

My, how you've grown:

A practical guide to modeling size transitions
for Integral Projection Model (IPM) applications

Tom E.X. Miller^{*a} and Stephen P. Ellner^b

^aDepartment of BioSciences, Rice University, Houston TX

^bDepartment of Ecology & Evolutionary Biology, Cornell University, Ithaca NY

Submitted to: *Ecology* (Statistical Report)

Keywords: demography; growth; integral projection model; kurtosis; skewness

Open Research Statement: Data are already published and publicly available, with those items properly cited in this submission. Three data sets are cited as data packages: (Miller, 2020) <https://doi.org/10.6073/pasta/dd06df3f950afe4a4642306182237d13>, (Ochocki *et al.*, 2023) <https://doi.org/10.6073/pasta/ca53c16f16dcf9fb11f3ee99ea5445ac>, (Winfield *et al.*, 2013b) <https://doi.org/10.5285/637d60d6-1571-49af-93f7-24c1279d884d>. Two other data sets are available in our Github repo, which also includes all of our code. The repo will be archived in a Zenodo package, with the DOI included at the end of the article, upon publication. During peer review, our data and code are available at https://github.com/texmiller/IPM_size_transitions.

^{*}Corresponding author. Department of BioSciences, Rice University, Houston, TX 77005-1827. Email: tom.miller@rice.edu
Phone: 713-348-4218

1 **Abstract**

2 Integral Projection Models (IPMs) are widely used for studying continuously size-structured populations.
3 IPMs require a growth sub-model that describes the probability of future size conditional on current size and
4 any covariates. Most IPM studies assume that this distribution is Gaussian, despite calls for non-Gaussian
5 models that accommodate skewness and excess kurtosis. We provide a general workflow for accommodating
6 non-Gaussian growth patterns while retaining important covariates and random effects. Our approach
7 emphasizes visual diagnostics from pilot Gaussian models and quantile-based metrics of skewness and kurtosis
8 that guide selection of a non-Gaussian alternative, if necessary. Across five case studies, skewness and excess
9 kurtosis were common features of growth data and non-Gaussian models consistently generated simulated data
10 that were more consistent with real data than pilot Gaussian models. However, effects of “improved” growth
11 modeling on IPM results were moderate to weak, and differed in direction or magnitude between different
12 outputs from the same model. Using tools not available when IPMs were first developed, it is now possible to fit
13 non-Gaussian models to growth data without sacrificing ecological complexity. Doing so, as guided by careful
14 interrogation of the data, will result in models that better represent the populations for which they are intended.

1 Introduction

Structured demographic models – matrix and integral projection models (MPMs and IPMs) – are powerful tools for data-driven modeling of population and community dynamics. In contrast to MPMs for populations with discrete structure (life stage, age class, etc.), IPMs (Easterling *et al.*, 2000) accommodate populations structured by continuous state variables, most commonly size. A related innovation of the IPM framework is its emphasis on regression-based modeling for parameter estimation, which often carries important advantages for making the most of hard-won data (Ellner *et al.*, 2022).

A standard workflow allows ecologists to assemble an IPM from data using familiar regression tools to describe growth, survival, reproduction, and other demographic transitions as functions of size (Coulson, 2012; Ellner *et al.*, 2016). The relative ease of regression analyses, accommodating covariates (e.g., environmental factors, experimental treatments) and complex variance structures (e.g., random effects, correlated errors), has facilitated a growing IPM literature that examines how biotic or abiotic factors affect population dynamics (e.g., Louthan *et al.*, 2022; Ozgul *et al.*, 2010) and explores the consequences of demographic heterogeneity associated with spatial, temporal, and individual variation (e.g., Compagnoni *et al.*, 2016; Crone, 2016; Plard *et al.*, 2018). The vital rate regressions (or “sub-models”) are the bridge between the individual-level data and the population-level model and its predictions; it is important to get those right.

Compared to other vital rates, growth is special. The survival and reproduction sub-models only need to provide a single predicted value as functions of size (we use “size” as the name for whatever continuous variable defines the population structure). But the growth model must specify the full probability distribution of subsequent size conditional on initial size, defining the growth ‘kernel’ $G(z', z)$ that gives the probability density of future size z' at time $t + 1$ conditional on current size z at time t . Whenever survival and reproduction are size-dependent, the entire distribution of size transitions can strongly influence IPM predictions because it governs how frequently size changes are much greater or much lower than average.

Easterling *et al.* 2000 provided the original template for modeling size transitions in IPMs. They first tried simple linear regression, assuming Normally distributed size changes with constant variance. Because the residuals from this regression exhibited non-constant variance, they used a two-step approach to estimate the size-dependence in mean squared residuals (better options soon became available, such as the `lme` function in R). However, even after accounting for non-constant variance, growth data may still be non-Normal. Size transitions are often skewed such that large decreases are more common than large increases (Peterson *et al.*, 2019; Salguero-Gómez & Casper, 2010), or vice versa (Stubberud *et al.*, 2019). Size transitions may also exhibit excess kurtosis (“fat tails”), where extreme growth or shrinkage is more common than predicted by the tails of the Normal distribution (Héault *et al.*, 2011).

The observation that the Normal (or Gaussian) distribution may poorly describe size transitions in real organisms has been made before, and several studies have emphasized that alternative distributions should be explored (Easterling *et al.*, 2000; Peterson *et al.*, 2019; Rees *et al.*, 2014; Williams *et al.*, 2012). For example, Peterson *et al.* 2019 showed that skewness in size transitions could be modeled through beta regression on transformed data (for reasons we describe below, this approach also has some drawbacks), or by fitting a skewed Normal distribution. They showed that incorporating skew could have important consequences for model-based inferences, and concluded that “testing of alternative distributions for growth... [should] become standard in the construction of size-structured population models.” Nonetheless, default use of Gaussian growth distributions (often with non-constant variance) remains the standard practice. The general state-of-the-art in the literature appears to remain where it was 20 or so years ago, using the default Gaussian model without examining critically whether or not it actually describes the data well. We are guilty of this, ourselves.

The persistence of Gaussian growth models is understandable. Popular packages such as `lme4` (Bates *et al.*, 2015), `mgcv` (Wood, 2017), and `MCMCglmm` (Hadfield *et al.*, 2010) make it easy to fit growth models with potentially complex fixed- and random-effect structures, but the possible distributions of continuous responses are limited, and default to Gaussian. Abandoning these convenient tools for the

sake of more flexible growth modeling means, it may seem, sacrificing the flexibility to model diverse sources of demographic variation, some of which may be the motivation driving the study in the first place.

Our goal here is to present and illustrate a practical “recipe” that moves growth modeling past the standards set over 20 years ago. Using software tools that are now readily accessible, ecologists can escape the apparent trade-off between realistically modeling non-Gaussian size transitions and flexibly including multiple covariates and random effects.¹ As with any recipe, users may need to make substitutions or add ingredients to suit their needs. We emphasize graphical diagnostics for developing and evaluating growth models, rather than a process centered on statistical tests or model selection. Through empirical case studies we demonstrate how tools that were nonexistent or not readily available when IPMs first came into use now make it straightforward and relatively easy to identify when the default model is a poor fit to the data, and to then choose and fit a better growth model that is no harder to use in practice. We illustrate our approach by revisiting three published case studies (and three additional case studies in Appendix S3), including examples from our own previous work. In each case, the Gaussian assumption does not stand up to close scrutiny. We illustrate how we could have done better, and the consequences of “doing better” for our ecological inferences. All analyses were carried out in R (R Core Team, 2022) version 4.0 or higher and may be reproduced from publicly available code and data (see *Data Availability Statement*).

