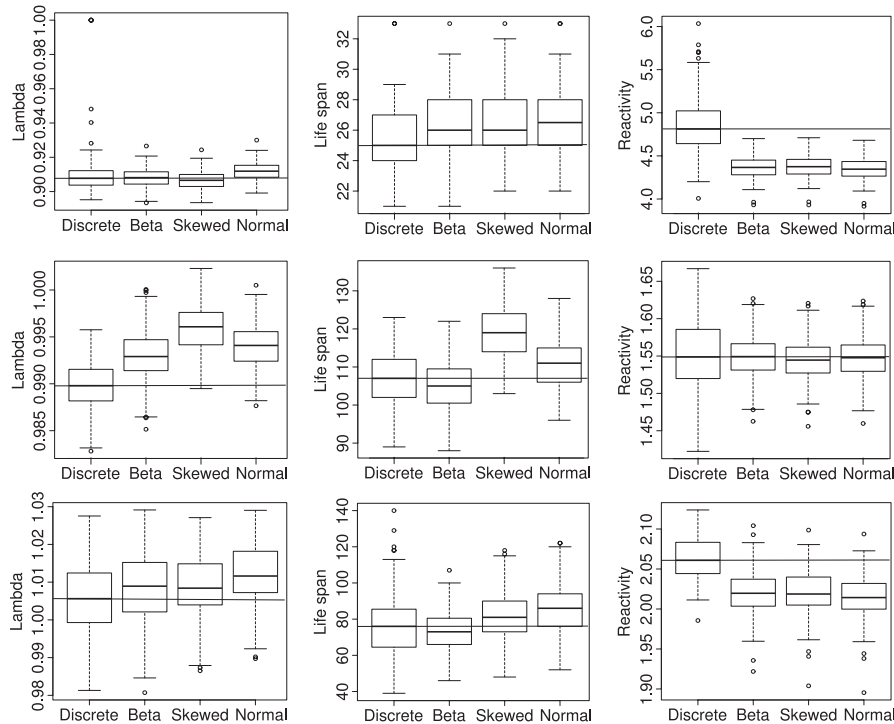
**FIGURE 3** Comparison of the ability of the beta, skewed normal and normal distributions to capture the moments of the distribution of size at time  $t + 1$ . Box plots give the median (black lines), interquartile range (boxes) and the most extreme data points within 1.5 times the interquartile range (whiskers) of the distribution of estimated moments. (Left) mean, (Centre) variance and (Right) skew of size at time  $t + 1$  for various starting sizes at time  $t$  for (Top) coral (*Paramuricea clavata*), (Centre) bistort (*Polygonum viviparum*) and (Bottom) lichen (*Vulpicida pinastri*). Estimates are from the best-supported models using the beta (red), skewed normal (green) or normal (blue) distributions, fit to 200 bootstrapped replicates drawn with replacement from the original datasets. Grey lines show the true moments of the original dataset, shown as a smoothing spline fit to moments calculated from a sliding window of 50 data points. Note that the normal distribution assumes 0 skew (blue line in right panels)

about dynamics and life-history traits that are commonly of interest to population biologists.

#### 4 | DISCUSSION

In many organisms, the distribution of growth from one size to the next will be skewed even after transformations meant to improve normality. We have outlined a highly flexible approach for modelling skewed growth rates by combining two widely available statistical

tools: quantile regression and beta regression. We demonstrate that both this approach and a rarely utilized alternative, the skewed normal, are able to more accurately capture the moments of real growth distributions for three biologically diverse species. Furthermore, we show that ignoring skew by assuming normally distributed growth, as is done in most published IPMs, can lead to overestimation of population growth rates and life span. Incorporating skew in growth transitions can be easily accomplished with existing statistical tools and is important for accurately capturing population dynamics, life-history attributes and population structure.



