

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 and Evolutionary Biology, Cornell University, Ithaca, New York

Running header: Better growth modeling for IPMs

Acknowledgements: This research was supported by US NSF grants DEB-1933497 to SPE and DEB-1754468, 2208857, and 2225027 to TEXM.

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

Data accessibility statement: No original data appear in this paper. Should the paper be accepted, all computer scripts supporting the results will be archived in a Zenodo package, with the DOI included at the end of the article. During peer review, our data and code are available at https://github.com/texmiller/IPM_size_transitions.

Conflict of interest statement: None.

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

¹ **Abstract**

- ² 1. Integral Projection Models (IPMs) are widely used for studying the dynamics of
³ continuously size-structured populations. IPMs require a growth sub-model that
⁴ describes the probability of future size conditional on current size. Over the past
⁵ two decades, most IPM studies have assumed that this probability is normally-
⁶ distributed, despite repeated calls for non-Gaussian approaches that accommodate
⁷ skewness and kurtosis known to occur in size transition data.
- ⁸ 2. We provide a general workflow for modeling size transitions that accommodates
⁹ non-Gaussian growth patterns while retaining the desirable features that Gaussian
¹⁰ approaches typically provide (ecologically important covariates and random effects).
¹¹ Our approach emphasizes visual diagnostics of residuals from pilot Gaussian mod-
¹² els and quantile-based metrics of skewness and kurtosis that vet the fit of the Gaus-
¹³ sian distribution and guide the selection of an alternative, if necessary. We illus-
¹⁴ trate our methods by reanalyzing size transition data from published IPM studies,
¹⁵ targeting a diversity of demographic quantities including population growth rate,
¹⁶ extinction risk, and evolutionarily stable life history strategies.
- ¹⁷ 3. Across five case studies, skewness and excess kurtosis were common features of
¹⁸ size transition data and non-Gaussian growth models consistently generated simu-
¹⁹ lated data that were more consistent with the real data than pilot Gaussian models.
²⁰ However, in these case studies, the effects of “improved” growth modeling on IPM
²¹ results were moderate to weak, and differed in direction or magnitude between dif-
²² ferent outputs from the same model.
- ²³ 4. Using tools that were not available when IPMs were first developed, it is now possi-
²⁴ ble to fit non-Gaussian models to size transition data without sacrificing ecological
²⁵ complexity; our worked examples demonstrate how, including open-access data and
²⁶ computing scripts. Doing so, as guided by careful interrogation of the data, will re-
²⁷ sult in a model that better represents the population for which it is intended.

²⁸ **Keywords**

²⁹ demography; growth; integral projection model; kurtosis; skewness

³⁰ Introduction

³¹ Structured demographic models – matrix and integral projection models (MPMs and
³² IPMs) – are powerful tools for data-driven modeling of population and community dy-
³³ namics that are widely used in basic and applied settings. In contrast to MPMs for pop-
³⁴ ulations with discrete structure (life stage, age class, etc.), IPMs (Easterling et al., 2000)
³⁵ readily accommodate populations structured by continuous state variables, most com-
³⁶ monly size. A related innovation of the IPM framework is its emphasis on regression-
³⁷ based modeling for parameter estimation, which often carries important advantages for
³⁸ making the most of hard-won data (Ellner et al., 2022).

³⁹ A standard workflow allows ecologists to assemble an IPM from data using famili-
⁴⁰ ar regression tools to describe growth, survival, reproduction, and other demographic
⁴¹ transitions as functions of size (Coulson, 2012; Ellner et al., 2016). The relative ease of
⁴² the regression-based approach, accommodating multiple covariates (e.g., environmental
⁴³ factors, experimental treatments) and complex variance structures (e.g., random effects,
⁴⁴ correlated errors), has facilitated a growing body of IPM literature that examines how
⁴⁵ biotic or abiotic factors affect population dynamics (e.g., Louthan et al., 2022; Ozgul
⁴⁶ et al., 2010; Schultz et al., 2017) and explores the consequences of demographic hetero-
⁴⁷ geneity associated with spatial, temporal, and individual variation (e.g., Compagnoni
⁴⁸ et al., 2016; Crone, 2016; Plard et al., 2018). The vital rate regressions (or “sub-models”)
⁴⁹ are the bridge between the individual-level data and the population-level model and its
⁵⁰ predictions; it is important to get these right.

⁵¹ Compared to other vital rates, growth is special. The regression sub-models for
⁵² survival and reproduction only need to provide a single mean value as functions of
⁵³ size (we use “size” as the name for whatever continuous variable defines the population
⁵⁴ structure, which could instead be immune competence, mother’s weight, etc.). But for
⁵⁵ modeling growth, the full probability distribution of subsequent size, conditioned on
⁵⁶ initial size, must be defined. This distribution defines the growth ‘kernel’ $G(z', z)$ that
⁵⁷ gives the probability density of any future size z' at time $t + 1$ conditional on current size
⁵⁸ z at time t . Whenever survival and reproduction are size-dependent, the entire distribu-
⁵⁹ tion of size transitions can strongly influence IPM predictions because this distribution
⁶⁰ governs how frequently size changes are much greater or much lower than average.

⁶¹ The original template for modeling size transitions in IPMs was provided by East-
⁶² erling et al. 2000. They first tried simple linear regression, assuming normally dis-
⁶³ tributed size changes with constant variance. Because the residuals from this regression
⁶⁴ exhibited non-constant variance, they used a two-step approach that estimated the size-

65 dependence in the residual variance (better options soon became available, such as the
66 `lme` function in R). However, even after accounting for non-constant variance, growth
67 data may still deviate from the assumption that size transitions are normally distributed.
68 Size transitions are often skewed such that large decreases are more common than large
69 increases (Peterson et al., 2019; Salguero-Gómez and Casper, 2010), or vice versa (Stub-
70 berud et al., 2019). Size transitions may also exhibit excess kurtosis ('fat tails'), where
71 extreme growth or shrinkage is more common than predicted by the tails of the normal
72 distribution (Hérault et al., 2011).

73 The observation that the normal (or "Gaussian") distribution may poorly describe
74 size transitions in real organisms has been made before, and several studies have empha-
75 sized that alternative distributions should be explored (Easterling et al., 2000; Peterson
76 et al., 2019; Rees et al., 2014; Williams et al., 2012). Nonetheless, default use of Gaussian
77 growth distributions (often with non-constant variance) remains the standard practice.
78 The general state-of-the-art in the literature appears to remain where it was 20 or so
79 years ago, using the default model without pausing to examine critically whether or not
80 it actually provides a good description of the data. We are guilty of this, ourselves.

81 The persistence of Gaussian growth modeling is understandable. There is a long
82 tradition of statistical modeling built on the assumption of normally distributed resid-
83 uals with constant variance. Popular packages such as `lme4` (Bates et al., 2007), `mgcv`
84 (Wood, 2017), and `MCMCglmm` (Hadfield et al., 2010) make it easy to fit growth models
85 with potentially complex fixed- and random-effect structures, but the possible distribu-
86 tions of continuous responses are limited, and default to Gaussian. Abandoning these
87 convenient tools for the sake of more flexible growth modeling means, it may seem,
88 sacrificing the flexibility to rigorously model diverse and potentially complex sources of
89 variation in growth, some of which may be the motivation driving the study in the first
90 place.

91 The question we address here is: how can ecologists escape the apparent trade-off
92 between realistically capturing the variance, skew, and kurtosis of size transition data
93 on the one hand, and flexibly including the multiple covariates and random effects that
94 often have substantial impacts on demographic rates? In this article, we offer an answer.

95 Our goal here is to present and illustrate a general and practical "recipe" that moves
96 growth modeling past the standards set over 20 years ago, using software tools available
97 now.¹ Like any recipe, users may need to make substitutions or add ingredients to
98 suit their needs. Our approach emphasizes graphical diagnostics for developing and

¹Our statements about what is available now are based on what tools reliably deliver in our experience, not on what they promise.

⁹⁹ evaluating growth models, rather than a process centered on statistical model selection.
¹⁰⁰ Through a set of empirical case studies we demonstrate how a simple workflow, using
¹⁰¹ tools that were nonexistent or not readily available when IPMs first came into use, makes
¹⁰² it straightforward and relatively easy to identify when the default model is a poor fit to
¹⁰³ the data, and to then choose and fit a substantially better growth model that is no harder
¹⁰⁴ to use in practice. We illustrate our approach with three published case studies (and
¹⁰⁵ two additional case studies in the Appendix), including data from our own previous
¹⁰⁶ work. In each case, the Gaussian assumption does not stand up to close scrutiny. We
¹⁰⁷ illustrate how we could have done better, and the consequences of “doing better” for
¹⁰⁸ our ecological inferences. All of our analyses may be reproduced from code and data
¹⁰⁹ that are publicly available (see Data accessibility statement).

¹¹⁰ A workflow for growth modeling

¹¹¹ The modeling workflow that we suggest runs as follows (Fig. 1):

- ¹¹² 1. *Fit a “pilot” model or models assuming a Gaussian distribution, but allowing for non-*
¹¹³ *constant variance.*

¹¹⁴ This step is familiar to most IPM users, as it is the start and end of the traditional
¹¹⁵ workflow. A well-fitted Gaussian model accurately describes the mean and variance
¹¹⁶ of future size conditional on current size and possibly on other measured covari-
¹¹⁷ ates or random effects. This step may include model selection to identify which
¹¹⁸ treatment effects or environmental drivers affect the mean and/or variance of future
¹¹⁹ size. Non-constant variance is often fitted in a two-stage process, first fitting mean
¹²⁰ growth assuming constant variance, then doing a regression relating the squared
¹²¹ residuals to initial size or the fitted mean of subsequent size. Fitting mean and
¹²² variance simultaneously, as can be done with R packages **mrgcv** and **nmle**, is ad-
¹²³ vantageous when possible because incorrectly assuming constant variance can affect
¹²⁴ model selection for the mean. But two-step fitting may be convenient when there
¹²⁵ are multiple fixed and random effects that can affect growth variance, because the
¹²⁶ fitted mean value implicitly accounts for all of them. We illustrate both one-step and
¹²⁷ two-step approaches in the case studies below.

¹²⁸ Allowing non-constant variance removes the need for transforming the data to
¹²⁹ stabilize the growth variance. Transformation remains an option when it does not
¹³⁰ create new problems (see Discussion²), and it may have advantages besides vari-

²*Discussion content needed.*

131 ance stabilization. In particular log-transformation is often appropriate for size data
132 (Ellner et al., 2016), and it helps to avoid eviction at small sizes (Williams et al., 2012).

133 2. *Use statistical and graphical diagnostics to identify if and how the standardized residuals*
134 *deviate from Gaussian, and to identify a more appropriate distribution.*

135 If the Gaussian pilot model is valid, the set of standardized (or “scaled”) residuals
136 (standardized by the standard deviation) should be Gaussian with mean zero and
137 unit variance, with no skew or excess kurtosis. This criterion provides a straightfor-
138 ward test for whether to accept a Gaussian growth model or explore alternatives. If
139 the standardized residuals are satisfactorily Gaussian, skip to the final step of the
140 workflow.

141 There are many ways that growth data may deviate from Gaussian, and the na-
142 ture of those deviations can guide the search for a better distribution. Frequentist
143 tests such as the D’Agostino test of skewness (D’Agostino, 1970) and the Anscombe-
144 Glynn test of kurtosis (Anscombe and Glynn, 1983) could be used to diagnose
145 whether the aggregate distribution of standardized residuals deviates from normal-
146 ity (R package **moments** (Komsta and Novomestky, 2015)). However, the aggregate
147 distribution of standardized residuals may be misleading if properties such as skew
148 and kurtosis vary with size or other covariates. For example, a change in the di-
149 rection of skewness from small to large sizes might produce zero overall skewness,
150 but really requires a distribution flexible enough to accommodate both positive and
151 negative skew, such as the skewed normal or Johnson S_U distributions. Alterna-
152 tively, growth data may lack skew but may exhibit leptokurtosis (in which case the t
153 distribution may be a good choice) or may shift from platykurtosis to leptokurtosis
154 depending on initial size (in which case the power exponential distribution may be
155 a good choice). It is therefore essential to visualize trends in distribution properties
156 with respect to size, either initial size (for simple models with only size-dependence)
157 or expected future size (for models with multiple fixed effects). Fig. 1 includes guid-
158 ance on how the skew and kurtosis properties of the standardized residuals suggest
159 options for an appropriate growth distribution. In our case studies we take advan-
160 tage of the many distributions provided in the **gamlss** R package (Stasinopoulos
161 et al., 2007), but any other distributions with the necessary properties can be used.

162 3. *Refit the growth model using the chosen distribution.*

163 In models with multiple covariates and/or random effects, each potentially affecting
164 several distribution parameters (location, scale, skew, kurtosis) in different ways,
165 “refit the model” could entail a massive model selection process to identify the

“right” or “best” non-Gaussian model. And with so many options, model uncertainty may be overwhelming and over-fitting becomes a significant risk even if precautions against it are taken. We therefore argue for adopting the more modest goal of remedying any evident defects in the Gaussian model. As we demonstrate below, the functional forms for the mean and standard deviation (or location and scale parameters) can often be carried over from the pilot Gaussian model into a non-Gaussian distribution, leaving skew and kurtosis as the targets for improvement.

Our recommendation for this step is based on the fact that parameter estimation using Gaussian regression models is generally robust to deviations from normality (Schielzeth et al., 2020), meaning that the mean of the Gaussian model is probably a good proxy for the mean of the non-Gaussian model (and if it is not, the next step in the workflow would 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., skewness switches from positive to negative with size).

4. *Test the final model through graphical diagnostics comparing simulated and real growth data.*

A good model will generate simulated data that look like the real data. Again, it is important to inspect the properties of simulated data conditional on initial size, rather than examining the aggregate distribution. We provide examples below of informative comparisons between simulated and real data, based mainly on quantiles. If the simulated data do not correspond well with real data, alternative (possibly more flexible) growth distributions, or more complex functions relating distribution parameters to size and other covariates, should be explored. However, we again caution against a full-blown model selection exercise. Instead, alternative models should be chosen to remedy observable discrepancies between real and simulated size transition data, and at most slightly modified based on final diagnostics and statistical tests.

How should skewness and kurtosis be measured?