2 Flexible growth modeling

The modeling process that we suggest runs as follows (Fig. 1):

1. Fit a “pilot” model assuming a Gaussian distribution, but allowing for non-constant variance. This step is familiar to most IPM users, as it is the start and end of the standard approach. It may include model selection to identify which treatment effects or environmental drivers affect the mean and/or variance of future size. Non-constant variance is often fitted in a two-stage process, first fitting mean growth assuming constant

¹Our statements about software availability are based on what current software reliably delivers in our personal experience, not on what they promise.

variance, then doing a regression relating the squared residuals to initial size or the fitted mean of subsequent size. Fitting mean and variance simultaneously as functions of initial size, as can be done with R packages **mgcv** and **nmle**, is advantageous when possible because incorrectly assuming constant variance can affect model selection for the mean. We illustrate both one-step and two-step approaches in the case studies below.

Allowing non-constant variance removes the need for transforming the data to stabilize growth variance. Transformation may still be useful if it does not create new problems such as making some state-fate relationships highly nonlinear. In particular, log-transformation often reduces or eliminates heteroskedasticity in growth data (Ellner *et al.*, 2016) and also helps avoid eviction at small sizes (Williams *et al.*, 2012).

The fitted mean and variance functions should be checked before going any further. If they are perfectly correct, standardized residuals (residuals scaled by the standard deviation) will have zero mean and unit variance overall, and will exhibit no trends in mean or variance with initial size or fitted mean value. However, estimates of the mean and variance functions are somewhat smoothed because of the inescapable bias-variance tradeoff, so scaled residuals will retain some variation in location and scale. Given enough data, statistical tests will detect that variation. So instead, we take for granted the presence of trends and assess their importance by fitting nonparametric spline regression models for residuals (trend in mean) and absolute residuals (trend in variance) as a function of initial size or fitted value. The mean and variance functions can be accepted if the regression curves for the scaled residuals are nearly flat.

2. Use graphical diagnostics to identify if and how the standardized residuals deviate from Gaussian, and to choose a more appropriate distribution. If the Gaussian growth model is valid, the standardized residuals should be Gaussian with zero skewness or excess kurtosis. Growth data may deviate from this in many ways, and the nature of the deviations can guide the search for a better distribution. Tests such as the D'Agostino test of skewness (D'Agostino, 1970) and the Anscombe-Glynn test of kurtosis (Anscombe & Glynn, 1983) can be used to diagnose whether the standardized residuals, in aggregate, deviate from normality (Komsta &

Novomestky, 2015). However, the aggregate distribution may be misleading if skewness or kurtosis vary with size or other covariates. Skewness changing from positive at small sizes to negative at large sizes might produce zero overall skewness, but really requires a distribution that can allow both positive and negative skew, such as the skewed Normal or Johnson S_U distributions. Alternatively, growth data may exhibit leptokurtosis (in which case the t distribution may be a good choice) or may shift from platykurtosis to leptokurtosis depending on initial size (in which case the power exponential distribution may be a good choice). It is therefore essential to visualize trends in distribution properties with respect to either initial size, or expected future size for models with multiple covariates. Fig. 1 includes guidance on how the skew and kurtosis properties of the standardized residuals suggest options for an appropriate growth distribution. In our case studies we exploit the many distributions in the **gamlss** R package (Stasinopoulos *et al.*, 2007), but other distribution families can be used.

3. Refit the growth model using the chosen distribution. In models with multiple covariates and/or random effects, each potentially affecting several distribution parameters, “refit the model” could entail a massive model selection process to identify the “best” non-Gaussian model. With so many options, model uncertainty may be overwhelming and over-fitting becomes a significant risk even when precautions against it are taken.

We therefore argue for adopting a more modest goal: remedy the defects evident in the standardized residuals of the Gaussian model. This recommendation is based on the finding that parameter estimation using Gaussian regression models is generally robust to deviations from normality of the residuals (Schielzeth *et al.*, 2020). That is, the fitted mean of the Gaussian model (as a function of covariates) is probably a very good approximation for the fitted mean in the corresponding non-Gaussian model — and if it is not, the next step in the modeling process will catch that. The functional forms for skew and kurtosis of the non-Gaussian model can be guided by the qualitative features of the graphical diagnostics (e.g., that skewness switches from positive to negative with increasing size). As we demonstrate below, the mean and standard deviation functions can often be carried over exactly from the pilot Gaussian model.

130 4. Evaluate the final growth model through graphical diagnostics comparing simulated and real

131 **growth data.** A good model will generate simulated data that look like the real data. Again, it is important

132 to inspect the properties of simulated data as a function of initial size, fitted mean, or other covariates

133 rather than examining the aggregate distribution. We again suggest below graphical diagnostics, based

134 mainly on quantiles, that can be used to compare simulated with real growth data. If the simulated data do

135 not correspond well with the real data, alternative or more flexible distribution families should be considered,

136 or more complex functions relating distribution parameters to size and other covariates.

137 3 How should skewness and kurtosis be measured?

138 Non-Gaussian growth modeling requires scrutinizing the skewness and kurtosis of standardized residuals,

139 so measurement of these properties warrants attention. The standard measures are based on the third and

140 fourth central moments, respectively, of the distribution: skewness = m_3/σ^3 , excess kurtosis = $m_4/\sigma^4 - 3$

141 where $m_k = \mathbb{E}(X - \bar{X})^k$ is the k^{th} central moment of a random variable X and σ^2 is the variance (second

142 central moment). A Gaussian distribution has zero skewness and zero excess kurtosis.

143 The standard measures are simple and easy to use, but they have poor sampling properties. Because

144 the measures involve high powers of data values, a few outliers can produce very inaccurate estimates. Figure

145 2 shows a simulated example, where the underlying data are samples of 200 values from a t distribution

146 with 8 degrees of freedom, repeated 5000 times; the true skew is 0, and the true excess kurtosis is 1.5.

147 The distance between the largest and smallest estimates (indicated by the dotted red vertical lines), relative

148 to the distance between the 5th and 95th percentiles, shows the broad extent of extreme values that can

149 occur even with a large sample, especially for kurtosis.

150 We therefore recommend nonparametric (NP) measures of skewness and kurtosis that are based on

151 quantiles and thus are less sensitive to a few extreme values. Let q_α denote the α quantile of a distribution

152 or sample (e.g., $q_{0.05}$ is the 5th percentile). For any $0 < \alpha < 0.5$, a quantile-based measure of skewness

is given by (McGillivray, 1986)

$$\text{NP Skewness} = \frac{q_{\alpha} + q_{1-\alpha} - 2q_{0.5}}{q_{1-\alpha} - q_{\alpha}}. \quad (1)$$

NP Skewness measures the asymmetry between the tails of the distribution above and below the median.

The size of the upper tail can be measured (for any $0 < \alpha < 0.5$) by $\tau_U = q_{1-\alpha} - q_{0.5}$; for $\alpha = 0.05$ this is the

difference between the 95th percentile and the median. The lower tail size is $\tau_L = q_{0.5} - q_{\alpha}$. The definition

above is equivalent to

$$\text{NP Skewness} = \frac{\tau_U - \tau_L}{(\tau_U + \tau_L)}. \quad (2)$$

An NP Skewness of ± 0.2 says that the difference in tail sizes is 20% of their total. The range of possible

values is -1 to 1. Both $\alpha = 0.25$ (sometimes called “Kelly’s skewness”) and $\alpha = 0.1$ (“Bowley’s skewness”)

are common choices. We used $\alpha = 0.1$.

An analogous quantile-based measure of kurtosis (?) is

$$\text{NP Kurtosis} = \frac{q_{1-\alpha} - q_{\alpha}}{q_{0.75} - q_{0.25}}. \quad (3)$$

For $\alpha = 0.05$, NP Kurtosis is the difference between the 95th and 5th percentiles, relative to the interquartile

range. To facilitate interpretation, we scale NP Kurtosis relative to its value for Gaussian distribution,

and subtract 1 so that the value for a Gaussian is zero. We call this “NP Excess Kurtosis”. A value of ± 0.2

means that the tails are on average 20% heavier than those of a Gaussian with the same interquartile range.

We calculate NP Kurtosis using $\alpha = 0.05$, to focus on the tail edges, but again this is somewhat arbitrary.