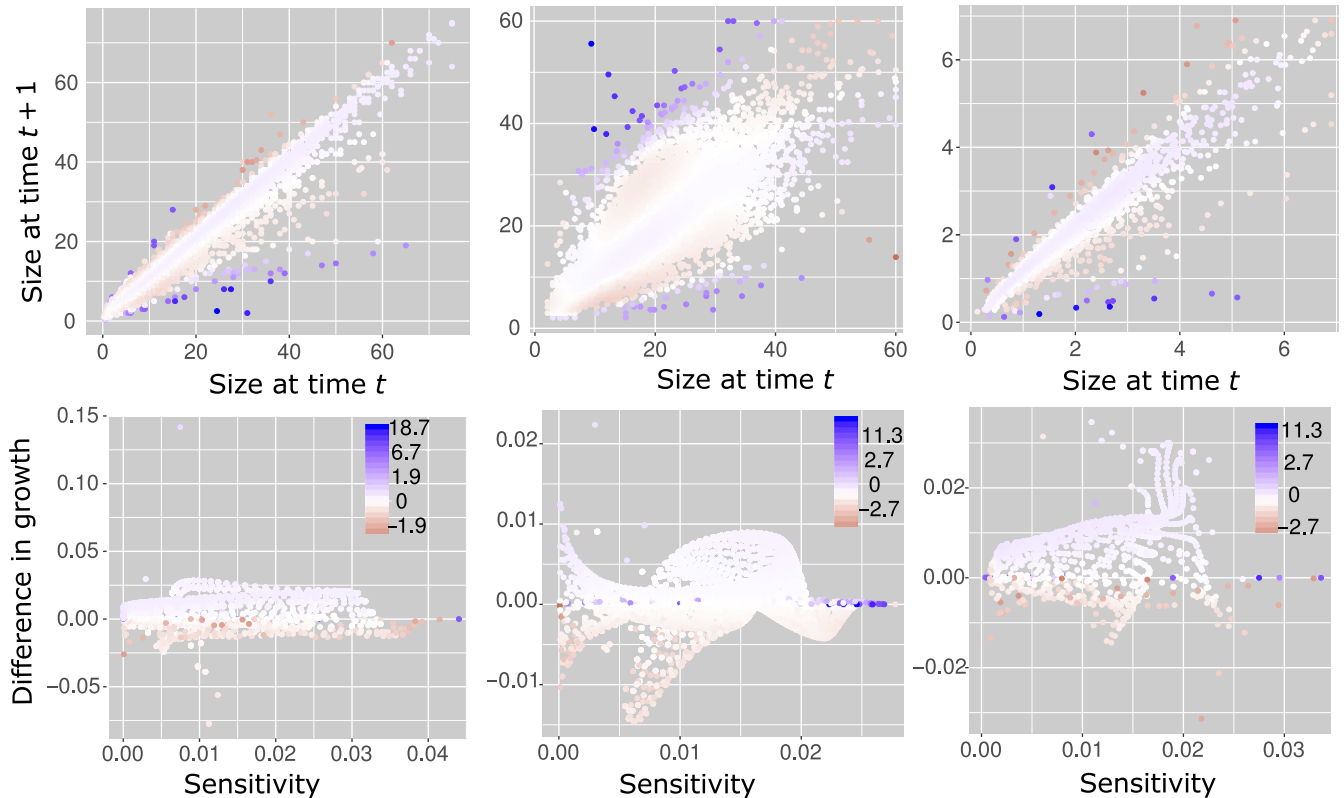
**FIGURE 4** Comparison of population parameters inferred from a high-dimension PPM (Discrete) to those from IPMs assuming beta, skewed normal or normal-distributed growth. Values are (Left) population growth rate  $\lambda$ , (Centre) life span or (Right) reactivity estimated from PPMs or IPMs fit to 200 bootstrap replicates drawn with replacement from the original datasets for (Top) coral (*Paramuricea clavata*), (Centre) bistort (*Polygonum viviparum*) or Bottom) lichen (*Vulpicida pinastri*). The IPMs were fit with the best-supported models based on either a beta, skewed normal or normal distribution, and all other vital rate models were identical. The PPMs estimated discrete vital rates for each of 50 size bins (see Supporting Methods for details). Box plots give the median (black lines), interquartile range (boxes) and the most extreme data points within 1.5 times the interquartile range (whiskers) of the distribution of each parameter. Horizontal lines show the median value from the PPMs for comparison. Abbreviation: IPMs, integral projection models; PPMs, projection matrix models