Improvement of a Gaussian model will involve scrutiny of skewness and kurtosis, so measurement of these properties warrants some attention. The standard measures of skewness and kurtosis (tail thickness) are based on the third and fourth central moments,



Figure 1: General workflow of recommendations for IPM 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 package `gamlss.dist`, except for the skewed generalized *t*, which is available in the package `sgt` (Davis, 2015).

197 respectively, of the distribution:

$$198 \quad \text{Skewness} = \frac{m_3}{\sigma^3}, \quad \text{Excess kurtosis} = \frac{m_4}{\sigma^4} - 3 \quad (1)$$

199 where $m_k = \mathbb{E}(X - \bar{X})^k$ is the k^{th} central moment of a random quantity X and σ^2 is the
200 variance (second central moment). A Gaussian distribution has zero skewness and zero
201 excess kurtosis.

202 The standard measures are easy to calculate but their use for choosing and eval-
203 uating growth models is hindered by their poor sampling properties. Because empirical
204 estimates involve high powers of data values, a few outliers can produce very inaccurate
205 estimates. Figure 2 shows a simulated example, where the underlying “data” are a sam-

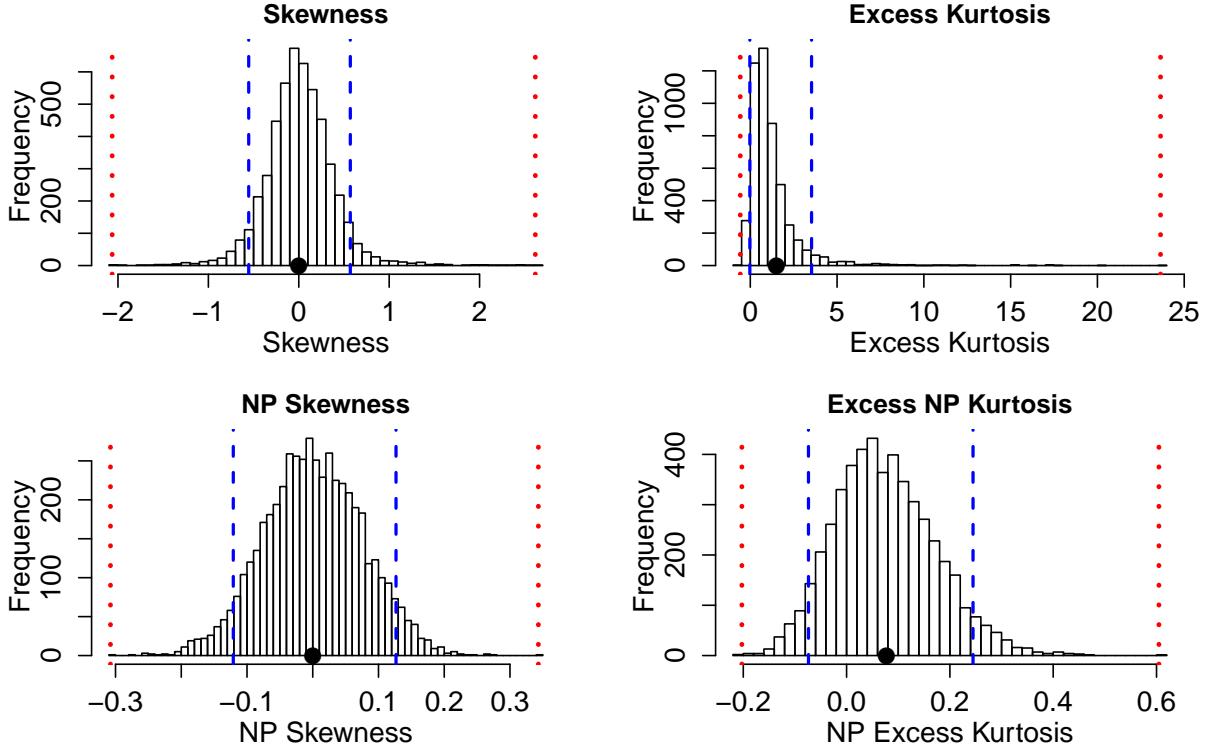


Figure 2: Histograms of skewness and kurtosis estimates using moment-based definitions (top two panels), compared with the nonparametric measures based on quantiles (bottom two panels). Note the very large differences in scale. Histograms are based on 5000 replicate draws of a sample of 200 independent 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`

ple of size 200 from a t distribution with 8 degrees of freedom; the true skew is 0, and the true excess kurtosis is 1.5. The distance between the largest and smallest estimates (indicated by the dotted red vertical lines), relative to the distance between the 5th and 95th percentiles, shows the broad extent of extreme values that can occur even with a large sample, especially for kurtosis.

We therefore use nonparametric (NP) measures of skew and kurtosis that are based on quantiles and thus are less sensitive to a few extreme values. Let q_α denote the α quantile of a distribution 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)

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

216 NP Skewness measures the asymmetry between the tails of the distribution above and
 217 below the median. The size of the upper tail can be measured (for any $0 < \alpha < 0.5$) by
 218 $\tau_U = q_{1-\alpha} - q_{0.5}$; for $\alpha = 0.05$ this is the difference between the 95th percentile and the
 219 median. The lower tail size is $\tau_L = q_{0.5} - q_\alpha$. The definition above is equivalent to

$$220 \quad \text{NP Skewness} = \frac{\tau_U - \tau_L}{(\tau_U + \tau_L)}. \quad (3)$$

221 An NP Skewness of ± 0.2 says that the difference in tail sizes is 20% of their total. The
 222 range of possible values is -1 to 1. Both $\alpha = 0.25$ (sometimes called “Kelly’s skewness”) and
 223 $\alpha = 0.1$ (“Bowley’s skewness”) are common choices. We used $\alpha = 0.1$, unless
 224 otherwise stated.

225 An analogous quantile-based measure of kurtosis (Jones et al., 2011) is

$$226 \quad \text{NP Kurtosis} = \frac{q_{1-\alpha} - q_\alpha}{q_{0.75} - q_{0.25}}. \quad (4)$$

227 For $\alpha = 0.05$, NP Kurtosis is the difference between the 95th and 5th percentiles, relative
 228 to the interquartile range. To facilitate interpretation, we scale NP Kurtosis relative to its
 229 value for Gaussian distribution, and subtract 1 so that the value for a Gaussian is zero.
 230 We call this “NP Excess Kurtosis”. A value of ± 0.2 means that the tails are on average
 231 20% heavier (or lighter) than those of a Gaussian with the same interquartile range. We
 232 calculate NP Kurtosis using $\alpha = 0.05$ unless otherwise stated, to focus on the tail edges,
 233 but again this is somewhat arbitrary.

234 Figure 2C,D illustrate how, applied to exactly the same simulated samples, the non-
 235 parametric measures produce a smaller fraction of highly inaccurate estimates caused
 236 by a few extreme values in the sample. But also note that, in contrast to the moment-
 237 based measures, numerically small values of the nonparametric measures (e.g., 0.1 or 0.2)
 238 should not be disregarded, because they are both scaled so that a value of 1 indicates
 239 extremely large departures from a Gaussian distribution.

240 Quantile-based estimation of skewness and kurtosis carries the added value that
 241 quantile regression methods may be used to derive these properties of size transitions as
 242 continuous functions of initial size or expected future size. In the examples below, we use
 243 the **qgam** package (Fasiolo et al., 2020) to fit smooth additive quantile regression models,
 244 which have the flexibility to accommodate nonlinear size-dependence in skewness and
 245 kurtosis. One risk of a gam-based approach is that fitted quantiles may be too “wiggly”
 246 without constraints on their complexity. In the examples below, we limit complexity by
 247 fitting splines with $k = 4$ basis functions. For the gam-averse, other quantile regression
 248 models may be suitable.

249 For consistency with nonparametric skewness and kurtosis, in comparisons of real
 250 and simulated data below and in the Appendix, we use quantile-based measures of
 251 mean and standard deviation, and use quantile regression to visualize these as functions
 252 of size. Specifically, following Wan et al. (2014),

$$253 \quad \text{NP Mean} = \frac{q_{0.25} + q_{0.5} + q_{0.75}}{3} \quad (5)$$

254 and

$$255 \quad \text{NP SD} = \frac{q_{0.75} - q_{0.25}}{1.35}. \quad (6)$$

256 1 Case study: lichen, *Vulpicida pinastri*

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

267 With initial size as the only predictor, a simple way to fit a Gaussian model with
 268 nonconstant variance is the `gam` function in `mgcv` library (Wood, 2017) using the `gaulss`
 269 family. Following a bit of model selection, we fit the mean and standard deviation of
 270 future size as second-order polynomials of current size, then derived the scaled residuals
 271 from the fitted mean and standard deviation. Here, the first argument to `gam()` is a two-
 272 element list that defines the linear predictors for mean and sd:

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

279 The data and fitted mean and standard deviation are shown in Fig. 3A. Quantile re-
 280 gression on the scaled residuals generates the diagnostics shown in Fig. 3B (see script

281 *Vulpicida_IPMS.R*). As expected based on previous analyses, visual analysis of the stan-
282 dardized residuals indicated negative skew, especially at larger sizes (Fig. 3B). We also
283 find positive excess kurtosis for all sizes.

284 We turned to the Johnson's *S-U* (JSU) distribution for improvement. The JSU is
285 a four-parameter, leptokurtic distribution that can accommodate positive or negative
286 skew; it also has the convenient property that location and scale parameters *mu* and
287 *sigma* are the mean and standard deviation, respectively, which facilitates a natural cor-
288 respondence to the pilot Gaussian model. The JSU is not available as a distribution
289 family in any of the standard linear or additive modeling packages, to our knowledge,
290 but that need not be a barrier for this or any other distribution as long as we can write
291 a likelihood function (*dJSU()* is provided by **gamlss**). Following the best-fit Gaussian
292 model, we defined *mu* and *sigma* of the JSU as second-order polynomials of initial size
293 and, based on signals of skewness and kurtosis in the standardized residuals (Fig. 3B),
294 we define parameter *nu* (which controls skewness) as a linear function of size and *tau*
295 (which controls kurtosis) as a positive constant; the likelihood function therefore has
296 nine parameters to estimate. We fit the model using the **maxLik** package and starting
297 values for *mu* and *sigma* based on estimates from the pilot Gaussian model:

```
298 ## define function that returns the JSU negative log-likelihood
299 LogLikJSU=function(pars){
300   dJSU(t1,
301     mu=pars[1]+pars[2]*t0+pars[3]*t0^2,
302     sigma=exp(pars[4]+pars[5]*t0+pars[6]*t0^2),
303     nu = pars[7]+pars[8]*t0,
304     tau = exp(pars[9]), log=TRUE)
305 }
306 ## starting parameter values
307 p0<-c(coef(fitGAU)[1:6],0,0,0)
308 ## fit with maxlik
309 outJSU=maxLik(logLik=LogLikJSU,start=p0*exp(0.2*rnorm(length(p0))),
310 method="BHHH",control=list(iterlim=5000,printLevel=2),finalHessian=FALSE);
```

311 Data simulation from the fitted JSU model indicates a compelling improvement over the
312 best Gaussian model, not only in skewness and kurtosis (Fig. 4C-D) but also standard
313 deviation (4B).

314 To understand the practical consequences of improved growth modeling, we as-
315 sembled the remainder of the lichen IPM following Shriver et al. 2012. The asymptotic

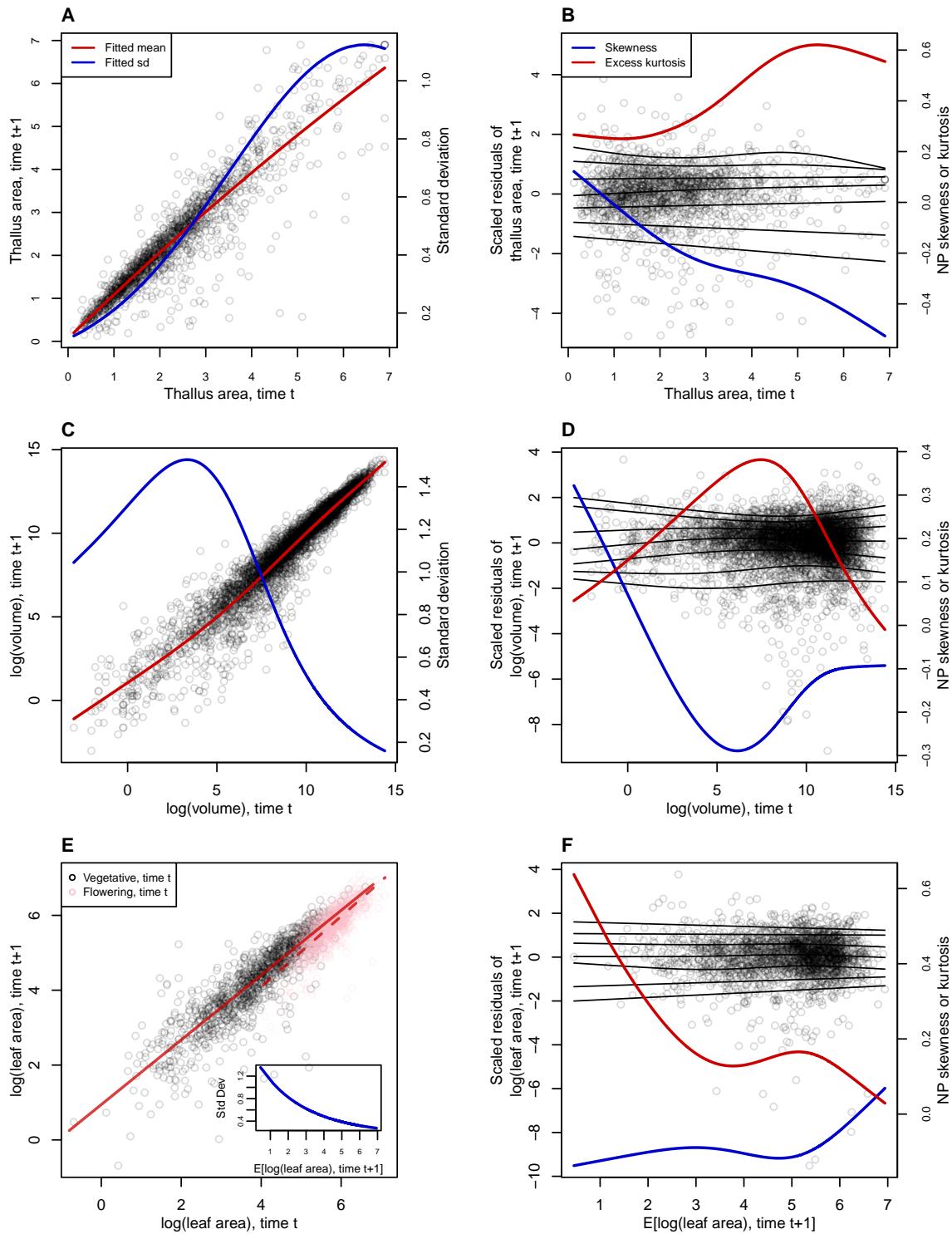


Figure 3: Best Gaussian models and diagnostics of standardized residuals for lichen (*Vulpicida pinastri*) **A,B**, cactus *Cylindropuntia imbricata* **C,D**, and orchid *Orchis purpurea* **E,F** case studies. **A,C**, fitted mean (red) and standard deviation (blue) of size at time $t+1$ conditional 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 `crosspp-growth.R`.