Figure 2C,D illustrate how, applied to the same simulated samples, the nonparametric measures

produce a smaller fraction of highly inaccurate estimates caused by a few extreme values. Also note that,

in contrast to the moment-based measures, numerically small values of the nonparametric measures (e.g., 0.1

or 0.2) should not be disregarded, because both measures are scaled so that a value of 1 indicates extremely

large departures from a Gaussian distribution.

Using quantile-based measures carries the added value that quantile regression can be used to estimate how they vary with initial size or expected future size. In the examples below, we use the **qgam** package (Fasiolo *et al.*, 2020) to fit spline quantile regression models, which accommodate nonlinear size-dependence in skewness and kurtosis. One risk of spline regression is that fitted quantiles may be excessively “wiggly” without constraints on their complexity; with realistic amounts of data, we can hope to estimate broad trends in distribution shape, but not fine-scale variation. In the examples below, we limit complexity by fitting splines with $k = 4$ basis functions unless otherwise noted. Parametric quantile regression is also an option.

For consistency we also use quantile-based measures of mean and standard deviation when comparing real and simulated data, and use quantile regression to visualize their trends. Specifically, following Wan *et al.* (2014),

$$\text{NP Mean} = \frac{q_{0.25} + q_{0.5} + q_{0.75}}{3}, \quad \text{NP SD} = \frac{q_{0.75} - q_{0.25}}{1.35}. \quad (4)$$

4 Case study: lichen, *Vulpicida pinastri*

We begin with a simple example where current size is the only predictor of future size. Growth data for the epiphytic lichen *Vulpicida pinastri* were analyzed first by Shriver *et al.* 2012 and again by Peterson *et al.* 2019 in their study of skewed growth distributions. We therefore had an *a priori* expectation of deviation from normality. The data set includes 1,542 inter-annual transitions in thallus area (cm^2) observed from 2004 to 2009 in Kennicott Valley, AK. Shriver *et al.* 2012 used a mixture distribution that separated “normal growth or shrinkage” from “extreme shrinkage”. We aimed to fit a single growth model that could realistically accommodate both types of size transition without requiring *ad hoc* decisions about which observations of shrinkage were “extreme” or not.

With initial size as the only predictor, a convenient way to fit a Gaussian model with non-constant variance is the `gam` function in **mgcv** library (Wood, 2017) using the `gauss` family. Following a bit of model selection, we fit the mean and standard deviation of future size as second-order polynomials

of current size², then calculated the scaled residuals from the fitted mean and standard deviation. Here, the first argument to `gam()` is a two-element list that defines the linear predictors for mean and sd:

```
# d is the data frame; t0,t1 are initial & final thallus area, respectively
fitGAU <- gam(list(t1~ t0 + I(t0^2), ~ t0 + I(t0^2)), data=d, family=gaulss())
d$fitted_mean = predict(fitGAU,type="response")[,1]
d$fitted_sd <- 1/predict(fitGAU,type="response")[,2]
d$scaledResids=residuals(fitGAU,type="response")/d$fitted_sd
```

The data and fitted mean and standard deviation are shown in Fig. 3A, and the corresponding diagnostic plots are in Fig. 4A,B. Our diagnostic plots are similar to plots made by R's `plot.lm` function, except that we use spline regression to allow data-driven choice of curve smoothness, and use absolute residuals (rather than their square roots) so that the standard deviation of the regression curve is on the same scale as the residuals. The spline curves are not exactly flat – their standard deviations, given above each panel, are positive – but the trends are much too small to be worth fixing.

Quantile regression on the scaled residuals generates the skewness and kurtosis diagnostics shown in Fig. 3B. As expected based on previous analyses, the graphical analysis of the standardized residuals indicates negative skew, especially at larger sizes (Fig. 3B). We also find positive excess kurtosis for all sizes.

We turned to the Johnson's *S-U* (JSU) distribution for improvement. The JSU is a four-parameter leptokurtic distribution allowing positive or negative skew, with the convenient property that its location and scale parameters `mu` and `sigma` are the mean and standard deviation, respectively, which greatly facilitates the transition from a pilot Gaussian model. JSU is not available in any standard linear or additive modeling packages, to our knowledge. But that is not a barrier because we can write a likelihood function using the `dJSU()` function in the **`gamlss.dist`** package. Following the best-fit Gaussian model, we defined `mu` and `sigma` of the JSU as quadratic polynomials of initial size and, based on Fig. 3B) we define

²`gam()` is most commonly used to fit smooth splines (`s()`) for predictor variables, but it can also fit parametric regressions.

the skewness parameter μ as a linear function of size and kurtosis parameter τ as a positive constant.

The likelihood function therefore has nine parameters to estimate. We fit the model using the **maxLik**

package³ with starting coefficient values for μ and σ based on the pilot Gaussian model:

```
## define function that returns the JSU negative log-likelihood
```

```
LogLikJSU=function(pars){
```

```
  dJSU(t1, mu=pars[1]+pars[2]*t0+pars[3]*t0^2,
```

```
  sigma=exp(pars[4]+pars[5]*t0+pars[6]*t0^2),
```

```
  nu = pars[7]+pars[8]*t0, tau = exp(pars[9]), log=TRUE)
```

```
}
```

```
## starting parameter values
```

```
p0<-c(coef(fitGAU)[1:6],0,0,0)
```

```
## fit with maxlik, adding some noise to starting values
```

```
outJSU=maxLik(logLik=LogLikJSU,start=p0*exp(0.2*rnorm(length(p0))),
```

```
method="BHHH",control=list(iterlim=5000,printLevel=2),finalHessian=FALSE);
```

Simulating data from the fitted JSU model indicates a compelling improvement over the best Gaussian model, not only in skewness and kurtosis (Fig. 5C-D) but also the nonparametric standard deviation (5B).

Note, in these data simulation figures Gaussian and non-Gaussian data are offset by an arbitrary amount to more easily visualize their correspondence to the real data (black lines in Fig. 5).

To understand the practical consequences of improved growth modeling, we assembled the remainder of the lichen IPM following Shriver et al. 2012. The asymptotic population growth rate λ based on Gaussian growth differs from the JSU growth model by about 1% annual population growth (Table 1), in line with results of Peterson et al. 2019. However, even this modest difference can lead to biased estimates of extinction

³We chose **maxLik** because it offers the BHHH optimization method, which works well for non-Gaussian likelihoods in our experience.

risk from the Gaussian model, particularly over longer time horizons (Fig. 6). We also explored differences in other life history metrics (Table 1) using functions from Hernández *et al.* (2024). For example, the JSU growth model predicts values for mean lifespan, mean lifetime reproductive success, and generation time that are 15–25% lower than the Gaussian growth model. In this case study, properly modeling non-normal size transitions – which was easy to do with a few extra lines of code – can influence ecological inferences, at least based on point estimates. However, Table 1 also provides bias-corrected, bootstrapped confidence intervals (Diciccio & Efron, 1996), and these are heavily overlapping between the Gaussian and JSU models for all life history traits, suggesting that effects of “improved” growth modeling are small relative to our uncertainty in model parameters.

One could argue that this example was a convenient “straw man” to disqualify Gaussian growth, because it was recognized by the original and subsequent analysts that size transitions are strongly skewed (Peterson *et al.*, 2019; Shriver *et al.*, 2012). In all remaining case studies, including those in Appendix S3, we re-examine growth data that were modeled as Gaussian in the original published analysis.