We have shown that IPMs built with normally distributed versus skewed growth transitions produce different conclusions about key population parameters, but which approach is more correct? There are two ways in which this could be assessed. The first is to compare the statistical fit of each model to the data and to test whether the additional parameters required by either the beta approach or a skewed normal model are justified using AIC (see Appendix S1 for an example). Across the three datasets here, AIC strongly supports the skewed normal, but not the beta approach, relative to normal models. However, there are several considerations when using AIC to evaluate growth models for use in IPMs. The first is whether a given growth model has been fit by simultaneously estimating all parameters by maximum likelihood, or whether it combines parameters fit separately by multiple models. For example, the beta approach is fit by separately estimating the minimum and maximum sizes with quantile regression followed by estimation of the mean and precision; the AIC for such a model will be an upper bound compared to a joint estimation approach. The same is true of the normal approach when fit with separate regressions for the mean and variance. While it is straightforward to estimate the normal mean and variance simultaneously with maximum likelihood, this is more difficult to accomplish for the beta approach we outline here. Although it is possible to extend the four-parameter beta distribution likelihood function

to enable size-dependent values of the mean, precision, minimum and maximum to be estimated simultaneously (Wang, 2007), it is not straightforward to implement as the minimum and maximum values must be constrained within some biologically plausible range. However, a first exploratory attempt to do this does suggest that it can improve the likelihood of the entire model to the point where beta models outperform normal models (see Figure S5 and Appendix S3 for example).

A second consideration is that, in practice, unbounded growth distributions are truncated for use in IPMs to avoid prediction outside the modelled size range. The two common solutions to this problem – adding probability density that falls outside the size bounds to the most extreme size class or renormalizing the truncated probabilities to sum to 1 – could either increase or decrease the likelihood of an unbounded model, depending on the fit of the truncated distribution to the data.

The final and perhaps most important consideration is that statistical measures of fit such as AIC weight the prediction of all data points (and equivalently, all parts of the predicted growth distributions) equally. However, errors in different portions of the growth distribution can have very different effects on the population growth rate and other parameters of interest from IPMs. For example, likelihood-based measures may support models that best



**FIGURE 5** Fit of the beta versus skewed normal models to different segments of the growth distribution. Top row: size at time  $t + 1$  against size at time  $t$ , with each data point coloured by the difference in its log-likelihood under the skewed normal approach versus the beta approach (bluer values indicate greater log likelihood of the skewed normal model and redder values indicate greater log likelihood of the beta model). In all three cases, greater statistical support for the skewed normal stems from its ability to better predict extreme outliers, as well as to slightly better predict the majority of data points close to the mean growth, whereas the beta approach is a better fit to the bands of moderate negative or positive skew relative to the mean. Bottom row: For each observed growth transition, the difference in its predicted probability from the skewed normal versus beta approach (positive values indicate greater probability under the skewed normal model and negative values indicate greater probability under the beta model) against the sensitivity of lambda to that transition. Points are coloured by the difference in their log-likelihood, as in the top row. The strongest log-likelihood support for the skewed normal approach (bluer values) comes from data points for which the methods produce similar predictions, whereas the beta approach has log-likelihood support (redder values) for data points that are predicted to have very different probabilities under the two methods. From left to right, columns show results for coral (*Paramuricea clavata*), bistort (*Polygonum viviparum*) and lichen (*Vulpicida pinastri*)

explain the majority of the data points, potentially at the expense of capturing rarer but biologically important transitions, such as extreme shrinkage. Indeed, we found that the beta approach better explained the majority of skewed size transitions, to which population growth is sensitive, relative to either the normal or skewed normal models, suggesting that the beta approach may better capture the aspects of growth that most strongly impact population dynamics.