316 population growth rate λ based on Gaussian growth differs from the JSU growth model
 317 by about 1% annual population growth (Table 1), in line with results of Peterson et al.
 318 2019. However, even this modest difference can lead to biased estimates of extinction
 319 risk from the Gaussian model, particularly over longer time horizons (Fig. 5). We also
 320 explored differences in other life history metrics (Table 1).³ For example, the JSU growth
 321 model predicts values for mean lifespan, mean lifetime reproductive success, and gen-
 322 eration time that are 15–25% lower than the Gaussian growth model. In this case study,
 323 properly modeling non-normal size transitions – which was easy to do with a few extra
 324 lines of code – can have important effects on ecological inferences.

| Species | Growth model | λ | Lifespan | Lifetime reproductive output | Age at reproduction | Generation time |
|--|--------------|-----------|-----------|------------------------------|---------------------|-----------------|
| Lichen (<i>Vulpicida pinastri</i>) | Gaussian | 1.001 | 6.443 | 1.031 | 5.588 | 33.869 |
| | Improved | 0.992 | 5.395 | 0.773 | 5.39 | 29.051 |
| Cactus (<i>Cylindriopuntia imbricata</i>) | Gaussian | 0.992 | 2.002 | 0.023 | 19.108 | 162.438 |
| | Improved | 0.993 | 2.002 | 0.019 | 21.676 | 179.474 |
| Orchid (<i>Orchis purpurea</i>) | Gaussian | 1.091 | 1.081 | 20.009 | 5.064 | 104.125 |
| | Improved | 1.09 | 1.079 | 19.378 | 5.027 | 100.753 |
| Pike (<i>Esox Lucius</i>) | Gaussian | 1.762 | 1.122 | 1.172 | 1.311 | 4.807 |
| | Improved | 1.764 | 1.123 | 1.236 | 1.303 | 4.788 |
| Creosote (<i>Larrea tridentata</i>) | Gaussian | 1.039 | 21651.948 | 1998.486 | 29.338 | 241517.676 |
| | Improved | 1.04 | 19613.824 | 1814.89 | 31.668 | 215330.883 |

Table 1: Life history attributes derived from IPM kernels that included Gaussian or “improved” growth sub-models for five case studies. The improved distributions were JSU (lichen, creosote), SHASH (cactus, pike), and skewed t (orchid). Table can be reproduced from script `crosspp-growth.R`.

325 One could argue that the lichen data set was a convenient “straw man” to disqualify
 326 Gaussian growth, since it was recognized by the original and subsequent IPM analysts
 327 that this species requires a skewed distribution of size transitions (Peterson et al., 2019;
 328 Shriver et al., 2012). In all remaining case studies, including those in the Appendix,
 329 we re-examine growth data that were modeled as Gaussian by the data originators in
 330 published IPM studies.

331 2 Case study: tree cholla cactus, *Cylindriopuntia imbricata*

332 The next case study, focusing on the tree cholla cactus *Cylindriopuntia imbricata* at the
 333 Sevilleta Long-Term Ecological Research site in central New Mexico, adds a new feature
 334 on top of the simple size-dependent regressions in the previous study: random effects

³What is the best way to cite Chrissy Hernandez' life history functions?

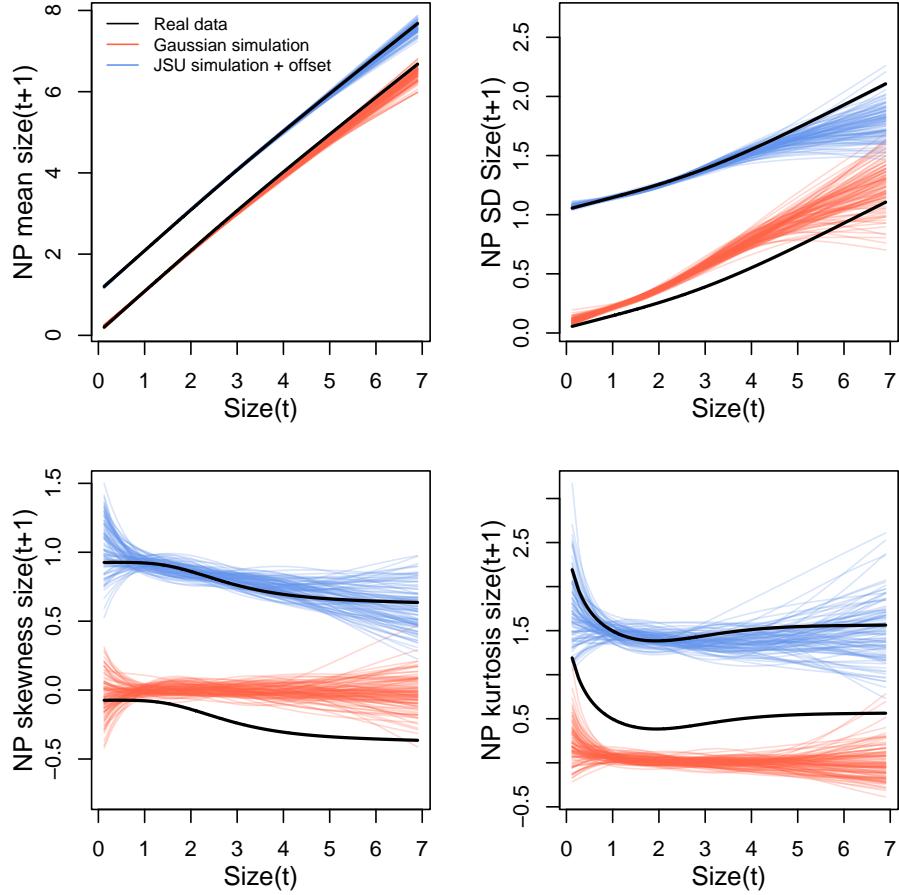


Figure 4: Comparisons among real lichen data and data simulated from Gaussian and JSU growth models for NP mean, NP standard deviation, NP skewness, and NP excess kurtosis of future size conditional on current size. Colored lines show 100 simulated data sets from the fitted Gaussian (red) or JSU (blue) 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`.

associated with temporal (year) and spatial (plot) environmental heterogeneity. This long-term study of cactus demography 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 and Miller, 2020; Elderd and Miller, 2016; Miller et al., 2009; Ohm and Miller, 2014). In fact, (Elderd and Miller, 2016) presented a Gaussian growth model fit to the cactus data as an example of a well fit growth function, based on a marginal distribution of residuals that appeared approximately 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 (Elderd and

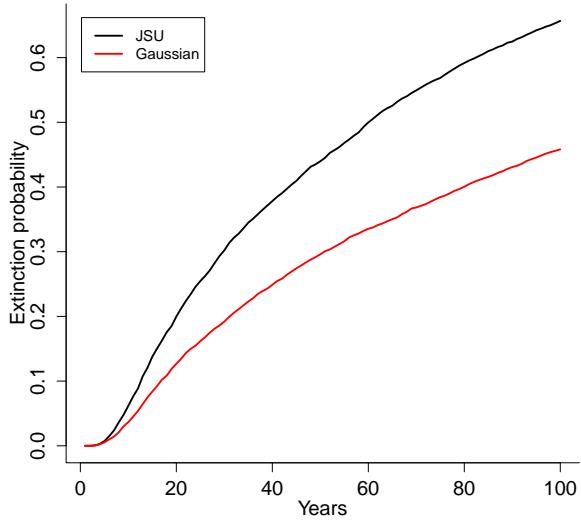


Figure 5: Extinction risk estimated from individual-based simulation of IPMs based on Gaussian and Johnson’s S-U (JSU) growth distributions. Figure made by script `Vuplicida_IPMs.R`.

344 Miller, 2016)). While PPCs and the associated “Bayesian P-value” are popular diagnostic
 345 tools, they are often considered to be too conservative (Conn et al., 2018; Zhang, 2014),
 346 failing to reject marginally bad models even though they are very effective in rejecting
 347 models that are terrible. The choice of discrepancy function (the statistic used to com-
 348 pare real and simulated data) can also be limiting: in our previous work, we used a
 349 discrepancy function focused on variance (the sum of the squared residuals), so we had
 350 a built-in blind-spot for mismatches in higher moments. In the clarity of hindsight, the
 351 PPC gave a false sense of security; the Gaussian was a poor choice all along.

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

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

361 `# cactus is the data frame`

```

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

```

365 The standardized residuals, accounting for variance that peaks at small-to-medium sizes
366 (Fig. 3C), show clear signals of negative skew and positive excess kurtosis across most
367 of the size distribution but strongest in the middle of the size distribution (Fig. 3D).

368 To better capture size transitions, we need a distribution with negative skew and
369 positive excess kurtosis, but both of which may be negligible at some sizes. We first tried
370 Johnson's S_U and then the skewed t distributions, both of which are limited to positive
371 excess kurtosis. Both distributions provided some improvement over the Gaussian, but
372 were not happy with the fit of either. Iterating through the workflow (Fig. 1), we arrived
373 at the SHASH distribution, which is more flexible than either the JSU or skewed t , capa-
374 ble of capturing a greater range of kurtosis for a given amount of skew, and vice versa
375 (Jones and Pewsey (2009); Appendix S.1). Furthermore, SHASH is available as an **mgcv**
376 family, allowing for flexible, non-monotonic size-dependence in skewness and kurtosis
377 without the need for model selection on specific size-dependent functions; through it-
378 erations of trial and error, we found this flexibility was necessary to generate simulated
379 data that compared favorably to the real data. Here, the first argument to **gam()** is now
380 a four-element list giving the linear predictors for mean, standard deviation, skewness,
381 and kurtosis:

```

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

```

388 Data simulated from the SHASH model compared favorably to the real data (Fig.
389 S-7). Similar to the lichen case study, we see that correctly modeling skewness and
390 kurtosis improved estimation of the mean and standard deviation (Fig. S-7A,B), yielding
391 a growth model that is truer to the data than the pilot Gaussian fit.

392 We explored how improved growth modeling influenced IPM results. The λ values
393 predicted by Gaussian and SHASH growth functions, corresponding to the average plot
394 and year, were nearly identical (Table 1) but we could also leverage structure of the study
395 design to quantify demographic variance associated with temporal and spatial hetero-
396 geneity. We used the fitted random effects from the vital rate models to estimate the

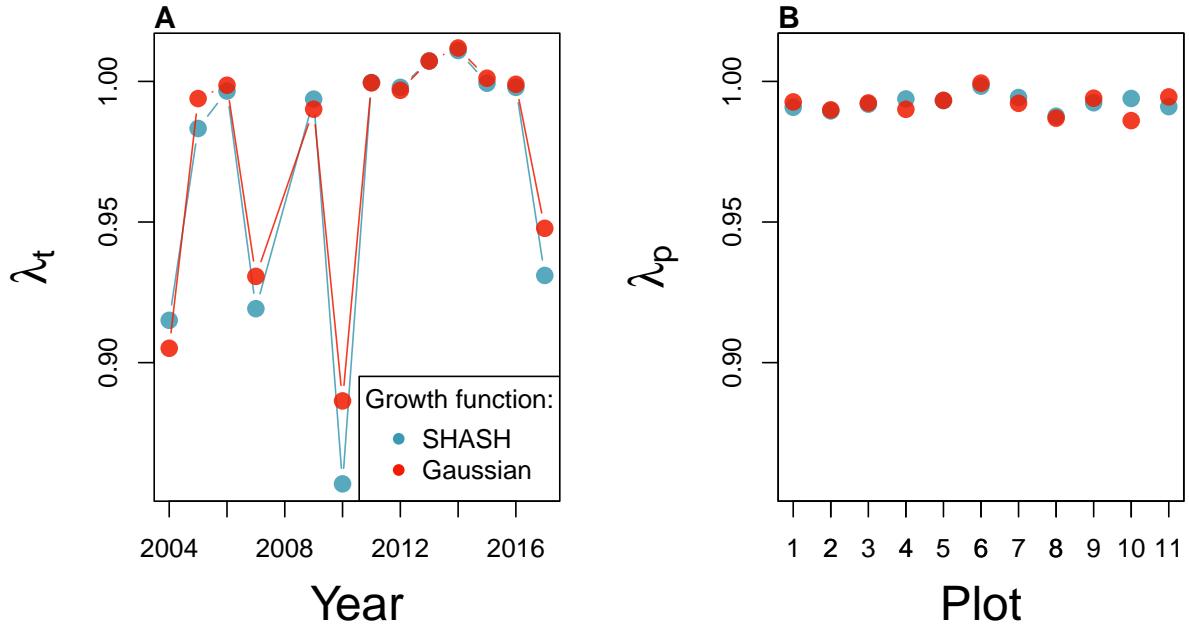


Figure 6: Temporal (A) and spatial (B) heterogeneity in fitness for the tree cholla cactus (*Cylindriopuntia imbricata*) predicted by IPMs using Gaussian or SHASH growth models. Figure made by script `cactus_growth_modeling_qgam.R`.