5 Case study: tree cholla cactus, *Cylindriopuntia imbricata*

The next case study, focused on the tree cholla cactus *Cylindriopuntia imbricata* at the Sevilleta Long-Term Ecological Research site in central New Mexico, adds a new feature to the simple size-dependent regressions in the previous study: random effects associated with temporal (year) and spatial (plot) environmental heterogeneity. This long-term study was initiated in 2004 and different subsets of the data have been analyzed in various IPM studies, all using Gaussian growth kernels (Compagnoni *et al.*, 2016; Czachura & Miller, 2020; Elderd & Miller, 2016; Miller *et al.*, 2009; Ohm & Miller, 2014). In fact, Elderd and Miller 2016 presented a Gaussian growth model as an example of a well fit growth function, based on an overall distribution of residuals that appeared Gaussian and posterior predictive checks (PPCs) of a Bayesian model that suggested consistency between the real data and data simulated from the fitted model (Fig. 4 in (Elder & Miller, 2016)). While PPCs and the associated “Bayesian P-value” are popular diagnostic tools, they are often too

conservative (Conn *et al.*, 2018; Zhang, 2014), failing to reject marginally bad models even though they are very effective in rejecting terrible models. The choice of discrepancy function (the statistic used to compare real and simulated data) can also be limiting: in our previous work, we used a discrepancy function focused on variance (the sum of squared residuals), creating a blind spot for poor modeling of higher moments.

The data includes 4844 size transition observations from 929 individuals spanning 13 transition years (2004–2018) and 11 spatial replicates (three spatial blocks in years 2004–2008 and eight 30m-by-30m plots in years 2009–2018). The data are provided in Miller (2020). Following previous studies, we quantified size as the natural logarithm of plant volume (cm^3), derived from height and width measurements.

We begin growth modeling, as above, with a generalized additive model with the mean and standard deviation of size in year $t + 1$ modeled as smooth function of size in year t , with random intercepts for year and plot and assuming normally-distributed residuals:

```
# t0 and t1 are initial and final log(volume), respectively
fitGAU <- gam(list(t1 ~ s(t0,k=4) + s(plot,bs="re") + s(year,bs="re"),
                  ~ s(t0,k=6)), data=cactus, family=gaulss())
```

Note that here we fitted the standard deviation function with $k = 6$ basis functions rather than our default of $k = 4$ because, in a preliminary analysis, we found a moderate variance trend in the standardized residuals using $k = 4$, suggesting a need for greater flexibility. With $k = 6$, spline regression detected essentially no trend in the mean of the resulting standardized residuals (Fig. 4C,D).

The growth variance is estimated to peak at small to medium sizes (Fig. 3C). The standardized residuals show clear signals of negative skew and positive excess kurtosis across most of the size distribution, but strongest in the middle (Fig. 3D). We therefore need a distribution family allowing negative skew and positive excess kurtosis, both of which may be negligible at some sizes. We first tried Johnson’s S_U and then the skewed t distributions, which provided some improvements but there were still visible discrepancies

between simulated and real data. We next turned to the SHASH distribution, which allows a greater range of kurtosis for a given amount of skew, and vice versa (?; Appendix S1). This flexibility proved necessary to generate simulated data that compared favorably to the real data, so we proceeded with the SHASH. Conveniently, SHASH is available as an **mgcv** family, allowing for flexible size-dependence in skewness and kurtosis without having to select specific size-dependent functions.

Here, the first argument to `gam()` is now a four-element list specifying the linear predictors for the four parameters of the SHASH distribution.

```
fit_shash <- gam(list(t1 ~ s(t0,k=4) +
  s(plot,bs="re") + s(year_t,bs="re"), # location
  ~ s(t0,k=4), # log-scale
  ~ s(t0,k=4), # skewness
  ~ s(t0,k=4)), # log-kurtosis
  data = cactus, family = shash,optimizer = "efs")
```

Data simulated from the SHASH model compared favorably to the real data (Appendix S4, Fig. S-1). Similar to the lichen case study, we see that correctly modeling skewness and kurtosis improved estimation of the nonparametric mean and standard deviation (Appendix S4, Fig. S-1A,B), yielding a growth model that is truer to the data.

We next explored how improved growth modeling influenced IPM results. The λ values predicted by Gaussian and SHASH growth functions, corresponding to the average plot and year, were nearly identical (Table 1) but we could also leverage structure of the study design to quantify demographic variance associated with temporal and spatial heterogeneity. We used the fitted random effects from the vital rate models to estimate the asymptotic growth rate for each year (λ_t), centered on the average plot, and for each plot (λ_p), centered on the average year. Estimates of λ_t from the Gaussian growth

model were often greater than estimates from the SHASH growth model, particularly in some of the harshest years (Fig. 7A), and therefore the Gaussian model predicted lower temporal variance in fitness ($SD(\lambda_{t(Gaussian)}) = 0.04$, $SD(\lambda_{t(SHASH)}) = 0.048$). Plot-to-plot variation was more similar between the two models ($SD(\lambda_{p(Gaussian)}) = 0.0026$, $SD(\lambda_{p(SHASH)}) = 0.0028$), although spatial variation in fitness was much lower than temporal variation (Fig. 7B). The difference in temporal variance would suggest that Gaussian growth modeling would predict a higher stochastic growth rate λ_S , because temporal variance has a negative effect on λ_S . However, the stochastic growth rate from the Gaussian growth model ($\lambda_S = 0.992$) was nearly identical to that of the SHASH growth model ($\lambda_S = 0.991$). This is likely because temporal fluctuations in vital rates, which is where the SHASH growth model would make a difference, have a weaker influence on λ_S than the temporal fluctuations in size structure that they generate (Compagnoni *et al.*, 2016; Ellis & Crone, 2013). The SHASH and Gaussian growth models predicted small differences in other life history traits but, as in the lichen case study, these differences were small relative to the uncertainty captured by bootstrapped confidence intervals (Table 1). Interestingly, the SHASH model had wider uncertainty intervals, particularly for lifespan and lifetime reproductive output, presumably because the additional parameters it requires introduce additional sources of uncertainty in these estimates. Thus, in this case study, modeling non-Gaussian size transitions with a Gaussian growth model may or may not influence IPM results depending on the target of the analysis, and whether the emphasis is on point estimates or uncertainty intervals.

6 Case study: lady orchid, *Orchis purpurea*

Our final case study examines selection on life history strategies in the lady orchid *Orchis purpurea*. In a prior study, Miller et al. 2012 analyzed how costs of reproduction (flowering or not in year t) affected growth from year t to $t+1$. The two growth kernels for flowering and non-flowering were then used in an IPM to quantify the optimal flowering size that balances the benefits of waiting to flower at larger sizes against the greater risk of death before flowering. The original study assumed Gaussian size transitions

with non-constant variance depending on initial size. Here we re-visit that analysis to derive improved growth kernels. We use this case study to illustrate several new elements and challenges, including modeling skewness and kurtosis as functions of expected future size.

The data, originated by Dr. Hans Jacquemyn, come from 368 plants in a Belgian population censused annually from 2003 through 2011. Here we use data only from the “light” habitat in the original study. We used the natural logarithm of total leaf area as the size variable in the IPM.

As a variation on software, we fitted the pilot Gaussian model using the `lmer` function in the **lme4** package, as in the original study. We fit three candidate linear models that included fixed effects of size in year t (model 1), additive effects of size and flowering status in year t (model 2), or an interaction between size and flowering (model 3), all including random intercepts for year. The interaction model was strongly favored ($\Delta AIC = 10.5$). Unlike our previous case studies, here we have multiple fixed effects (initial size and flowering status) that may influence the variance of future size. In cases such as this it is convenient to model variance as a function of expected future size, rather than initial size as we did with the lichens and cacti. The expected (or “fitted”) values reflect the combined influence of all fixed and random effects, and therefore implicitly account for multiple sources of variation in the variance.

Models where error variance is a function of fitted values cannot be fitted directly with `lme4` (nor in the `mgcv` functions for generalized additive models). But it can still be done with `lmer` through an iterative re-weighting approach, as follows. In `lmer`, weights w_i can be used to indicate that the observations y_i have error variance proportional to $1/w_i^2$. The iterative steps are as follows, and code that executes these steps is in `orchid_growth_modeling.R`.

1. Fit the expected value assuming Gaussian-distributed residuals with constant variance.
2. Fit the standard deviation of the residuals as a function of the corresponding fitted value.
3. Re-fit the model, with weights equal to the inverse of the standard deviation estimated in step 2.