Another method to assess which approach is more accurate is to compare the outputs from each IPM to those from a high-dimension PPM that uses the observed size transitions directly. Although a PPM is also an approximation of the ‘true’ population dynamics, there are several aspects of this study that makes the PPM outputs a reasonable benchmark for comparison. First, we used large demographic datasets ( $N = 1,621$ – $11,882$  individual transitions) to minimize the potential for substantial influences of sampling errors in the observed transition frequencies and, second, we used a bootstrapping approach to compare model outputs across many

randomly sampled datasets. Under these conditions and in the absence of independent knowledge, we can roughly assume that IPMs that more closely approximate the results of a high-dimension PPM are also more accurately representing the underlying population. By this measure, IPMs assuming normally distributed growth consistently overestimated population growth rates and life spans, whereas IPMs using the beta approach performed equivalently or better than those using a skewed normal model of growth. Thus, the beta models’ predictions of growth better match behaviour inferred without recourse to assumptions about the distributional properties of the vital rates. This suggests that the beta approach better captures population-level outputs when growth transitions are skewed, even if it is not supported by likelihood-based measures of model fit such as AIC.

Structured population models are used to address a wide range of ecological and evolutionary questions, including quantifying population growth and extinction risk (Crouse, Crowder, & Caswell, 1987;

Lande, 1988), describing basic aspects of species' life history (Franco & Silvertown, 1996), and predicting transient dynamics following disturbance or management interventions (Ezard et al., 2010; Stott et al., 2011). Here, we show that incorporating versus ignoring skew in growth rates has strong and consistent effects on estimates of population growth and life span, with implications for both basic inference as well as management decisions. For example, extreme shrinkage in the coral we examine here is largely caused by diving activity and warming-induced die-back. Importantly for species like this, assuming normally distributed growth will tend to underestimate the severity of human impacts, with potentially severe consequences for management. Interestingly, we did not find consistent effects of skewed growth on transient dynamics, as measured by the reactivity. However, the impact on transient dynamics could vary depending on the life history of a particular organism, and should be tested across a wider range of species.

The main statistical cost of the beta approach relative to either the normal or skewed normal approaches is the estimation of minimum and maximum size bounds. While this requires additional parameter estimates, we argue that this statistical cost is simply accounting for a very real aspect of the demographic process: the existence of biological bounds on potential size transitions. However, identifying the best size bounds is likely to be a compromise between statistical fit and biological realism; as our example in Appendix S3 shows, widening the size bounds can often improve the likelihood and AIC of the model, but will also allow some non-zero probability of unobserved – and unlikely to occur – growth transitions (Figure S5). For this reason, we have proposed using quantile regression to estimate size-dependent minimum and maximum values that are close to the observed range of growth transitions. However, it is important to note that this approach will always yield size ranges that are slightly less than or equal to those observed in the data, and therefore could slightly shrink the variance when back-transformed to the original data scale. Although we did not find this to be an issue in our datasets, with the fitted variance closely matching the observed variance, it is possible that this could be more problematic for sparser or noisier datasets for which estimating size-dependent minima and maxima would be more difficult. For this reason, we recommend comparing the fitted moments, especially the variance and skew, to those observed in the data, and potentially adjusting the minimum and maximum values slightly if necessary.