397 asymptotic growth rate for each year (λ_t), centered on the average plot, and for each plot
 398 (λ_p), centered on the average year. We found that the Gaussian growth model tended to
 399 over-estimate λ_t , particularly in the harshest years (Fig. 6A), and thus under-estimated
 400 temporal variance in fitness ($SD(\lambda_{t(Gaussian)}) = 0.042$, $SD(\lambda_{t(SHASH)}) = 0.048$). The op-
 401 posite was true for plot-to-plot variation ($SD(\lambda_{p(Gaussian)}) = 0.0037$, $SD(\lambda_{p(SHASH)}) =$
 402 0.0028), although spatial variation in fitness was much lower than temporal variation
 403 (Fig. 6B). The difference in temporal variance would suggest that Gaussian growth
 404 modeling would lead to over-estimation of the stochastic growth rate λ_S , since temporal
 405 variance has a negative influence on λ_S . However, this was not the case: stochastic IPMs
 406 based on Gaussian and SHASH growth models had nearly identical stochastic growth
 407 rates ($\lambda_S(Gaussian) = 0.9907$, $\lambda_S(Gaussian) = 0.991$). This is likely because temporal
 408 fluctuations in vital rates, which is where the SHASH growth model would make a dif-
 409 ference, have a weaker influence on λ_S than the temporal fluctuations in size structure
 410 that they generate (Compagnoni et al., 2016; Ellis and Crone, 2013). Thus, depending
 411 on the target of one's analysis, modeling non-Gaussian size transitions with a Gaussian
 412 growth model could bias results in either direction, or make no difference at all.

⁴¹³ 3 Case study: lady orchid, *Orchis purpurea*

⁴¹⁴ Our final case study examines selection on life history strategies in the lady orchid *Or-*
⁴¹⁵ *chis purpurea*. In a prior study, Miller et al. 2012 contrasted the growth trajectories from
⁴¹⁶ year t to $t + 1$ for plants that did or did not flower in year t , as a way to quantify costs of
⁴¹⁷ reproduction. The different growth kernels were then used in an IPM to quantify evo-
⁴¹⁸ lutionarily stable life history strategies: the optimal flowering size that balances benefits
⁴¹⁹ of waiting to flower at larger sizes against the risk of dying before reaching those sizes.
⁴²⁰ The original study assumed a Gaussian distribution of size transitions and allowed for
⁴²¹ non-constant variance with respect to initial size. Here we re-visit that analysis applying
⁴²² our growth modeling workflow to derive improved growth kernels for flowering and
⁴²³ non-flowering orchids. We use this case study to illustrate several new elements and
⁴²⁴ challenges, including modeling skewness and kurtosis as functions of expected future
⁴²⁵ size (instead of initial size).

⁴²⁶ The data, originated by Dr. Hans Jacquemyn and used here with permission, come
⁴²⁷ from 368 plants in a Belgian population that was censused annually from 2003 through
⁴²⁸ 2011 (for this reanalysis we are using data only from the “light” habitat). Size was mea-
⁴²⁹ sured as leaf area (cm^2) summed over all leaves, and we analyzed the natural logarithm
⁴³⁰ of total leaf area as the size variable of the IPM.

⁴³¹ As a variation on software, our pilot Gaussian approach used the **lme4** package to
⁴³² fit three candidate linear models for size in year $t + 1$ that included fixed effects of size
⁴³³ in year t (model 1), additive effects of size and flowering status in year t (model 2), or
⁴³⁴ an interaction between size and flowering (model 3), all including random intercepts for
⁴³⁵ year. The interaction model with strongly favored ($\Delta AIC = 10.5$). Unlike our previous
⁴³⁶ case studies, here we have multiple fixed effects (initial size and flowering status) that
⁴³⁷ may influence the variance of future size. In cases such as this, it is often convenient
⁴³⁸ to model variance as a function of expected future size, rather than initial size as we
⁴³⁹ did with the lichens and cacti. The expected (or “fitted”) values reflect the combined
⁴⁴⁰ influence of all fixed and random effects, and therefore implicitly account for multiple
⁴⁴¹ sources of variation in the variance. While there are several software packages for simul-
⁴⁴² taneously modeling Gaussian mean and variance as functions of independent variables
⁴⁴³ (**mgcv** for gam models as we saw above, **nlme** for linear models), modeling variance as
⁴⁴⁴ a function of the mean is trickier because they cannot easily be fit simultaneously. Here
⁴⁴⁵ we use an iterative re-weighting approach – which is not elegant, but it works.

446 For Gaussian models, weights w_i can be used to indicate that the observations y_i
 447 vary in their dispersion around the mean. In general, the iterative steps are as follows,
 448 and code to execute these steps may be found in `orchid_growth_modeling.R`.

1. Fit the expected value and normally-distributed residuals with constant standard deviation σ :

$$y_i = \mu_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma)$$

2. Fit the standard deviation of the residuals as a function of the expected value.
 Weights are derived as the inverse of the fitted variance:

$$\epsilon_i \sim N(0, f(\mu_i))$$

$$w_i = 1/f(\mu_i)^2$$

3. Re-fit the observation model, weighting the residual variance according to step 2:

$$y_i = \mu_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma \times \sqrt{w_i})$$

449 We iterated steps 2 and 3 until the weights did not change. In step 2, we modeled the
 450 standard deviation as a second-order polynomial of the expected value ($\log(f(\mu_i)) =$
 451 $\beta_0 + \beta_1 \mu_i + \beta_2 \mu_i^2$); in exploratory analyses we found that the the second-order term
 452 provided necessary flexibility to fit the standard deviation. We did this for all candidate
 453 models and, for fair AIC comparison, we re-fit all candidate models with the same
 454 weights, estimated from the top model.

455 The updated model selection continued to favor the size \times flowering interaction
 456 model (3), but now with a weaker improvement over the next-best model ($\Delta AIC =$
 457 6.7). The fitted mean (a function of initial size and flowering status) and fitted standard
 458 deviation (a function of the fitted mean) are shown in Fig. 3E. The best Gaussian model
 459 indicated a growth cost associated with flowering at the start of the census interval and a
 460 decline in growth variance with increasing expected values. The standardized residuals
 461 indicated negative skewness (10–20% difference in tail weight) and excess kurtosis (10–
 462 40% fatter than Gaussian) across much of the size distribution but both negligible at
 463 large expected sizes (Fig. 3F).

464 As improvements, we explored the skewed *t* and JSU distributions, both leptokurtic
 465 distributions with flexible skewness. We were happier with the skewed *t*, which we fit

466 with a custom likelihood function similar to the JSU growth model fit to the lichen data.
 467 However, rather than re-fitting all the parameters of the skewed t model, as we did with
 468 the lichen JSU, here we build a “hybrid” likelihood function that uses the fitted mean
 469 and standard deviation from the best Gaussian model, and estimates parameters that
 470 control skewness and kurtosis as linear functions of expected future size. This is easy
 471 because the **gamlss.dist** package provides a parameterization of the skewed t in which
 472 the location parameter μ is the mean and scale parameter σ is the standard deviation
 473 (Rigby et al., 2019). The hybrid likelihood looks like this:

```

474 ## t1 and t0 are the size (log(leaf area))obervations
475 ## GAU_fitted and GAU_sd are mean and standard deviation from best Gaussian
476 ## pars is a vector of free parameters to be estimated
477 SSTLogLik=function(pars){
478   dSST(log_area_t1,
479     mu=GAU_fitted,
480     sigma=GAU_sd,
481     nu = exp(pars[1] + pars[2]*GAU_fitted),
482     tau = exp(pars[3] + pars[4]*GAU_fitted)+2,
483     log=TRUE)
484 }
485 ## starting parameters
486 p0<-c(0,0,0,0,0)
487 ## fit with maxLik
488 SSTout=maxLik(logLik=SSTLogLik,start=p0*exp(0.2*rnorm(length(p0))))

```

489 Based on diagnostics of the standardized residuals, parameters that control skew-
 490 ness and kurtosis are defined as linear functions of the mean, and those coefficients
 491 are estimated by maximum likelihood (note that the tau parameter uses a $\log(x - 2)$
 492 link function). This approach relies on the robustness of Gaussian models to deviations
 493 from normality, which implies that the fitted mean and variance from a Gaussian model
 494 are good approximations for the mean and variance of a corresponding non-Gaussian
 495 model. If one is skeptical of this approach, it is possible to simultaneously re-fit all pa-
 496 rameters of the skewed t . However, recall that unlike the lichen case study, our pilot
 497 Gaussian approach included random effects for year, and therefore the expected values
 498 getting passed into dSST account for this source of variation. Coding random effects
 499 “from scratch” into a custom likelihood model is possible (we provide guidance on one
 500 way to do this, using the “shrinkage” approach, in the Supporting Information) but

should generally not be necessary. Instead, a key advantage of the hybrid approach is convenient 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 (it was critical that we used 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 through the simulation step of the workflow. 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 (Fig. S-8).

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. 7). 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. 7). Similarly, there were very small differences between growth functions in other metrics of orchid life history (Table 1).

4 Discussion

Much of the appeal of integral projection models has stemmed from their embrace of continuous size structure through reliance on regression-based approaches, and the potentially complex fixed- and random-effect structures that these 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: IPMs are good models of the populations they are intended represent only insofar as the statistical models provide good fits to the underlying data. The growth sub-model is the trickiest part of “getting it right” 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.

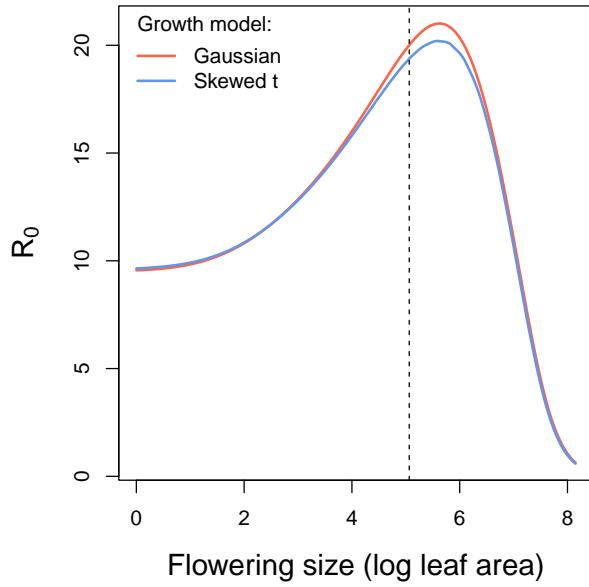


Figure 7: Orchid life history results from IPMs using Gaussian or skewed t growth models. Lifetime reproductive success (R_0) is shown as a function of mean size of flowering. Dashed vertical line shows the observed mean flowering size.

535 In our case studies (chosen simply because we had the data at our fingertips) and, we
 536 suspect, more broadly, skewness and excess kurtosis were common features of size trans-
 537 ition data: shrinking was more common than growing, and large changes in size were
 538 more common than a Gaussian model would predict. Our most important message is
 539 that the standard assumption of normally-distributed size transitions should be aban-
 540 doned and a more inquisitive process of growth modeling should take its place.

541 We have attempted to lay out a general workflow for what that process should
 542 look like, guided by visual diagnostics of standardized residuals that characterize the
 543 ways in which growth data may deviate from Gaussian. One implication of relying on
 544 visual diagnostics is that goodness of fit is in the eye of the beholder. This approach
 545 can empower IPM users to make informed choices, but it is not very prescriptive; we
 546 have not suggested any hard rules for when one or another distribution should be used,
 547 only that a good growth model should generate data that look like the real thing. Al-
 548 ternatively, model selection could be used to identify best-fitting growth distributions
 549 and best-fitting functions for higher moments. However, model selection among growth
 550 distributions with 3-5 parameters, each of which may be functions of multiple state vari-
 551 ables or fitted values, can quickly explode in complexity, and we are not convinced it is
 552 worth the trouble. It should be possible to find a good growth model without worrying
 553 about which one is “best”.

Our work follows the important contribution of Peterson et al. 2019, who were motivated by a similar problem (inadequacy of the Gaussian distribution for skewed size transitions) but arrived at different recommendations for dealing with the problem. These authors developed a creative approach in which size data are transformed onto a [0, 1] scale and then size transitions are modeled using beta regression. The beta distribution can accommodate positive, negative, or zero skew. As Peterson et al. demonstrate, this approach can be a viable option for skewed growth data. However, beta regression also has some downsides: common beta regression packages do not fit random effects (e.g., **betareg** (Cribari-Neto and Zeileis, 2010)) or do not do so reliably (in our experience **gamlss** regressions are numerically unstable); the two-parameter beta distribution has rigid mean-variance and skewness-kurtosis relationships that may not describe the data well; **and**⁴. Rather than shoe-horn size transition data into a default beta distribution, we find it more natural and appropriate to leverage the vast arsenal of real-valued probability distributions – all of them at one’s fingertips with a few lines of code – and let the data and their particular deviations from normality guide the choice of a better distribution.

While the arsenal of candidate distributions is indeed vast, in our analyses for this paper we found ourselves coming back time and again to a few usual suspects. The four-parameter SHASH distribution, for example, is able to flexibly accommodate independent, size-dependent variation in variance, skewness, and kurtosis, and it is available as a distribution family in the well-developed **mgcv** package. In our case study analyses it was consistently among the top non-Gaussian candidates and was our model of choice for several data sets. While we have emphasized the importance of moving away from a single default distribution of size transitions, if one were to want or need a default distribution one could do worse than the SHASH. In cases where size transitions are leptokurtic but consistently so across the size distribution, the Johnson’s S-U (used for lichens) and skewed *t* (used for orchids) distributions were easy to fit with custom likelihood functions. All of the distributions we have used (and the specific parameterizations we have implemented) share the property that their location and scale parameters correspond to the mean and standard deviation, which is not essential but it facilitates interpretation and an intuitive connection to the pilot Gaussian model. The five-parameter skewed generalized *t* (sgt) generalizes many other real-valued distributions and is therefore another, highly flexible option, but it does not share the location=mean and scale=sd property, and in our experience can be hard to fit. Finally, finding an appropriate non-Gaussian alternative does not solve all the problems of growth modeling. “Eviction”

⁴Steve, I recall you have a beef with transformation to [0,1], so that would go here.

589 from the approximating matrix of the IPM kernel is an ever-present danger and requires
590 vigilance to detect and correct (Williams et al., 2012).