We iterated steps 2 and 3 until the root mean square change in weights was below 10^{-6} . This is not elegant, but it works and converges quickly. In step 2, we modeled the log of the standard deviation (because standard deviations cannot be negative) as a quadratic polynomial in the fitted mean. In exploratory analyses we found that the quadratic term was necessary to fit the standard deviation. We did this for all candidate models and, for a fair AIC comparison, we then re-fit all candidate models with the weights estimated from the top model.

The updated model selection continued to favor the size \times flowering interaction model (3), but now with a weaker improvement over the next-best model ($\Delta AIC = 6.7$). The fitted mean (a function of initial size and flowering status) and fitted standard deviation (a function of the fitted mean) are shown in Fig. 3E. Spline regression found no trend in the mean of the resulting standardized residuals, and only small variation in the variance (Fig. 4E,F).

The best Gaussian model indicated a growth cost associated with flowering at the start of the census interval and a decline in growth variance with increasing expected values (Fig. 3E). The standardized residuals indicated negative skewness (10–20% difference in tail weight) and excess kurtosis (10–40% fatter than Gaussian) across much of the size distribution but both negligible at large expected sizes (Fig. 3F).

As possible improvements, we explored the skewed t and JSU distributions, both leptokurtic distributions with flexible skewness. Based on comparisons between real and simulated data we were happier with the skewed t , which we fit with a custom likelihood function similar to the JSU growth model for the lichen data. However, rather than re-fitting all parameters of the skewed t model, as we did with the lichen JSU, we built a “hybrid” likelihood function that uses the fitted mean and standard deviation from the best Gaussian model, and estimates parameters that control skewness and kurtosis as linear functions of expected future size. This is easy because the **gamlss.dist** package provides a parameterization of the skewed t in which the location parameter μ is the mean and scale parameter σ is the standard deviation (Rigby *et al.*, 2019). The hybrid likelihood looks like this:

```

381  ## GAU_fitted and GAU_sd are mean & standard deviation from the best Gaussian.
382  SSTLogLik=function(pars){
383      dSST(log_area_t1,
384      mu=GAU_fitted, sigma=GAU_sd,
385      nu = exp(pars[1] + pars[2]*GAU_fitted),
386      tau = exp(pars[3] + pars[4]*GAU_fitted)+2, log=TRUE)
387  }
388  p0<-c(0,0,0,0) ## default starting parameters
389  SSTout=maxLik(logLik=SSTLogLik,start=p0) ## fit with maxLik

```

Based on diagnostics of the standardized residuals, parameters that control skewness and kurtosis are defined as linear functions of the mean (note that the τ parameter uses a $\log(x-2)$ link function). This approach relies on the robustness of fitted Gaussian models to deviations from normality, which implies that the fitted mean and variance from a Gaussian model are good approximations for the mean and variance of the corresponding non-Gaussian model. If one is skeptical of this approach, it is possible to simultaneously re-fit all parameters of the skewed t . However, recall that unlike the lichen case study, the pilot Gaussian model here includes random year effects, and the expected values getting passed into dSST account for this source of variation. Estimating random effects “from scratch” with a custom likelihood model is possible (we provide guidance on doing this with a “shrinkage” approach, in Appendix S2), but generally should not be necessary. Instead, a key advantage of the hybrid approach is retention of the fitted random effects and associated variance components, which get shuttled from the Gaussian model into the non-Gaussian model without any fuss (though it was critical to use a parameterization of the skewed t for which μ is the mean and σ is the standard deviation). And, if this approach does not “work” (i.e., deviations from normality biased the fitted values of the Gaussian model) one would quickly find out when comparing simulated with real data. In this case, size transition data simulated

from this model corresponded favorably to the real data, much better than the pilot Gaussian model, including improvements in the standard deviation, skewness, and kurtosis of future size (Appendix S4, Fig. S-2).

Finally, we used the improved growth model to revisit key results of the original study. Miller et al. (2012) used the orchid IPM to estimate the evolutionarily stable strategy (ESS) as the mean size at flowering that maximizes lifetime reproductive success (R_0), given the constraint that flowering when small reduces growth and thus elevates mortality risk. Repeating that analysis here, we found that improved growth modeling has virtually no influence on predictions for optimal life history strategies (Fig. 8). ESS flowering sizes were nearly identical between IPMs with Gaussian vs skewed t growth models, and both aligned well with the observed mean flowering size (dashed vertical line in Fig. 8). Similarly, there were very small differences between growth functions in other metrics of orchid life history and, again, these differences were overwhelmed by uncertainty associated with parameter estimation (Table 1).

7 Discussion

Much of the appeal of IPMs has stemmed from their embrace of continuous size structure through regression-based approaches, and the potentially complex fixed- and random-effect structures that those approaches allow. Using familiar statistical tools and with relatively few parameters to estimate, IPM users can incorporate important sources of variation in demography and interrogate their influence on ecological and evolutionary dynamics. With this opportunity comes the burden of getting it right: an IPM is only as good as the statistical sub-models for the underlying data. The growth sub-model is the trickiest part because it defines a distribution of future size conditional on current size. Distributions have many properties – “moments” – and a good growth model should recapitulate the properties of real size transitions. The default assumption of Normally distributed size transitions, employed overwhelmingly across 20+ years of IPM studies, is an arbitrary historical precedent. In our case studies and, we suspect, more broadly, skewness and excess kurtosis were common

features of size transitions. Our most important message is that the assumption of normally-distributed size transitions can easily be abandoned, and a more inquisitive process of growth modeling should take its place.

We have attempted to lay out what that process should look like, emphasizing visual diagnostics to characterize how data deviate from Gaussian. One implication of relying on visual diagnostics is that goodness of fit is in the eye of the beholder. This empowers IPM users to make informed choices, but it is not very prescriptive; we have not suggested any hard rules for choosing among distributions, only that a good growth model should generate data that look like the real thing. Alternatively, model selection could be used to identify best-fitting growth distributions and best-fitting functions for higher moments. However, model selection among growth distributions with 3-5 parameters, each of which may be functions of multiple state variables or fitted values, can quickly explode in complexity, and we are not convinced it is worth the trouble.

Our work follows the important contribution of Peterson et al. 2019, who were similarly motivated by inadequacy of the Gaussian model but arrived at different recommendations. These authors developed a creative approach in which size data are transformed onto a $[0,1]$ scale and size transitions on that scale are modeled using beta regression. The beta distribution can accommodate positive, negative, or zero skew, potentially varying with size, so the Peterson et al. approach is a flexible option for skewed growth data. However, beta regression also has some limitations: common beta regression packages do not fit random effects (e.g., **betareg** (Cribari-Neto & Zeileis, 2010)) or do not do so reliably (in our experience **gamlss** regressions with random effects are numerically unstable); and the two-parameter beta distribution does not allow skewness and kurtosis to be fitted independently. Additionally, the initial transformation onto $[0,1]$ scale requires estimating extreme quantiles of the growth distribution (e.g., 0.01 and 0.99) as a function of initial size. In our experience those quantile estimates can be very sensitive to how size-dependence is modeled, and model selection is challenging for extreme quantiles where data are (by definition) very sparse. Rather than picking one distribution as a new default, users can leverage the vast arsenal of continuous probability distributions – all at one’s fingertips with a few lines of code – so that the data and their particular deviations from

normality can guide the choice of a better distribution. It is also possible to use mixtures of multiple growth distributions, as done by Shriver et al. 2012 to model “normal” and “extreme” types of lichen shrinkage. In re-analyzing that data set, we found that a single, flexible distribution (Johnson’s *S-U*) could recapitulate the observed size transitions, but there may be other cases where mixture distributions are preferable or necessary.