Although we argue that a bounded distribution is biologically preferable for modelling growth, the skewed normal model could be a good alternative in some cases, such as for sparser datasets, as it does not require estimates of the minimum or maximum sizes. We found that IPMs based on the skewed normal distribution performed as well as the beta approach in some cases, although it overestimated population growth in one species and overestimated life span in two species. The relative performance of the beta approach versus skewed normal may depend on the details of a particular dataset. Interestingly, we observed the biggest differences in bistorts, for which we had the most data ( $N = 11,882$  individual transitions). Since both approaches are easy to implement with existing statistical tools, it may often be preferable to simply fit both growth models to compare their results.

Finally, as with any other approach to continuous vital rate estimation, it is important to carefully consider the fit of models to the data. This is even more critical for growth because the entire distribution of size transitions, not merely the mean, is used to parameterize size-structured population models. When assessing model fit, particular attention should be paid to the tails of the data range, where data are often sparse, as well as to rare but biologically important transitions, such as extreme growth or shrinkage. Several population ecologists have advocated the use of more flexible spline or GAM-based approaches to modelling vital rates (Dahlgren et al., 2011; Ellner et al., 2016), and this can be extended to the beta and skewed normal approaches we outline here. For example, the `GAMLSS` package in R enables both the mean and dispersion of beta-distributed data, as well as all three parameters of the skewed normal distribution, to be modelled as nonparametric functions of explanatory variables, such as cubic splines (Stasinopoulos & Rigby, 2007). Across all three species, nonparametric models were supported over parametric models, indicating that nonparametric models may be better able to capture shifts in the location, scale and skewness of growth distributions relative to parametric approaches. Finally, it is important to note that other aspects of growth distributions, such as kurtosis, will also violate assumptions of normality and can be addressed with several other distributions (e.g. the  $t$  distribution; Rees et al., 2014). A wide range of distributions can be easily fit by maximum likelihood or with the `GAMLSS` package, yet the normal distribution remains a pervasive default for modelling growth (Merow et al., 2014; Metcalf et al., 2013). Size transformations, such as taking the log, are often assumed to yield normally distributed growth. However, this can still leave substantial skew or other aspects of non-normality, as shown by the three species we examine here, all of which were transformed to improve the normality of size prior to our analyses but still produced biased IPM outputs under assumptions of normally distributed growth. For these reasons, we reiterate prior calls (e.g. Easterling et al., 2000; Ellner et al., 2016; Rees et al., 2014) to carefully compare multiple alternative growth distributions for a given dataset.

Many organisms will have patterns of growth that violate assumptions of normality. In such cases, the beta distribution offers a flexible alternative that can be easily implemented with existing statistical software and incorporated into IPMs. Given the need for accuracy in the predictions of population models, we suggest that testing of alternative distributions for growth, such as the beta approach we outline here, become standard in the construction of size-structured population models.

## ACKNOWLEDGEMENTS

This work was supported by NSF DEB 1242355, 1242558, 1753954, 1753980, 1340024. We thank Steve Ellner and an anonymous reviewer for helpful comments that greatly improved the manuscript.

## AUTHORS' CONTRIBUTIONS

M.L.P., D.D. and W.M. conceived the ideas. D.D., W.M. and C.L. collected the data. M.L.P. performed the analyses and wrote the first

draft of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

## DATA ACCESSIBILITY

Data: Dryad Digital Repository <https://doi.org/10.5061/dryad.t6c3573> (Peterson, Morris, Linares, & Doak, 2019). R scripts: uploaded as Supporting Information and archived in Dryad Digital Repository <https://doi.org/10.5061/dryad.t6c3573>.