591 In all of our case studies, non-Gaussian growth models always yielded more sat-
592 isfying fits to size transition data than the Gaussian models published in those papers.
593 However, much to our relief, none of these re-analyses yielded a “gotcha” result that
594 overturned results of the original study. In this small sampling of case studies, improved
595 growth modeling had weak to modest effects on IPM results, similar in magnitude to
596 the results of Peterson et al. (2019). We caution against taking too much comfort in this
597 outcome; we can imagine other scenarios in which the choice of the growth distribution
598 could be more consequential. It is worth noting that most of our case studies focused
599 on perennial life histories (perennial plants and lichens) characterized by relatively slow
600 growth, heavy losses during recruitment, and high survival once established, and these
601 species all had mean lifespans between one and six years and generation times on the or-
602 der of decades. Life histories such as these may be relatively robust to subtle features of
603 the growth kernel. In the Supporting Information we present two additional case studies
604 that broaden our life history coverage, including pike (*Esox lucius*), a fish with a gener-
605 ation time of four to five years and creosotebush (*Larrea tridentata*), a desert shrub that
606 is virtually immortal once established with a generation time exceeding 200,000 years.
607 Life history metrics from the “fast” fish population were no more sensitive to improved
608 growth modeling than those of the perennial plants and lichens, while the creosotebush
609 generation time differed by > 25,000 years between Gaussian and improved growth
610 models (Table 1). More systematic comparative analyses may provide insight into which
611 types of species and life histories are more likely to exhibit strong skewness and kurtosis,
612 and and which demographic quantities are more or less sensitive to these features of size
613 transition. It is worth noting, as we saw in several case studies, that different outputs
614 from the same model can be more or less sensitive to the choice of growth distribution.

615 Across our case studies we have attempted to illustrate a diversity of software pack-
616 ages and computational approaches to model fitting, to reflect the diversity of prefer-
617 ences and habits that the community of IPM users bring to their own analyses. We like
618 generalized additive models (gams) for their flexibility and for **mgcv**’s numerous op-
619 tions for distribution families and overall speed and reliability. However, there are some
620 applications for which classical parametric regression would be preferable because the
621 coefficients carry biological meaning. For example, regression coefficients may be tar-
622 gets of natural selection (Rees and Ellner, 2016) and may combine to influence traits of
623 interest such as the expected size at flowering (e.g. in Fig. 7A), a function of the intercept
624 and slope of the size-dependent flowering function (Metcalf et al., 2003). Some poten-

tially useful but relatively obscure distributions may not be available in linear modeling software packages, but that should not be a barrier to their use: as we have illustrated in several case studies, custom likelihood functions open up diverse possibilities for non-Gaussian growth modeling without sacrificing the complex, multi-level features that one might be accustomed to fitting in `lme4`, for example. We have illustrated fitting growth models using maximum likelihood but Bayesian analysis is another option that may further broaden the options of 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 the vital rate sub-models through the predictions of the IPM (Elder and Miller, 2016).

This paper has focused on size transitions, but IPMs have been increasingly extended in ways that capture other continuous state variables, and the same problems and solutions we propose should apply in those cases. For example, IPMs can be used to model infectious disease dynamics, where hosts may exhibit continuous variation in infection load (e.g., parasite density), and host vital rate processes depend on infection load (Metcalf et al., 2016; Wilber et al., 2016). Such models must define probabilities of future infection load conditional on current load, and would therefore benefit from the same modeling workflow that we have outlined for size transitions.

4.1 Conclusion

Our focus on modeling size transitions is admittedly deep in the weeds of demographic analysis that “casual

Acknowledgements

This research was supported by US NSF grants DEB-1933497 to SPE and DEB-1754468, 2208857, and 2225027 to TEXM.

5 Authorship statement

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

652 6 Data accessibility statement

653 No original data appear in this paper. Should the paper be accepted, all computer scripts
654 supporting the results will be archived in a Zenodo package, with the DOI included at
655 the end of the article. During peer review, our data and code are available at https://github.com/texmiller/IPM_size_transitions.
656

657 **Literature Cited**

- 658 Anscombe, F. J. and Glynn, W. J. (1983). Distribution of the kurtosis statistic b 2 for
659 normal samples. *Biometrika*, 70(1):227–234.
- 660 Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2007). The lme4 package. *R package
661 version*, 2(1):74.
- 662 Bruno, J. F., Ellner, S. P., Vu, I., Kim, K., and Harvell, C. D. (2011). Impacts of aspergillosis
663 on sea fan coral demography: modeling a moving target. *Ecological Monographs*,
664 81(1):123–139.
- 665 Compagnoni, A., Bibian, A. J., Ochocki, B. M., Rogers, H. S., Schultz, E. L., Sneck, M. E.,
666 Elderd, B. D., Iler, A. M., Inouye, D. W., Jacquemyn, H., et al. (2016). The effect of
667 demographic correlations on the stochastic population dynamics of perennial plants.
668 *Ecological Monographs*, 86(4):480–494.
- 669 Conn, P. B., Johnson, D. S., Williams, P. J., Melin, S. R., and Hooten, M. B. (2018). A guide
670 to bayesian model checking for ecologists. *Ecological Monographs*, 88(4):526–542.
- 671 Cooch, E. G. and White, G. C. (2020, accessed 5/17/2020). *Program MARK - a 'gentle
672 introduction'*. Available at phidot.org.
- 673 Coulson, T. (2012). Integral projections models, their construction and use in posing
674 hypotheses in ecology. *Oikos*, 121(9):1337–1350.
- 675 Cribari-Neto, F. and Zeileis, A. (2010). Beta regression in r. *Journal of statistical software*,
676 34:1–24.
- 677 Crone, E. E. (2016). Contrasting effects of spatial heterogeneity and environmental
678 stochasticity on population dynamics of a perennial wildflower. *Journal of Ecology*,
679 104(2):281–291.
- 680 Czachura, K. and Miller, T. E. (2020). Demographic back-casting reveals that subtle
681 dimensions of climate change have strong effects on population viability. *Journal of
682 Ecology*.
- 683 D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g1.
684 *Biometrika*, pages 679–681.
- 685 Davis, C. (2015). *sgt: Skewed Generalized T Distribution Tree*. R package version 2.0.

- 686 Drees, T., Ochocki, B. M., Collins, S. L., and Miller, T. E. (2023). Demography and
 687 dispersal at a grass-shrub ecotone: a spatial integral projection model for woody plant
 688 encroachment. *Ecological Monographs*, page e1574.
- 689 Easterling, M. R., Ellner, S. P., and Dixon, P. M. (2000). Size-specific sensitivity: applying
 690 a new structured population model. *Ecology*, 81(3):694–708.
- 691 Elderd, B. D. and Miller, T. E. (2016). Quantifying demographic uncertainty: Bayesian
 692 methods for integral projection models. *Ecological Monographs*, 86(1):125–144.
- 693 Ellis, M. M. and Crone, E. E. (2013). The role of transient dynamics in stochastic popula-
 694 tion growth for nine perennial plants. *Ecology*, 94(8):1681–1686.
- 695 Ellner, S. P., Adler, P. B., Childs, D. Z., Hooker, G., Miller, T. E., and Rees, M. (2022).
 696 A critical comparison of integral projection and matrix projection models for demo-
 697 graphic analysis: Comment. *Ecology*, 103(10):e3605.
- 698 Ellner, S. P., Childs, D. Z., and Rees, M. (2016). *Data-driven Modeling of Structured Popula-
 699 tions: A Practical Guide to the Integral Projection Model*. Springer, New York.
- 700 Fasiolo, M., Wood, S. N., Zaffran, M., Nedellec, R., and Goude, Y. (2020). qgam: Bayesian
 701 non-parametric quantile regression modelling in r. *arXiv preprint arXiv:2007.03303*.
- 702 Gould, W. R. and Nichols, J. D. (1998). Estimation of temporal variability of survival in
 703 animal populations. *Ecology*, 79:2531 – 2538.
- 704 Gu, C. (2013). *Smoothing Spline ANOVA Models*. Springer Science+Business Media, New
 705 York, 2 edition.
- 706 Hadfield, J. D. et al. (2010). Mcmc methods for multi-response generalized linear mixed
 707 models: the mcmcglmm r package. *Journal of Statistical Software*, 33(2):1–22.
- 708 Héault, B., Bachelot, B., Poorter, L., Rossi, V., Bongers, F., Chave, J., Paine, C. T., Wagner,
 709 F., and Baraloto, C. (2011). Functional traits shape ontogenetic growth trajectories of
 710 rain forest tree species. *Journal of ecology*, 99(6):1431–1440.
- 711 Jones, M. and Pewsey, A. (2009). Sinh-arcsinh distributions. *Biometrika*, 96:761 – 780.
- 712 Jones, M. C., Rosco, J. F., and Pewsey, A. (2011). Skewness-invariant measures of kurtosis.
 713 *The American Statistician*, 65(2):89 – 95.

- 714 Komsta, L. and Novomestky, F. (2015). Moments, cumulants, skewness, kurtosis and
715 related tests. *R package version*, 14(1).
- 716 Link, W. A. and Nichols, J. D. (1994). On the importance of sampling variance to inves-
717 tigations of temporal variation in animal population size. *Oikos*, 69(3):539 – 544.
- 718 Louthan, A. M., Keighron, M., Kiekebusch, E., Cayton, H., Terando, A., and Morris, W. F.
719 (2022). Climate change weakens the impact of disturbance interval on the growth rate
720 of natural populations of venus flytrap. *Ecological Monographs*, 92(4):e1528.
- 721 McGillivray, H. (1986). Skewness and asymmetry: measures and orderings. *Annals of
722 Statistics*, 14:994–1011.
- 723 Metcalf, C. J. E., Ellner, S. P., Childs, D. Z., Salguero-Gómez, R., Merow, C., McMahon,
724 S. M., Jongejans, E., and Rees, M. (2015). Statistical modelling of annual variation for
725 inference on stochastic population dynamics using Integral Projection Models. *Methods
726 in Ecology and Evolution*, 6:1007–1017.
- 727 Metcalf, C. J. E., Graham, A. L., Martinez-Bakker, M., and Childs, D. Z. (2016). Oppor-
728 tunities and challenges of i ntegral p rojection m odels for modelling host-parasite
729 dynamics. *Journal of Animal Ecology*, 85(2):343–355.
- 730 Metcalf, J. C., Rose, K. E., and Rees, M. (2003). Evolutionary demography of monocarpic
731 perennials. *Trends in Ecology & Evolution*, 18(9):471–480.
- 732 Miller, T. E. (2020). Long-term study of tree cholla demography in the los pinos
733 mountains, sevilleta national wildlife refuge. [https://doi.org/10.6073/pasta/
734 dd06df3f950afe4a4642306182237d13](https://doi.org/10.6073/pasta/dd06df3f950afe4a4642306182237d13).
- 735 Miller, T. E., Louda, S. M., Rose, K. A., and Eckberg, J. O. (2009). Impacts of insect
736 herbivory on cactus population dynamics: experimental demography across an envi-
737 ronmental gradient. *Ecological Monographs*, 79(1):155–172.
- 738 Miller, T. E., Williams, J. L., Jongejans, E., Brys, R., and Jacquemyn, H. (2012). Evolution-
739 ary demography of iteroparous plants: incorporating non-lethal costs of reproduction
740 into integral projection models. *Proceedings of the Royal Society B: Biological Sciences*,
741 279(1739):2831–2840.
- 742 Ochocki, B. M., Drees, T., and Miller, T. E. (2023). Density-dependent demography of
743 creosote bush (*larrea tridentata*) along grass-shrub ecotones. <https://doi.org/10.6073/pasta/ca53c16f16dcf9fb11f3ee99ea5445ac>.

- 745 Ohm, J. R. and Miller, T. E. (2014). Balancing anti-herbivore benefits and anti-pollinator
746 costs of defensive mutualists. *Ecology*, 95(10):2924–2935.
- 747 Ozgul, A., Childs, D. Z., Oli, M. K., Armitage, K. B., Blumstein, D. T., Olson, L. E., Tul-
748 japurkar, S., and Coulson, T. (2010). Coupled dynamics of body mass and population
749 growth in response to environmental change. *Nature*, 466(7305):482–485.
- 750 Peterson, M. L., Morris, W., Linares, C., and Doak, D. (2019). Improving structured
751 population models with more realistic representations of non-normal growth. *Methods
752 in Ecology and Evolution*, 10(9):1431–1444.
- 753 Plard, F., Schindler, S., Arlettaz, R., and Schaub, M. (2018). Sex-specific heterogene-
754 ity in fixed morphological traits influences individual fitness in a monogamous bird
755 population. *The American Naturalist*, 191(1):106–119.
- 756 Rees, M., Childs, D. Z., and Ellner, S. P. (2014). Building integral projection models: a
757 user's guide. *Journal of Animal Ecology*, 83(3):528–545.
- 758 Rees, M. and Ellner, S. P. (2016). Evolving integral projection models: evolutionary
759 demography meets eco-evolutionary dynamics. *Methods in Ecology and Evolution*,
760 7(2):157–170.
- 761 Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z., and De Bastiani, F. (2019). *Distributions
762 for modeling location, scale, and shape: Using GAMLS in R*. CRC press.
- 763 Salguero-Gómez, R. and Casper, B. B. (2010). Keeping plant shrinkage in the demo-
764 graphic loop. *Journal of Ecology*, 98(2):312–323.
- 765 Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky,
766 C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., and Araya-Ajoy, Y. G. (2020).
767 Robustness of linear mixed-effects models to violations of distributional assumptions.
768 *Methods in ecology and evolution*, 11(9):1141–1152.
- 769 Schultz, E. L., Eckberg, J. O., Berg, S. S., Louda, S. M., and Miller, T. E. (2017). Native
770 insect herbivory overwhelms context dependence to limit complex invasion dynamics
771 of exotic weeds. *Ecology letters*, 20(11):1374–1384.
- 772 Shriver, R. K., Cutler, K., and Doak, D. F. (2012). Comparative demography of an epi-
773 phytic lichen: support for general life history patterns and solutions to common prob-
774 lems in demographic parameter estimation. *Oecologia*, 170:137–146.