In all of our case studies, non-Gaussian growth models always yielded more satisfying fits to size transition data than the Gaussian models published in those papers. However, to our relief, none of these re-analyses yielded a “gotcha” result that overturned results of the original study. In fact, in this small sampling of case studies, improved growth modeling had weak to modest effects on IPM results, similar in magnitude to the results of Peterson *et al.* (2019), and for most species and life history metrics these effects were overwhelmed by the uncertainty associated with parameter estimation (Table 1). For some case studies, one might argue that non-Gaussian modeling was not worth the trouble – only it was almost no trouble at all, and we could not have known whether or not a non-Gaussian model would have made a difference before fitting it. Even where Gaussian and improved growth models differed in IPM results, we do not know “true” values to compare them against, only that one describes the data better than another, and a closer match to the data does not necessarily translate to better predictive ability due to the bias-variance trade-off.

We caution against taking too much comfort in weak effects of “improved” growth modeling; in other scenarios the choice of growth distribution could be more consequential. We focused on life history metrics such as mean lifespan and mean lifetime reproductive success (Table 1). It is possible that higher moments of those traits (e.g., variance or skewness of lifespan and lifetime reproduction) are more sensitive to the tails of the growth distribution. It is also worth noting that most of our case studies focused on perennial life histories (perennial plants and lichens) characterized by relatively slow growth, heavy losses during recruitment, and high survival once established, and these species all had mean lifespans between one and six years and generation times on the order of decades. Life histories such as these may be relatively robust to subtle features of the growth kernel. In Appendix S3 we present three additional case studies

that broaden our life history coverage, including pike (*Esox lucius*), a fish with a generation time of four to five years and creosotebush (*Larrea tridentata*), a desert shrub that is virtually immortal once established. Life history metrics from these “fast” and “slow” populations were no more sensitive to improved growth modeling than those of the perennial plants and lichens (Table 1). More systematic comparative analyses may provide insight into which types of species and life histories are more likely to exhibit strong skewness and kurtosis, and which demographic quantities are more or less sensitive to these features of size transition.

Our case studies illustrate a diversity of software packages and computational approaches, to reflect the diversity of preferences and habits that the community of IPM analysts bring to their own problems. We like spline generalized additive models (gams) for their flexibility and for **mgcv**’s numerous options for distribution families and overall speed and reliability. However, there are some applications for which classical parametric regression would be preferable because the coefficients carry biological meaning. For example, regression coefficients may be targets of natural selection (Rees & Ellner, 2016) and may combine to influence traits of interest such as the expected size at flowering (e.g. in Fig. 8A), a function of the intercept and slope of the size-dependent flowering function (Metcalf *et al.*, 2003). Some potentially useful distributions are not available in linear modeling software packages, but that should not be a barrier to their use: as in several of our case studies, custom likelihood functions allow non-Gaussian models without sacrificing the complex, multi-level features that one might be accustomed to fitting in **lme4**, for example. Bayesian analysis may further broaden the options for non-Gaussian candidate distributions and may help estimate hard-to-fit parameters through the brute force of sampling algorithms. Bayesian analysis also provides a natural way to propagate uncertainty from vital rate sub-models to full model predictions (Elder & Miller, 2016), *as an alternative to our bootstrapping approach. However, as of this writing, most of the non-Gaussian distributions that we have discussed are not available in popular Bayesian software packages such as Stan or JAGS. While user-defined distributions can always be coded from scratch, this may be a significant technical barrier for many IPM analysts.*

From the outset there have been concerns about “how well these methods [IPM growth kernels] can deal with different patterns of growth, stasis, and shrinkage” (Morris & Doak, 2002, p. 200), compared to “binning” methods that use observed transition frequencies between user-defined size classes as the transition probabilities in a (possibly large) matrix model (Doak *et al.*, 2021). The non-Gaussian models that we have considered here are not a panacea. For example, none of them allow bimodal growth, such as might occur if herbivore- or pathogen-attached individuals experience rapid tissue loss. When the shape of the growth distribution is nearly the same for all initial sizes, a nonparametric IPM growth kernel can be defined from a kernel density estimate for scaled residuals (Ellner *et al.*, 2016, p. 288). Outside that special situation, nonparametric approaches require choosing multiple smoothing parameters, which is very challenging. We are currently exploring whether “targeted learning” approaches developed for causal inference (van der Laan & Rose, 2011) can be used to circumvent smoothing parameter selection. Targeted learning starts with a pilot model and updates it iteratively to achieve unbiased estimates and valid confidence intervals for a particular “target” quantity, such as λ or mean lifespan. Preliminary results suggest that targeted learning with a deliberately under-smoothed pilot model works well for complex growth patterns (Zhou & Hooker, 2024). But nonparametric methods are data-hungry, so when departures from Gaussian are quantitative rather than qualitative, parametric modeling as developed here will make more efficient use of limited data.

Conclusion

Gaussian-distributed size transitions are probably the exception in nature, not the rule, yet two decades of IPM studies have relied overwhelmingly on Gaussian growth models. Using tools not available when IPMs were first developed, it should often be possible now to make major improvements over a Gaussian model, without worrying about finding the “best” alternative. By generating predicted size transitions that are truer to the data, IPM analysts can narrow the gap between model and nature.

Acknowledgements: This research was supported by US NSF grants DEB-1933497 to SPE and DEB-1754468, 2208857, and 2225027 to TEXM. The Sevilleta LTER (source of the cactus and creosote case studies) is supported by DEB-1655499 and DEB-1748133. Giles Hooker gave us the very good idea to use quantile regression instead of binning to estimate trends in skewness and kurtosis. Ali Campbell and Jacob Moutouama provided helpful discussion and comments on the manuscript.

Authorship statement: All authors discussed all aspects of the research and contributed to developing methods, analyzing data, and writing and revising the paper.

Conflict of interest statement: The authors have none to declare.

Literature Cited

Adler, P.B., Kleinhesselink, A., Hooker, G., Taylor, J.B., Teller, B. & Ellner, S.P. (2019) Data from: Weak interspecific interactions in a sagebrush steppe? Conflicting evidence from observations and experiments. Dryad Data set. <https://doi.org/10.5061/dryad.96dn293>.

Anscombe, F.J. & Glynn, W.J. (1983) Distribution of the kurtosis statistic b_2 for normal samples. *Biometrika* **70**, 227–234.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48.

Compagnoni, A., Bibian, A.J., Ochocki, B.M., Rogers, H.S., Schultz, E.L., Sneek, M.E., Elder, B.D., Iler, A.M., Inouye, D.W., Jacquemyn, H. *et al.* (2016) The effect of demographic correlations on the stochastic population dynamics of perennial plants. *Ecological Monographs* **86**, 480–494.

Conn, P.B., Johnson, D.S., Williams, P.J., Melin, S.R. & Hooten, M.B. (2018) A guide to Bayesian model checking for ecologists. *Ecological Monographs* **88**, 526–542.

540 Cooch, E.G. & White, G.C. (2020, accessed 5/17/2020) *Program MARK - a 'gentle introduction'*. phidot.org.

541 Coulson, T. (2012) Integral projections models, their construction and use in posing hypotheses in ecology.
542 *Oikos* **121**, 1337–1350.

543 Cribari-Neto, F. & Zeileis, A. (2010) Beta regression in R. *Journal of statistical software* **34**, 1–24.

544 Crone, E.E. (2016) Contrasting effects of spatial heterogeneity and environmental stochasticity on population
545 dynamics of a perennial wildflower. *Journal of Ecology* **104**, 281–291.

546 Czachura, K. & Miller, T.E. (2020) Demographic back-casting reveals that subtle dimensions of climate
547 change have strong effects on population viability. *Journal of Ecology* **108**, 2557–2570.

548 D'Agostino, R.B. (1970) Transformation to normality of the null distribution of g_1 . *Biometrika* **57**, 679–681.

549 Davis, C. (2015) *sgt: Skewed Generalized T Distribution Tree*. R package version 2.0.

550 DiCiccio, T.J & Efron, B. (1996) Bootstrap Confidence Intervals. *Statistical Science* **11**, 189–228.

551 Doak, D.F., Waddle, E., Langendorf, R.E., Louthan, A.M., Isabelle Chardon, N., Dibner, R.R., Keinath,
552 D.A., Lombardi, E., Steenbock, C., Shriver, R.K., Linares, C., Begoña Garcia, M., Funk, W.C., Fitzpatrick,
553 S.W., Morris, W.F. & DeMarche, M.L. (2021) A critical comparison of integral projection and matrix
554 projection models for demographic analysis. *Ecological Monographs* **91**, e01447.