## ORCID

Megan L. Peterson  <https://orcid.org/0000-0002-5010-2721>

## REFERENCES

- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*, 171–178.
- Caswell, H. (2001). *Matrix population models: construction, analysis, and interpretation*. Sunderland, Massachusetts: Sinauer Associates.
- Cribari-Neto, F., & Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software*, *34*, 1–24.
- Crouse, D. T., Crowder, L. B., & Caswell, H. (1987). A stage-based population model for Loggerhead sea turtles and implications for conservation. *Ecology*, *68*, 1412–1423. <https://doi.org/10.2307/1939225>
- Dahlgren, J. P., Garcia, M. B., & Ehrlén, J. (2011). Nonlinear relationships between vital rates and state variables in demographic models. *Ecology*, *2011*, 5. <https://doi.org/10.1890/10-1184.1>
- Doak, D. F., & Morris, W. F. (2010). Demographic compensation and tipping points in climate-induced range shifts. *Nature*, *467*, 959–962. <https://doi.org/10.1038/nature09439>
- Easterling, M. R., Ellner, S. P., & Dixon, P. M. (2000). Size-specific sensitivity: applying a new structured population model. *Ecology*, *81*, 694–708. [https://doi.org/10.1890/0012-9658\(2000\)081\[0694:SSAAN\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[0694:SSAAN]2.0.CO;2)
- Ellner, S. P., Childs, D. Z., & Rees, M. (2016). *Data-driven modelling of structured populations: A practical guide to the integral projection model*. Cham, Switzerland: Springer International Publishing.
- Ellner, S. P., & Rees, M. (2006). Integral projection models for species with complex demography. *American Naturalist*, *167*, 410–428. <https://doi.org/10.1086/499438>
- Ezard, T. H. G., Bullock, J. M., Dalgleish, H. J., Millon, A., Pelletier, F., Ozgul, A., & Koons, D. N. (2010). Matrix models for a changeable world: The importance of transient dynamics in population management. *Journal of Applied Ecology*, *47*, 515–523. <https://doi.org/10.1111/j.1365-2664.2010.01801.x>
- Ferrari, S., & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, *31*, 799–815. <https://doi.org/10.1080/0266476042000214501>
- Franco, M., & Silvertown, J. (1996). Life history variation in plants: An exploration of the fast-slow continuum hypothesis. *Philosophical Transactions of the Royal Society B*, *351*, 1341–1348.
- Geraci, M. (2014). Linear quantile mixed models: The lqmm package for Laplace quantile regression. *Journal of Statistical Software*, *57*, 1–29.
- Gray, L., & Alava, M. H. (2018). A command for fitting mixture regression models for bounded dependent variables using the beta distribution. *The Stata Journal*, *18*, 51–75. <https://doi.org/10.1177/1536867X1801800105>
- Gross, K., Morris, W. F., Wolosin, M. S., & Doak, D. F. (2005). Modeling vital rates improves estimation of population projection matrices. *Population Ecology*, *48*, 79–89. <https://doi.org/10.1007/s10144-005-0238-8>
- Kim, M., & Wolt, J. D. (2011). Probabilistic risk assessment of dietary cadmium in the South Korean population. *Food Additives & Contaminants. Part A, Chemistry, Analysis, Control, Exposure & Risk Assessment*, *28*, 62–70. <https://doi.org/10.1080/19440049.2010.529620>
- Koenker, R. (2009). *quantreg: Quantile regression*. Retrieved from <http://CRAN.R-project.org/package=quantreg>
- Lande, R. (1988). Demographic models of the northern spotted owl (*Strix occidentalis caurina*). *Oecologia*, *75*, 601–607. <https://doi.org/10.1007/BF00776426>
- Linares, C., & Doak, D. F. (2010). Forecasting the combined effects of disparate disturbances on the persistence of long-lived gorgonians: A case study of *Paramuricea clavata*. *Marine Ecology Progress Series*, *402*, 59–68. <https://doi.org/10.3354/meps08437>
- Linares, C., Doak, D., Coma, R., Díaz, D., & Zabala, M. (2007). Life history and viability of a long-lived marine invertebrate: The octocoral *Paramuricea clavata*. *Ecology*, *88*, 918–928.