- 775 Stasinopoulos, D. M., Rigby, R. A., et al. (2007). Generalized additive models for location
776 scale and shape (gamlss) in r. *Journal of Statistical Software*, 23(7):1–46.
- 777 Stubberud, M. W., Vindenes, Y., Vøllestad, L. A., Winfield, I. J., Stenseth, N. C., and Lan-
778 gangen, Ø. (2019). Effects of size-and sex-selective harvesting: An integral projection
779 model approach. *Ecology and Evolution*, 9(22):12556–12570.
- 780 Wan, X., Wang, W., Liu, J., and Tong, T. (2014). Estimating the sample mean and stan-
781 dard deviation from the sample size, median, range and/or interquartile range. *BMC*
782 *medical research methodology*, 14:1–13.
- 783 Wilber, M. Q., Langwig, K. E., Kilpatrick, A. M., McCallum, H. I., and Briggs, C. J. (2016).
784 Integral projection models for host–parasite systems with an application to amphibian
785 chytrid fungus. *Methods in Ecology and Evolution*, 7(10):1182–1194.
- 786 Williams, J. L., Miller, T. E., and Ellner, S. P. (2012). Avoiding unintentional eviction from
787 integral projection models. *Ecology*, 93(9):2008–2014.
- 788 Wood, S. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and
789 Hall/CRC, 2 edition.
- 790 Zhang, J. L. (2014). Comparative investigation of three bayesian p values. *Computational*
791 *Statistics & Data Analysis*, 79:277–291.

Appendices

792 S.1 The Jones-Pewsey distribution

793 Jones and Pewsey (2009) introduced a simple, tractable generalization of the Normal dis-
794 tribution with two additional parameters determining asymmetry (skewness), and tail
795 weight (kurtosis) which can be either lighter or heavier than the Gaussian. It is defined
796 as a transformation of a Normal(0,1) random variable using the hyperbolic sine func-
797 tion (\sinh) and its inverse (asinh), as follows. The distribution family's base probability
798 density $f_{\epsilon,\delta}$ is the probability density of the random variable $X_{\epsilon,\delta}$ where

799

$$Z = \sinh(\delta \text{ asinh}(X_{\epsilon,\delta}) - \epsilon) \quad (\text{S.1})$$

800 and Z has a Normal(0,1) distribution. Equivalently,

801

$$X_{\epsilon,\delta} = \sinh\left(\frac{1}{\delta} \text{ asinh}(Z) + \frac{\epsilon}{\delta}\right). \quad (\text{S.2})$$

802 Parameters $\delta = 1, \epsilon = 0$ give the Normal(0,1) distribution. Skewness has the sign of ϵ ,
803 and $\delta > 0$ controls tail weight, with heavier than Gaussian tails for $\delta < 1$ and lighter
804 than Gaussian tails for $\delta > 1$. A formula for the density $f_{\epsilon,\delta}$ is given by Jones and Pewsey
805 (2009, eqn. 2). The general four-parameter family with location parameter μ and scale
806 parameter σ is defined as the probability densities of $\mu + \sigma X_{\epsilon,\delta}$. We refer to this as the
807 JP distribution family.

808 As is unfortunately the case for most four-parameter distributions μ is not the mean,
809 σ is not the standard deviation, ϵ is not the skew and δ is not the kurtosis. All else being
810 equal, larger μ gives a larger mean, larger σ gives a higher standard deviation, higher
811 ϵ gives higher asymmetry, and higher δ gives heavier tail weight. But each moment is
812 jointly determined by all four parameters.

813 The main advantage of the JP distribution is that the attainable combinations of
814 skewness and kurtosis are very broad, compared to other four-parameter families, and
815 come very close to the theoretical limits on kurtosis as a function of skewness (Jones and
816 Pewsey, 2009, Fig. 2). Additionally, being a transformation of the Normal makes it very
817 simple to generate random numbers from the distribution, and to compute probability
818 density, cumulative distribution, and quantile functions. There are also simple analytic
819 formulas for the first four moments (Jones and Pewsey, 2009, p. 764) which we use below
820 to define a centered and scaled version in which μ and σ are the mean and standard
821 deviation.

822 The definition (S.2) shows that the distribution depends on ϵ only through the ratio
823 ϵ/δ . We have found that this property can be problematic for estimating distribution
824 parameters. Even with good sized ($n = 250$ or 500) data sets generated from the distri-
825 bution with known parameters, both maximum likelihood and Bayesian estimation were
826 unstable for some values of ϵ and δ , occasionally yielding estimates far from the truth.
827 One cause was a ridge in the (ϵ, δ) likelihood surface with a constant of ϵ/δ . Another is
828 that when δ is large, changes in ϵ have little effect.

829 To avoid that problems, we reparameterize the distribution as follows:

830
$$X_{\lambda, \tau} = \sinh(e^{-\tau} \operatorname{asinh}(Z) + \lambda). \quad (\text{S.3})$$

831 Thus, the two parameterizations are related by

832
$$\delta = e^\tau, \epsilon = \delta\lambda = e^\tau\lambda. \quad (\text{S.4})$$

833 The definition of τ allows it to take any real value, with negative values giving thinner
834 than Gaussian tails and positive values giving fatter than Gaussian tails. λ also can take
835 any real value, and the distribution's skew has the same sign as λ . Because the sinh
836 function is nonlinear, it is still the case that the skew depends on τ as well as λ , but the
837 "crosstalk" between the kurtosis and skew parameters is weaker. As a result, we found
838 that maximum likelihood estimation of parameter values was generally more reliable if
839 the distribution is parameterized in terms of τ and λ .

840 S.2 Estimating mixed-effects models using shrinkage

841 Ecologists often fit demographic and other statistical models that include random effects
842 terms to quantify variation among years, spatial locations, individuals, etc. Random
843 effects are a natural choice when interest centers on the magnitude of variation (e.g., how
844 much does mortality vary among years?) rather than individual values (e.g., mortality
845 in 2013). They also allow each estimate to "borrows strength" from others, so that (for
846 example) the estimate from a year with small sample size (and thus large sampling
847 variability) is shifted towards the center of the overall distribution.

848 Specialized software is often used to fit such models, such as the **nlme**, **lme4**, **mgcv**
849 and **gamm4** libraries in R, but these only allow a small subset of the distribution families
850 we want to consider for modeling growth increments (the **gamLss** package allows many
851 distribution families, but in our experience, even when random effects are simple in
852 structure the fitting algorithms often fail to converge or fail to find the global optimum).

853 One way past this limitation is Bayesian estimation, using STAN with user-written
854 (or borrowed) code for the chosen growth distribution (see section XX for an example).
855 In this appendix we describe another option, introduced by Link and Nichols (1994)
856 and Gould and Nichols (1998): fitting a fixed-effects model by Maximum Likelihood,
857 followed by shrinkage of coefficient estimates. None of the ideas here are original. The
858 material overlaps Appendix S1 of Metcalf et al. (2015), but for completeness we make
859 it self-contained. Appendix D of Cooch and White (2020) (written by K.D. Burnham)
860 provides more details and examples in the context of capture-recapture analysis.

861 Here we explain shrinkage using a simple model based on our analysis of *Pseu-*
862 *doroegneria spicata*. That model includes random effects for between-year variation in
863 the slope and intercept of future size (log area) as a function of initial size. To keep
864 the example simple, we assume that initial size and year are the only covariates, and
865 we assume that growth increments follow a skew-Normal distribution with noncon-
866 stant variance and constant skew parameter. Code for this example is in the script
867 `SimpleShrinkageExample.R`. The first part of the script generates an artificial data set
868 by fitting the model to a subset of the growth data (20th century Control plots), and
869 randomly generating new “size next year” values for each individual in the actual data
870 set. The second part contains the “data” analysis.

871 As in our *P. spicata* analysis, we assumed that that the skew and kurtosis parameters
872 were functions of the location parameter; this dominated ($\Delta AIC \approx 30$) the alternate
873 model with skew and kurtosis depending on initial size. The analogous Gaussian model,
874 with constant variance, could be fitted as follows using `lmer`:

875 `lmer(new.size ~ init.size + (init.size|year), data=growthData, REML=TRUE);`
876 where `growthData` is a data frame holding the data with year as an unordered factor.
877 For our skew-Normal model, we instead use maximum likelihood with all between-year
878 variation included as fixed effects. The appropriate design matrix is easily constructed
879 using the `model.matrix` function:

880 `U = model.matrix(~ year + init.size:year - 1, data=growthData)`

881 If there are T years, the matrix `U` specified in this way has $2T$ columns corresponding to
882 n annual intercepts and T annual slopes.

883 Using this design matrix, we can readily write a log likelihood function for use with
884 the `maxLik` package, with a log link function for the variance because it is necessarily
885 positive:

886 `LogLik=function(pars,new.size,U){`

```

887 pars1 = pars[1:ncol(U)]; pars2=pars[-(1:ncol(U))];
888 mu = U%*%pars1;
889 sigma = exp(pars2[1]+pars2[2]*mu);
890 dSN1(new.size, mu=mu, sigma=sigma, nu=pars2[3], log=TRUE)
891 }

```

892 Parameters and their standard errors can then be estimated with `maxLik`, starting
893 from a random guess:

```

894 start=c(runif(ncol(U)), rep(0,3))
895 out=maxLik(logLik=LogLik,start=start, new.size=simData$new.size,U=U,
896 method="BHHH",control=list(iterlim=5000,printLevel=1),finalHessian=TRUE);
897 coefs = out$estimate; # parameters
898 V = vcov(out); SEs = sqrt(diag(V)); # standard errors

```

899 In real life we would repeat the optimization several times with several different starting
900 values, to be confident that the optimal parameter values had been found.

901 Focus now on the year-specific intercept parameters $\hat{a}_t, t = 1, 2, \dots, T$. We can view
902 the year-specific estimates \hat{a}_t as consisting of unobserved true values a_t plus sampling
903 error:

$$904 \quad \hat{a}_t = a_t + \varepsilon_t \quad (S.5)$$

905 Because of the sampling errors, the sample variance of the estimates \hat{a}_t is an upward-
906 biased estimate of the true across-year variance in the parameter. That is undesirable if
907 the model will be used to project how temporal variability affects population dynamics.
908 However, maximum likelihood estimation gives us an approximate variance-covariance
909 matrix \hat{V} of the sampling errors, V in the code above. With that information, we can
910 estimate the parameters of a random effects model for the intercept parameters, and
911 thereby improve the year-specific estimates and the estimate of the across-year variance.

912 The model is as follows. We make the standard mixed-models assumptions that the
913 a_t are drawn independently from some fixed distribution with unknown variance σ^2 .
914 We also assume that the estimates \hat{a}_t are unbiased, that is

$$915 \quad \mathbb{E}(\varepsilon_t | a_t) = 0. \quad (S.6)$$

916 These are optimistic assumptions, but not excessively optimistic. Some degree of tem-
917 poral correlation will often be present, and as we explain at the end, it is theoretically
918 possible to account for it. Maximum likelihood parameter estimates are not unbiased,
919 but if the assumptions of maximum likelihood are satisfied the bias is asymptotically

920 negligible compared to the standard error (the bias scales as the inverse of sample size,
 921 the standard error as the square root of the inverse of sample size).

922 Let S^2 denote the sample variance of the estimates \hat{a}_t . It can then be shown that

$$923 \quad \mathbb{E}(S^2) = \sigma^2 + \frac{1}{T} \sum_{t=1}^T \mathbb{E}Var(\varepsilon_t) - \frac{1}{T(T-1)} \sum_{i=1}^{j-1} \sum_{j=1}^T \mathbb{E}Cov(\varepsilon_i, \varepsilon_j). \quad (\text{S.7})$$

924 This is eqn. (1) in Gould and Nichols (1998) in our notation, without the term that results
 925 from temporal autocorrelation.

926 The terms besides σ^2 on the right-hand are the expected impact of sampling error
 927 on the across-year variance of the parameter estimates; their presence makes S^2 a biased
 928 estimated of σ^2 . However, all of those terms correspond to entries in the variance-
 929 covariance matrix V . We can therefore use our estimated variance-covariance matrix \hat{V}
 930 to removes the bias due to sampling variability:

$$931 \quad \hat{\sigma}^2 = S^2 - \frac{1}{T} \sum_{t=1}^T \hat{V}_{t,t} + \frac{1}{T(T-1)} \sum_{i=1}^{j-1} \sum_{j=1}^T \hat{V}_{i,j}. \quad (\text{S.8})$$

932 $\hat{\sigma}^2$ estimates the variance of the distribution from which the a_t are assumed to be drawn.

933 Using that estimate, we can adjust the year-specific estimates to reduce the ex-
 934 pected impact of sampling error. Depending on your purposes, there are two possible
 935 adjustments. The first option is the one used in the popular capture-recapture analysis
 936 software Mark Cooch and White (2020),

$$937 \quad \tilde{a}_t = \bar{a}_t + \sqrt{\frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \hat{V}_{t,t}}} (\hat{a}_t - \bar{a}_t). \quad (\text{S.9})$$

938 The name “shrinkage” comes from the fact that each estimate is adjusted towards the
 939 overall mean, with larger adjustments of values that have higher estimated sampling
 940 error variance, $\hat{V}_{t,t}$. This shrinkage estimate has the property that the expected sample
 941 variance of the adjusted estimates \tilde{a}_t is very close to $\hat{\sigma}^2$, so the \tilde{a}_t approximate the actual
 942 amount of parameter variation.

943 The second is to replace \hat{a}_t by the least-squares estimate of a_t under the additional
 944 assumption that the a_t are drawn from a Gaussian distribution; this is given by

$$945 \quad \tilde{a}_t = \bar{a}_t + \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \hat{V}_{t,t}} (\hat{a}_t - \bar{a}_t). \quad (\text{S.10})$$

946 This option is theoretically preferable if the Gaussian assumption is reasonable, and you
 947 are more interested in year-specific values rather than across-year variance. However,
 948 Metcalf et al. (2015) found that even (S.9), which does less shrinkage, resulted in a small
 949 downward bias in the temporal variance of population growth rates. This argues for
 950 always using the first option, and we do the same here.

951 We differ from MARK, however, in using (S.8) rather than an iterative method
 952 that takes (S.8) as its starting estimate and refines the estimate by using weighted least
 953 squares based on the current estimate. Metcalf et al. (2015) found, in simulation studies,
 954 that the iterative method was either slightly beneficial or wildly inaccurate. We therefore
 955 advise against it.