555 Easterling, M.R., Ellner, S.P. & Dixon, P.M. (2000) Size-specific sensitivity: applying a new structured
556 population model. *Ecology* **81**, 694–708.

557 Elder, B.D. & Miller, T.E. (2016) Quantifying demographic uncertainty: Bayesian methods for integral
558 projection models. *Ecological Monographs* **86**, 125–144.

- 559 Ellis, M.M. & Crone, E.E. (2013) The role of transient dynamics in stochastic population growth for nine
560 perennial plants. *Ecology* **94**, 1681–1686.
- 561 Ellner, S.P., Adler, P.B., Childs, D.Z., Hooker, G., Miller, T.E. & Rees, M. (2022) A critical comparison of
562 integral projection and matrix projection models for demographic analysis: Comment. *Ecology* **103**, e3605.
- 563 Ellner, S.P., Childs, D.Z. & Rees, M. (2016) *Data-driven Modeling of Structured Populations: A Practical*
564 *Guide to the Integral Projection Model*. Springer, New York.
- 565 Fasiolo, M., Wood, S.N., Zaffran, M., Nedellec, R. & Goude, Y. (2020) qgam: Bayesian non-parametric
566 quantile regression modelling in R. *arXiv preprint arXiv:2007.03303* .
- 567 Gu, C. (2013) *Smoothing Spline ANOVA Models*. Springer Science+Business Media, New York, 2 edn.
- 568 Hadfield, J.D. *et al.* (2010) MCMC methods for multi-response generalized linear mixed models: the
569 mcmcglmm r package. *Journal of Statistical Software* **33**, 1–22.
- 570 Hérault, B., Bachelot, B., Poorter, L., Rossi, V., Bongers, F., Chave, J., Paine, C.T., Wagner, F. & Baraloto,
571 C. (2011) Functional traits shape ontogenetic growth trajectories of rain forest tree species. *Journal*
572 *of ecology* **99**, 1431–1440.
- 573 Hernández, C.M., Ellner, S.P., Snyder, R.E. & Hooker, G. (2024) The natural history of luck: A synthesis
574 study of structured population models. *Ecology Letters* pp. 0 – 1.
- 575 Komsta, L. & Novomestky, F. (2015) moments: Moments, cumulants, skewness, kurtosis and related
576 tests. *R package version 0.14.1*.
- 577 Louthan, A.M., Keighron, M., Kiekebusch, E., Cayton, H., Terando, A. & Morris, W.F. (2022) Climate
578 change weakens the impact of disturbance interval on the growth rate of natural populations of venus
579 flytrap. *Ecological Monographs* **92**, e1528.

580 McGillivray, H. (1986) Skewness and asymmetry: measures and orderings. *Annals of Statistics* **14**, 994–1011.

581 Metcalf, C.J.E., Ellner, S.P., Childs, D.Z., Salguero-Gómez, R., Merow, C., McMahon, S.M., Jongejans,
582 E. & Rees, M. (2015) Statistical modelling of annual variation for inference on stochastic population
583 dynamics using Integral Projection Models. *Methods in Ecology and Evolution* **6**, 1007–1017.

584 Metcalf, J.C., Rose, K.E. & Rees, M. (2003) Evolutionary demography of monocarpic perennials. *Trends*
585 *in Ecology & Evolution* **18**, 471–480.

586 Miller, T.E. (2020) Long-term study of tree cholla demography in the los pinos mountains, sevilleta national
587 wildlife refuge. <https://doi.org/10.6073/pasta/dd06df3f950afe4a4642306182237d13>.

588 Miller, T.E., Louda, S.M., Rose, K.A. & Eckberg, J.O. (2009) Impacts of insect herbivory on cactus
589 population dynamics: experimental demography across an environmental gradient. *Ecological*
590 *Monographs* **79**, 155–172.

591 Miller, T.E., Williams, J.L., Jongejans, E., Brys, R. & Jacquemyn, H. (2012) Evolutionary demography
592 of iteroparous plants: incorporating non-lethal costs of reproduction into integral projection models.
593 *Proceedings of the Royal Society B: Biological Sciences* **279**, 2831–2840.

594 Morris, W.F. & Doak, D.F. (2002) *Quantitative Conservation Biology: Theory and Practice of Population*
595 *Viability Analysis*. Sinauer Associates, Sunderland, Mass.

596 Ochocki, B.M., Drees, T. & Miller, T.E. (2023) Density-dependent demography of creosote
597 bush (*larrea tridentata*) along grass-shrub ecotones. <https://doi.org/10.6073/pasta/ca53c16f16dcf9fb11f3ee99ea5445ac>.

598
599 Ohm, J.R. & Miller, T.E. (2014) Balancing anti-herbivore benefits and anti-pollinator costs of defensive
600 mutualists. *Ecology* **95**, 2924–2935.

601 Ozgul, A., Childs, D.Z., Oli, M.K., Armitage, K.B., Blumstein, D.T., Olson, L.E., Tuljapurkar, S. & Coulson,
602 T. (2010) Coupled dynamics of body mass and population growth in response to environmental change.
603 *Nature* **466**, 482–485.

604 Peterson, M.L., Morris, W., Linares, C. & Doak, D. (2019) Improving structured population models with
605 more realistic representations of non-normal growth. *Methods in Ecology and Evolution* **10**, 1431–1444.

606 Plard, F., Schindler, S., Arlettaz, R. & Schaub, M. (2018) Sex-specific heterogeneity in fixed morphological
607 traits influences individual fitness in a monogamous bird population. *The American Naturalist* **191**, 106–119.

608 R Core Team (2022) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical
609 Computing, Vienna, Austria.

610 Rees, M., Childs, D.Z. & Ellner, S.P. (2014) Building integral projection models: a user’s guide. *Journal*
611 *of Animal Ecology* **83**, 528–545.

612 Rees, M. & Ellner, S.P. (2016) Evolving integral projection models: evolutionary demography meets
613 eco-evolutionary dynamics. *Methods in Ecology and Evolution* **7**, 157–170.

614 Rigby, R.A., Stasinopoulos, M.D., Heller, G.Z. & De Bastiani, F. (2019) *Distributions for modeling location,*
615 *scale, and shape: Using GAMLSS in R*. CRC press.

616 Salguero-Gómez, R. & Casper, B.B. (2010) Keeping plant shrinkage in the demographic loop. *Journal*
617 *of Ecology* **98**, 312–323.

618 Schielzeth, H., Dingemanse, N.J., Nakagawa, S., Westneat, D.F., Alagüe, H., Teplitsky, C., Réale, D.,
619 Dochtermann, N.A., Garamszegi, L.Z. & Araya-Ajoy, Y.G. (2020) Robustness of linear mixed-effects
620 models to violations of distributional assumptions. *Methods in ecology and evolution* **11**, 1141–1152.

- 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.
- Stasinopoulos, D.M., Rigby, R.A. *et al.* (2007) Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software* **23**, 1–46.
- Stubberud, M.W., Vindenes, Y., Vøllestad, L.A., Winfield, I.J., Stenseth, N.C. & Langangen, Ø. (2019) Effects of size-and sex-selective harvesting: An integral projection model approach. *Ecology and Evolution* **9**, 12556–12570.
- van der Laan, M.J. & Rose, S. (2011) *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer Science and Business Media.
- Wan, X., Wang, W., Liu, J. & Tong, T. (2014) Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC medical research methodology* **14**, 1–13.
- Williams, J.L., Miller, T.E. & Ellner, S.P. (2012) Avoiding unintentional eviction from integral projection models. *Ecology* **93**, 2008–2014.
- Winfield, I., Fletcher, J. & James, J. (2013b) Pike growth data 1944-1995. NERC Environmental Information Data Centre, <https://doi.org/10.5285/637d60d6-1571-49af-93f7-24c1279d884d>.
- Wood, S. (2017) *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edn.
- Zhang, J.L. (2014) Comparative investigation of three Bayesian p values. *Computational Statistics & Data Analysis* **79**, 277–291.
- Zhou, Y. and Hooker, G. (2024) Targeted Maximum Likelihood Estimation for Integral Projection Models in Population Ecology. *arXiv preprint* 2411.08150.