- Merow, C., Dahlgren, J. P., Metcalf, C. J. E., Childs, D. Z., Evans, M. E. K., Jongejans, E., ... Ramula, S. (2014). Advancing population ecology with integral projection models: A practical guide. *Methods in Ecology and Evolution*, *5*, 99–110. <https://doi.org/10.1111/2041-210X.12146>
- Metcalf, C. J. E., McMahon, S. M., Salguero-Gómez, R., Jongejans, E., & Rees, M. (2013). IPMPack: An R package for integral projection models. *Methods in Ecology and Evolution*, *4*, 195–200.
- Montero-Serra, I., Garrabou, J., Doak, D. F., Figuerola, L., Hereu, B., Ledoux, J.-B., & Linares, C. (2017). Accounting for life-history strategies and timescales in marine restoration. *Conservation Letters*, *11*, e12341. <https://doi.org/10.1111/conl.12341>
- Morris, W. F., & Doak, D. F. (2002). *Quantitative conservation biology*. Sunderland, Massachusetts, USA: Sinauer.
- Needham, J., Merow, C., Chang-Yang, C. H., Caswell, H., & McMahon, S. M. (2018). Inferring forest fate from demographic data: From vital rates to population dynamic models. *Proceedings of the Royal Society B: Biological Sciences*, *285*, 20172050. <https://doi.org/10.1098/rspb.2017.2050>
- Neubert, M. G., & Caswell, H. (1997). Alternatives to resilience for measuring the responses of ecological systems to perturbations. *Ecology*, *78*, 653–665. [https://doi.org/10.1890/0012-9658\(1997\)078\[0653:ATRFMT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1997)078[0653:ATRFMT]2.0.CO;2)
- Peterson, M. L., Morris, W. F., Linares, C., & Doak, D. F. (2019). Data from: Improving structured population models with more realistic representations of non-normal growth. *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.t6c3573>
- R Core Development Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rees, M., Childs, D. Z., & Ellner, S. P. (2014). Building integral projection models: A user's guide. *Journal of Animal Ecology*, *83*, 528–545. <https://doi.org/10.1111/1365-2656.12178>
- Salguero-Gómez, R., & Casper, B. B. (2010). Keeping plant shrinkage in the demographic loop. *Journal of Ecology*, *98*, 312–323. <https://doi.org/10.1111/j.1365-2745.2009.01616.x>
- Shriver, R. K., Cutler, K., & Doak, D. F. (2012). Comparative demography of an epiphytic lichen: Support for general life history patterns and solutions to common problems in demographic parameter estimation. *Oecologia*, *170*, 137–146. <https://doi.org/10.1007/s00442-012-2301-4>
- Simas, A. B., Barreto-Souza, W., & Rocha, A. V. (2010). Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis*, *54*, 348–366. <https://doi.org/10.1016/j.csda.2009.08.017>
- Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, *11*, 54–71. <https://doi.org/10.1037/1082-989X.11.1.54>

- Stasinopoulos, D., & Rigby, R. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, 23, 1–46.
- Stott, I., Townley, S., & Hodgson, D. J. (2011). A framework for studying transient dynamics of population projection matrix models. *Ecology Letters*, 14, 959–970. <https://doi.org/10.1111/j.1461-0248.2011.01659.x>
- Wang, J. Z. (2007). A note on estimation in the four-parameter beta distribution. *Communications in Statistics - Simulation and Computation*, 34, 495–501. <https://doi.org/10.1081/SAC-200068514>
- Wikelski, M., & Thom, C. (2000). Marine iguanas shrink to survive El Niño. *Nature*, 403, 37–38.
- Williams, J. L., Miller, T. E., & Ellner, S. P. (2012). Avoiding unintentional eviction from integral projection models. *Ecology*, 93, 2008–2014. <https://doi.org/10.1890/11-2147.1>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Peterson ML, Morris W, Linares C, Doak D. Improving structured population models with more realistic representations of non-normal growth. *Methods Ecol Evol.* 2019;10:1431–1444. <https://doi.org/10.1111/2041-210X.13240>