956 Finally, as mentioned above, the estimate of σ^2 can account for temporal autocor-
 957 relation in the a_t . When present, those correlations add a term to eqn. (S.7) (see eqn.
 958 (1) in Gould and Nichols (1998)), which can be estimated from the sample autocorre-
 959 lation of the \hat{a}_t . We do not recommend doing this (and therefore omit the formulas)
 960 because the autocorrelations can only be reliably estimated if they fall to nearly zero
 961 within lag $m \ll T$, in which case the autocorrelation term is small (specifically, $O(m/T)$).
 962 Otherwise, the random error from using poorly estimated autocorrelations is likely to
 963 outweigh the small bias from omitting that term.

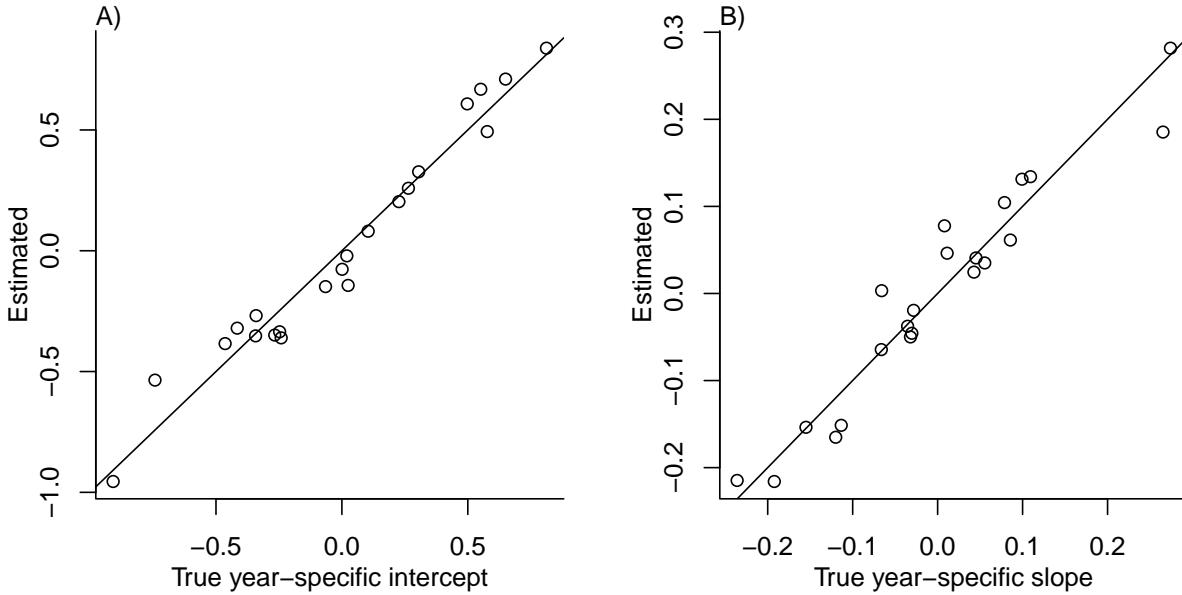
964 The take-home message is that estimating random effects from the regression coef-
 965 ficients is very simple:

```

966 # Variance-covariance matrices for intercepts and slopes
967 V1 = V[1:T,1:T]; V2 = V[(T+1):(2*T),(T+1):(2*T)];
968 # Extract year-specific intercepts, center them to zero
969 fixed.fx = coefs[1:T]; fixed.fx = fixed.fx-mean(fixed.fx);
970
971 # Estimate sigma^2
972 var.hat = mean(fixed.fx^2) - mean(diag(V1)) +
973           (sum(V1)-sum(diag(V1)))/(2*T*(T-1));
974
975 # Shrink deviations from the mean
976 shrinkRanIntercept = fixed.fx*sqrt(var.hat/(var.hat + diag(V1)));
977
978 # Do it all again for the slopes
979 fixed.fx2 = coefs[(T+1):(2*T)]; fixed.fx2 = fixed.fx2-mean(fixed.fx2);
980 var2.hat = mean(fixed.fx2^2) - mean(diag(V2)) +
981           (sum(V2)-sum(diag(V2)))/(2*T*(T-1));
  
```

```
982 shrinkRanSlope = fixed.fx2*sqrt(var2.hat/(var2.hat + diag(V2)));
```

983 The figure below shows the results for one artificial PSSP “data” set, having $T = 22$
984 years and growth measurements on about 175 individuals/year on average. The true
985 random year effects (the ones used to generate the data) are recovered with good accu-
986 racy and no bias. In particular there is no sign of extreme values being pulled in too
987 far towards the mean, which would cause an S-shaped graph of estimated versus true
988 values.



989 S.3 Additional case studies

990 S.3.1 Case study: Sea fan corals, *Gorgonia ventalina*

991 Bruno et al. (2011) developed an IPM to understand the rise and fall of a fungal pathogen
992 *Aspergillus sydowii* in Caribbean sea fan corals *G. ventalina*. The model was based on re-
993 peated observations of marked corals in permanent transects at several sites near Aku-
994 mal, Mexico, recording disease status (infected/uninfected) and the area of uninfected
995 tissue. The epidemic peak had passed and disease incidence was already low, so in-
996 fected fans were relatively infrequent. We therefore limit the analysis here to uninfected
997 individuals. Bruno et al. (2011) found statistically significant year and site effects, but
998 as those explained a very small fraction of the variation in growth increments, they
999 fitted a single growth model to data pooled across years and sites. We do the same
1000 here. The pooled data set consists of 358 observed size transitions. The data exhibited

1001 size-dependent variance in growth (change in area, cm^2). Bruno et al. (2011) chose to sta-
1002 bilize the variance by cube-root transforming size, and then fitting the standard model
1003 with Gaussian growth increments. Here we take a different approach, using natural log
1004 transformation of area and modeling size-dependent variance.

1005 With initial size as the only predictor, a simple way to fit a Gaussian model with
1006 nonconstant variance is the `gam` function in `mgcv` library (Wood, 2017) using the `gaulss`
1007 family. The mean and standard deviation are both fitted as smoothing spline functions
1008 of initial size, and the `predict` function returns the fitted mean and also the inverse of
1009 the fitted standard deviations with which we can compute the scaled residuals:

```
1010 # XH is a data frame holding the data
1011 # logarea.t0, .t1 denote initial and final values of log-transformed area
1012 fitGAU <- gam(list(logarea.t1~ s(logarea.t0), ~ s(logarea.t0)),
1013   data=XH, gamma=1.4, family=gaulss())
1014 fitted_all = predict(fitGAU,type="response");
1015 fitted_sd = 1/fitted_all[,2];
1016 scaledResids = residuals(fitGAU,type='response')/fitted_sd;
```

1017 Fig. S-1A shows the log-transformed data and Gaussian model. The mean function
1018 (solid red curve) is visually nearly linear, but the fitted spline is strongly favored over a
1019 linear model for the mean ($\Delta AIC \approx 9$). The spline for standard deviation σ versus initial
1020 size reflects the evident greater variability in growth at smaller sizes.

1021 There are no blatant signs of trouble in the pilot Gaussian model, but quantile re-
1022 gressions on the scaled residuals, and the NP Skewness and Kurtosis metrics derived
1023 from them (Eq. 3 and 4), suggest deviations from normality (Fig. S-1B). Specifically,
1024 skewness switches from negative to positive across the size range, with smaller corals
1025 more prone to extreme shrinkage and larger corals more prone to extreme growth. Kur-
1026 tosis also changes direction over the size distribution, with thinner tails than Gaussian
1027 at small sizes and fatter tails at large sizes. The fitted nonparametric moments suggest
1028 that the upper and lower tails of size transition probabilities may differ by up to 20%,
1029 and the weight of the tails may be 20% greater or less than Gaussian, depending on
1030 initial size – not overwhelming deficiencies, but not trivial either. Are these deviations
1031 from normality severe enough to warrant a second, non-Gaussian iteration of growth
1032 modeling? To answer that question, we simulated data from the fitted Gaussian model
1033 and examined whether key properties of the simulated data are consistent with those
1034 of the real data – this is the ultimate litmus test for a growth model's adequacy and
1035 should be a standard element of IPM construction, in our opinion. If the simulated data

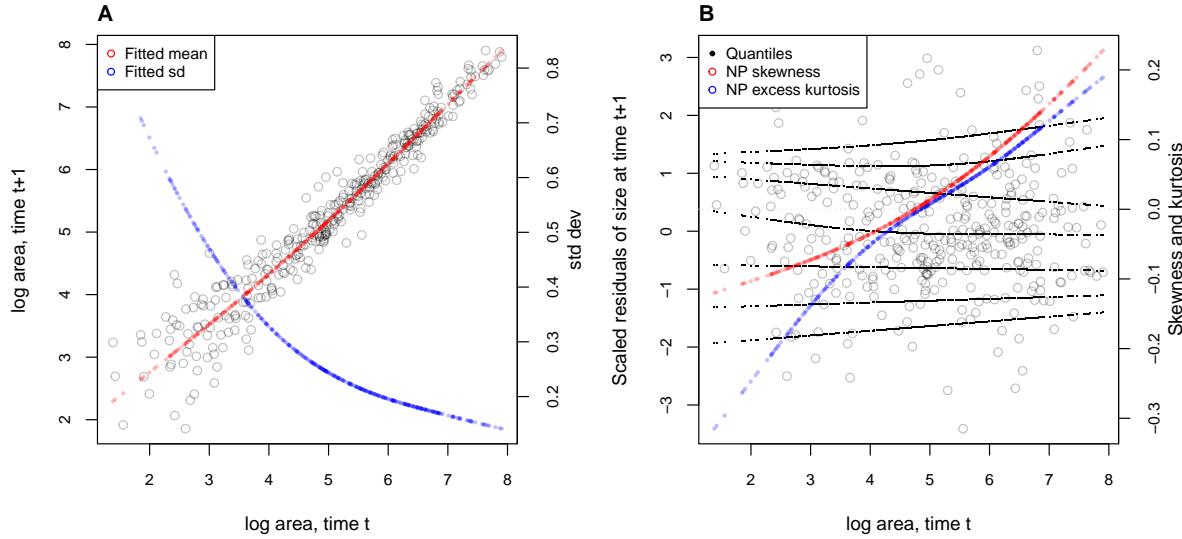


Figure S-1: **A**, Size transition data for sea fan corals, *Gorgonia ventalina*, and fitted gam with mean (red) and standard deviation (blue) of future size conditional on current size. **B**, Quantile regressions of scaled residuals on size and nonparametric estimates of skewness (red) and excess kurtosis (blue) derived from them. Black lines in **B** show the 5th, 10th, 25th, 50th, 75th, 90th, and 95th quantiles. Figure made by script AkumalCorals_qgam.R.

1036 are not consistent with the real data, it is time to choose a better distribution (Fig. 1).
 1037 In this case, most of 100 Gaussian model simulations are out of line with the skew at
 1038 smallest and largest sizes, and excess kurtosis observed at moderately large sizes (Fig.
 1039 S-2 CD). For at least some parts of the size distribution, a non-Gaussian model would
 1040 better capture size transitions.

1041 We sought a distribution that could accommodate the observed changes in the sign
 1042 of skewness and excess kurtosis. We chose the sinh-arcsinh (SHASH) distribution, a
 1043 four-parameter distribution that, conveniently, is included in **mgcv**'s **gam()** function.
 1044 For consistency with the Gaussian for location and scale, specification of basis functions
 1045 ($k = 4$) is limited to parameters for skewness and kurtosis:

```
1046 fitSHASH <- gam(list(logarea.t1 ~ s(logarea.t0), # <- location
1047   ~ s(logarea.t0), # <- log-scale
1048   ~ s(logarea.t0,k=4), # <- skewness
1049   ~ s(logarea.t0,k=4)), # <- log-kurtosis
1050   data = XH, gamma = 1.4, family = shash, optimizer = "efs")
```

1051 The fitted model's mean and variance are nearly identical to the Gaussian (Fig. S-2AB),
 1052 and the fitted trends in skewness and kurtosis are much less "wiggly" than the estimate

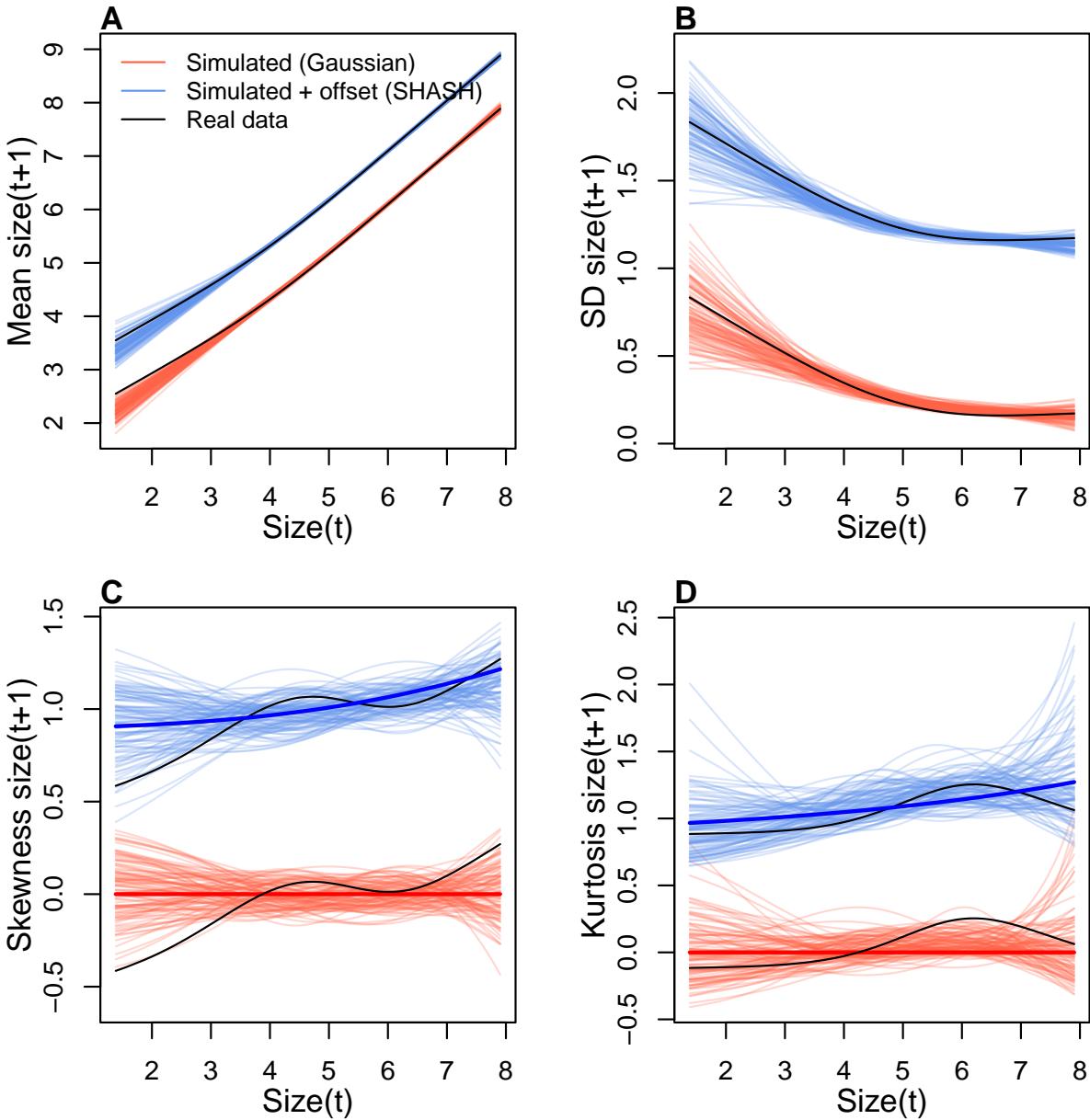


Figure S-2: Comparisons among real coral data and data simulated from Gaussian and SHASH growth models for mean, standard deviation, NP skewness, and NP excess kurtosis of future size conditional on current size. Note that plotted values for the SHASH are offset by one unit to allow comparisons. In the skewness and kurtosis panels, the darker solid curves show the values for the fitted growth models. Figure made by script AkumalCorals_qgam.R.