Table 1: Life history attributes derived from IPM kernels that included Gaussian or “improved” growth sub-models for six case studies. The improved distributions were JSU (lichen, creosote), SHASH (cactus, pike, coral), and skewed t (orchid). Pike, creosote and coral case studies are presented in Appendix S3. The original coral case study assumed an open population with constant recruitment from a large source region, so some life history attributes cannot be computed from the published model. Values in parenthesis are 95% bootstrap confidence intervals, specifically the bias-corrected (BC) bootstrap confidence intervals Diccio & Efron (1996). Table can be reproduced from scripts `crossspp_growth.R`, `Vulpicida_boot.R`, `Akumal_corals_boot.R`.

Species	Growth model	λ	Lifespan	Lifetime reproductive output	Age at reproduction	Generation time
Lichen (<i>V. pinastri</i>)	Gaussian	1.01 (0.99, 1.04)	6.4 (3.6, 11.1)	1.4 (0.5, 3.1)	6.5 (5.7, 7.3)	40.8 (30.5, 57.4)
	Improved	1.00 (0.98, 1.03)	5.4 (3.1, 9.7)	1 (0.4, 2.4)	6.4 (5.4, 7.3)	36.6 (27.5, 48.6)
Cactus (<i>C. imbricata</i>)	Gaussian	0.994 (0.99, 0.996)	6.11 (3.66, 8.63)	21.8 (8.27, 49.4)	17.6 (1.75, 22.7)	189 (131, 266)
	Improved	0.993 (0.991, 0.998)	5.38 (3.34, 16.3)	13.4 (5.72, 251)	20.3 (1.21, 22.2)	179 (133, 298)
Orchid (<i>O. purpurea</i>)	Gaussian	1.09 (1.08, 1.1)	1.08 (1.06, 1.11)	20.0 (12.6, 31.0)	5.07 (4.78, 5.31)	104 (73.1, 150)
	Improved	1.09 (1.08, 1.1)	1.08 (1.06, 1.1)	19.3 (12.0, 29.9)	5.03 (4.75, 5.3)	100.7 (71.0, 145.0)
Pike (<i>E. Lucius</i>)	Gaussian	1.62 (1.35, 1.89)	1.2 (1.09, 1.35)	5.75 (2.9, 9.7)	1.09 (1.03, 1.18)	4.96 (4.26, 5.84)
	Improved	1.62 (1.35, 1.88)	1.2 (1.09, 1.35)	5.76 (2.91, 9.73)	1.09 (1.03, 1.18)	4.94 (4.30, 5.84)
Creosote (<i>L. tridentata</i>)	Gaussian	1.033 (1.029, 1.04)	4.52 $\times 10^6$ (2.14 $\times 10^5$, 1.82 $\times 10^8$)	3.19 $\times 10^5$ (1.27 $\times 10^4$, 1.24 $\times 10^7$)	32.7 (29.2, 36.0)	5.27 $\times 10^6$ (2.50 $\times 10^5$, 1.95 $\times 10^8$)
	Improved	1.034 (1.03, 1.04)	3.26 $\times 10^5$ (1.98 $\times 10^3$, 1.66 $\times 10^7$)	2.31 $\times 10^4$ (5.83 $\times 10^2$, 1.27 $\times 10^6$)	32.8 (29.3, 36.0)	3.7 $\times 10^5$ (2.63 $\times 10^3$, 1.93 $\times 10^7$)
Coral (<i>G. ventalina</i>)	Gaussian	—	17.3 (11.9, 24.3)	—	10.5 (9.3, 11.8)	31.6 (28.3, 36.7)
	Improved	—	17.5 (12.1, 24.3)	—	10.7 (9.4, 12.2)	30.9 (27.4, 35.3)

8 Figure Legends

Figure 1 Recommended steps in growth modeling (left) and guide to common non-Gaussian distributions of size x for $x \in \mathbb{R}$ that can accommodate different combinations of skewness and kurtosis (right). All of these distributions (often including multiple versions or parameterizations of each) are available in the R package **gamlss.dist**, except for the skewed generalized t , which is available in the package **sgt** (Davis, 2015).

Figure 2 Histograms of skewness and kurtosis estimates using moment-based definitions (top two panels), compared with the nonparametric measures based on quantiles (bottom two panels). Note the very large differences in scale. Histograms are based on 5000 replicate draws of a sample of 200 independent values, from a t distribution with 8 degrees of freedom. Dotted red vertical lines mark the minimum and maximum of sample estimates, and dashed blue lines show the 5th and 95th percentiles. The true value is indicated by a black dot on the x -axis. Figure drawn by script `NPmoments.R`

Figure 3 Best Gaussian models and diagnostics of standardized residuals for lichen (*Vulpicida pinastri*) **A,B**, cactus *Cylindriopuntia imbricata* **C,D**, and orchid *Orchis purpurea* **E,F** case studies. **A,C**, fitted mean (red) and standard deviation (blue) of size at time $t+1$ conditional on initial size at time t . **E**, fitted means for plants that were vegetative (solid line) or flowering (dashed line) at the start of the census interval and standard deviation as a function of the fitted mean (inset). **B,D,F** Quantile regressions of scaled residuals (lines show 5%, 10%, 25%, 50%, 75%, 90%, and 95% quantiles) and non-parametric measures of skewness (blue) and excess kurtosis (red) derived from them. In **B,D** scaled residuals are shown with respect to initial size and in **F** they are shown with respect to fitted values. Figure made by script `crossspp_growth.R`.

Figure 4 Diagnostic plot for trends in the mean (left column) or variance (right column) of scaled residuals from a pilot Gaussian model, for the lichen (*Vulpicida pinastri*) **A,B**, cactus *Cylindriopuntia imbricata* **C,D**, and orchid *Orchis purpurea* **E,F** case studies. In **A,C,E** the standardized residuals are plotted, and in **B,D,F** the absolute values of standardized residuals, as functions of fitted mean subsequent size values. The solid

curves are cubic splines (R function `smooth.spline`) fitted by generalized cross-validation with a modest over-penalization of model degrees of freedom to prevent overfitting (`penalty=1.4` as recommended by Gu (2013)). The numbers appearing above each panel are the standard deviation of the values on the spline regression curve, evaluated at all of the fitted values. Figure made by script `crossspp_diagnose_pilot.R`.

Figure 5 Comparisons among real lichen data and data simulated from Gaussian and JSU growth models for NP mean, NP standard deviation, NP skewness, and NP excess kurtosis of future size conditional on current size. Colored lines show 100 simulated data sets from the fitted Gaussian (red) or JSU (blue) growth models. Thick black line shows the real data. Gaussian and JSU data are offset by one unit and the real data line is duplicated with a one-unit offset for ease of visualization. Figure made by script `Vuplicida_IPMs.R`.

Figure 6 Extinction risk estimated from individual-based simulation of IPMs based on Gaussian and Johnson's S-U (JSU) growth distributions. Figure made by script `Vuplicida_IPMs.R`.

Figure 7 Temporal (A) and spatial (B) heterogeneity in fitness for the tree cholla cactus (*Cylindriopuntia imbricata*) predicted by IPMs using Gaussian or SHASH growth models. Figure made by script `cactus_growth_modeling_qgam.R`.

Figure 8 Orchid life history results from IPMs using Gaussian or skewed t growth models. Lifetime reproductive success (R_0) is shown as a function of mean size of flowering. Dashed vertical line shows the observed mean flowering size.