from the data (Fig. S-2CD). Nonetheless, data simulated from the SHASH model are more consistent with the real data, with more SHASH data sets matching or exceeding the largest skewness and kurtosis values observed (Fig. S-2CD). If one cares to quantify

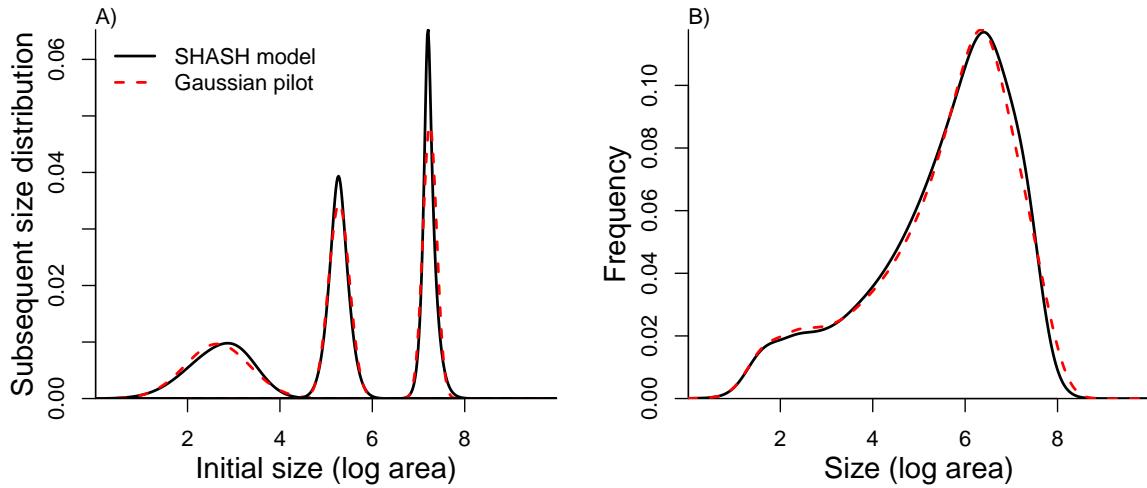


Figure S-3: Comparisons between the fitted SHASH growth model (solid black curves) and the Gaussian pilot model (dashed red curves) for sea fans *G. ventalina*. A) Predicted frequency distributions of size in year $t + 1$ for three different values of size in year t . The leftmost pair of curves are for initial size equal to the median size of a new recruit (data from (Bruno et al., 2011)); central pair is for the median initial size of uninfected individuals; rightmost pair are for the 95th percentile initial size for uninfected individuals. B) Steady-state size distributions resulting from a constant unit input of new recruits. As in Bruno et al. (2011) we assume that larvae are very widely dispersed, so that the number of recruits arriving at any one location is independent of the local population abundance or structure. Size distribution of recruits was described by a kernel density estimate based on the measured sizes of known new recruits ($n = 9$). Figure made by script AkumalCoralsIPMs.R.

1056 the difference between models, the SHASH model is clearly favored by AIC ($\Delta AIC =$
 1057 5.45) despite having twice as many parameters to fit.

1058 What, then, have we gained by fitting a better growth model? Fig. S-3A compares
 1059 the predicted distributions of subsequent size in the fitted model and Gaussian pilot
 1060 models, for the median size of a new recruit (leftmost pair of curves), the median initial
 1061 size (central curves), and the 95th percentile of initial size in the data (rightmost
 1062 curves). The differences are small, and most pronounced for the smallest size, where recruits
 1063 are predicted to grow slightly larger under the SHASH model than the Gaussian
 1064 model. The direction of this difference was surprising, because the SHASH has negative
 1065 skew at small sizes in the data. However, the SHASH model also gives a better prediction
 1066 of mean growth at small sizes than the Gaussian model. At intermediate sizes the
 1067 predictions are nearly identical; at large sizes the SHASH has slightly lower standard
 1068 deviation, but fatter tails (excess kurtosis). Fig. S-3B shows the predicted steady-state
 1069 size distributions resulting from a constant unit input of recruits. Again, the differences

1070 are very subtle. Finally, the Gaussian and SHASH growth models predict very similar
1071 mean life span (17.7 and 17.9 years, respectively).

1072 From these outputs, there is little evidence that improved modeling of coral growth
1073 meaningfully improved biological inferences from the IPM. One could argue that it was
1074 not worth the trouble, even though it was almost no trouble at all. But before fitting
1075 the SHASH model, we could not have known whether or not it would have made a
1076 difference.

1077 In this case study we used `gam` to fit both the Gaussian and SHASH models because
1078 that obviated model selection on functions for mean, variance, and higher moments.
1079 However, `gam` should be used with caution. Nonparametric regression models notori-
1080 ously “wag their tails” because the ends of the fitted curve can be pulled close to the
1081 outermost data points. This is especially problematic for growth modeling, because data
1082 are typically sparse near the bounds of the size distribution. To minimize the risk of
1083 overfitting we specified the number of “knots” (`k=4`) and used `gamma=1.4` to overweight
1084 model degrees of freedom as suggested by Gu (2013, sec. 3.2). But it is always impor-
1085 tant to plot the fitted splines and make sure they do not wag unrealistically. If they do,
1086 parametric regression may be a better choice.

1087 S.3.2 Case study: creosotebush, *Larrea tridentata*

1088 Our next case study comes from our studies of the woody shrub creosotebush (*Larrea tri-*
1089 *dentata*) at the Sevilleta Long-Term Ecological Research (LTER) site in central New Mex-
1090 ico, US. At this site as elsewhere in the Southwest US, creosotebush is encroaching into
1091 desert grassland habitats. The data described here were collected along transects span-
1092 ning grass-shrub ecotones to understand patterns of density dependence in creosotebush
1093 demography. Specifically, we asked whether fitness is maximized approaching zero den-
1094 sity at the leading edge of the expansion front (consistent with ‘pulled’ expansion), or
1095 whether there is a demographic advantage for shrubs at higher density due to positive
1096 feedbacks expected for ecosystem engineers (leading to ‘pushed’ expansion). Our pub-
1097 lished study (Drees et al., 2023) used a spatial integral projection model (SIPM) to predict
1098 the speed of shrub encroachment, assuming normally-distributed size transitions. Here
1099 we step through our suggested workflow to ask whether a non-Gaussian model would
1100 have been more faithful to the data, and how such an improvement would influence
1101 predictions for the speed of encroachment.

1102 Growth data come from 522 shrubs censused longitudinally over four years (2013-
1103 2017). Census individuals occurred along 12 replicate transects (200 to 600 m in length)

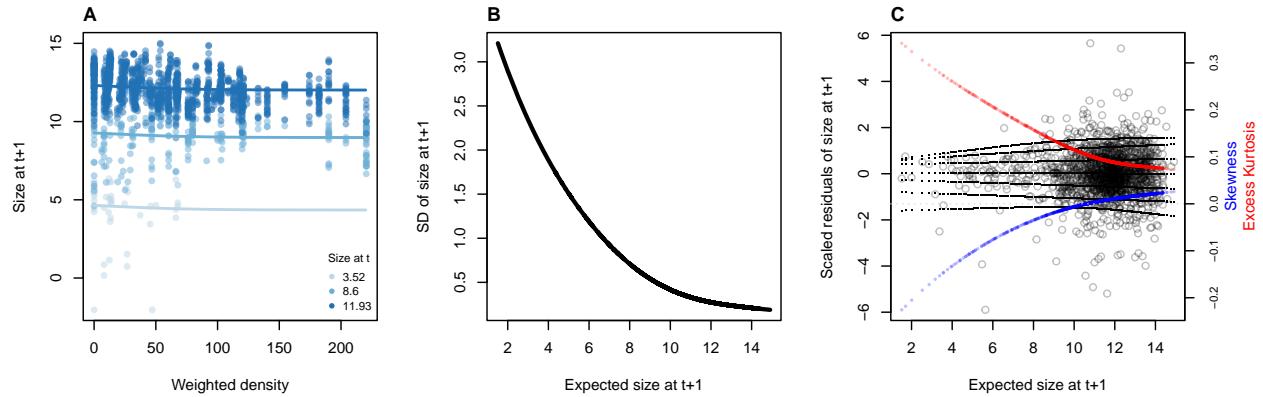


Figure S-4: **A**, Creosotebush size transition data with respect to initial size (colors) and local weighted density (sum of sizes of all plants within a five-meter transect window). Size is quantified as the natural logarithm of plant volume (cm^3). **B**, Standard deviation of size at time $t + 1$ as a function of expected size at $t + 1$ (the fitted values), estimated by iterative re-weighting. **C**, Quantile regressions of scaled residuals on size and nonparametric estimates of skewness (blue) and excess kurtosis (red) derived from them. Black lines in **C** show the 5th, 10th, 25th, 50th, 75th, 90th, and 95th quantiles. All figures made by script `creosote_growth_modeling.R`.

that spanned gradients of shrub density along shrub-grass ecotones. Size was measured as volume of an elliptical cone based on height and width measurements; the size variable of the IPM was the natural logarithm of volume (cm^3). For each census individual, we recorded the size and density of all conspecifics within the five-meter transect “window” in which it occurred, and took the sum of all sizes within the window as a weighted measure of local density. The data are available in Ochocki et al. (2023).

As an initial Gaussian approach, and following the approach of Drees et al. 2023, we first fit a generalized additive model with `mgcv` that included smooth terms for initial size and weighted density (constrained to four basis functions), plus the random effect of transect. We used the `gaulss` family and, as a starting point, fit a constant standard deviation.

```
1115 LATR_GAU <- gam(list(log_volume_t1~ s(log_volume_t,k=4) +
1116 s(dens_scaled,k=4) + s(unique.transect,bs="re"), ~ 1),
1117 family="gaulss", data=LATR_grow, method="ML", gamma=1.4)
```

Using the fitted values from this initial model, we updated the standard deviation to be a smooth function of fitted values, and iterated the fitting until the weights stopped changing, following the same steps as in the orchid case study.

The resulting Gaussian growth model predicts strong initial size-dependence and weak and slightly nonlinear (but monotonic) negative density dependence (Fig. S-4A).

1123 The model accounts for non-constant variance, which indicate greater dispersion for
1124 smaller values of expected size (Fig. S-4B). Quantiles of the standardized residuals indi-
1125 cate that skew and excess kurtosis are both greater at smaller sizes (Fig. S-4C). Skewness
1126 is close to zero for larger plants (the best-sampled size range) but excess kurtosis re-
1127 mains positive for large plants (ca. 10% heavier tails than Gaussian). As a candidate
1128 for improvement, we turned to the Johnson's S_U (JSU) distribution, a four-parameter,
1129 leptokurtic distribution capable of skew in either direction.

1130 Following our suggested workflow, rather than re-fitting a JSU model from scratch,
1131 we parameterize a model where the residuals from the Gaussian model are fitted by
1132 a JSU distribution. This is relatively easy because the **gamlss.dist** package provides a
1133 parameterization of the JSU in which the location parameter μ is the mean and scale
1134 parameter σ is the standard deviation (Rigby et al., 2019). We fit the "hybrid" model by
1135 writing a likelihood function that uses the fitted mean and standard deviation functions
1136 from Gaussian pilot model, and estimates the parameters that control skewness and
1137 kurtosis as linear functions of predicted future size. The "hybrid" likelihood looks like
1138 this:

```
1139 JSULogLik=function(pars){  
1140   dJSU(LATR_grow$log_volume_t1,  
1141     mu=LATR_grow$GAU_mean,  
1142     sigma=LATR_grow$GAU_sd,  
1143     nu = pars[1]+pars[2]*LATR_grow$GAU_mean,  
1144     tau = exp(pars[3]+pars[4]*LATR_grow$GAU_mean), log=TRUE)  
1145 }
```

1146 The mean and standard deviation of the JSU are set to those of the best Gaussian
1147 model and parameters controlling skewness and kurtosis were fit independently, follow-
1148 ing our approach to the orchid data. The hybrid JSU model performed well, generating
1149 simulated data that aligned with the real data better than the best Gaussian model, par-
1150 ticularly in the standard deviation and kurtosis (Fig. S-5). The JSU model has exactly the
1151 same mean and standard deviation of future size as the Gaussian, but Fig. S-4 uses the
1152 quantile-based nonparametric mean and standard deviation. The results show that even
1153 though the JSU was not fitted to match those, it comes closer than the Gaussian model
1154 as a result of accounting for the skew and kurtosis.

1155 The improvement of the JSU over the Gaussian growth model, while visually sat-
1156 isfying, had only weak influence on SIPM results. The Gaussian model slightly over-
1157 estimated the low-density growth rate, but models using either Gaussian or JSU growth

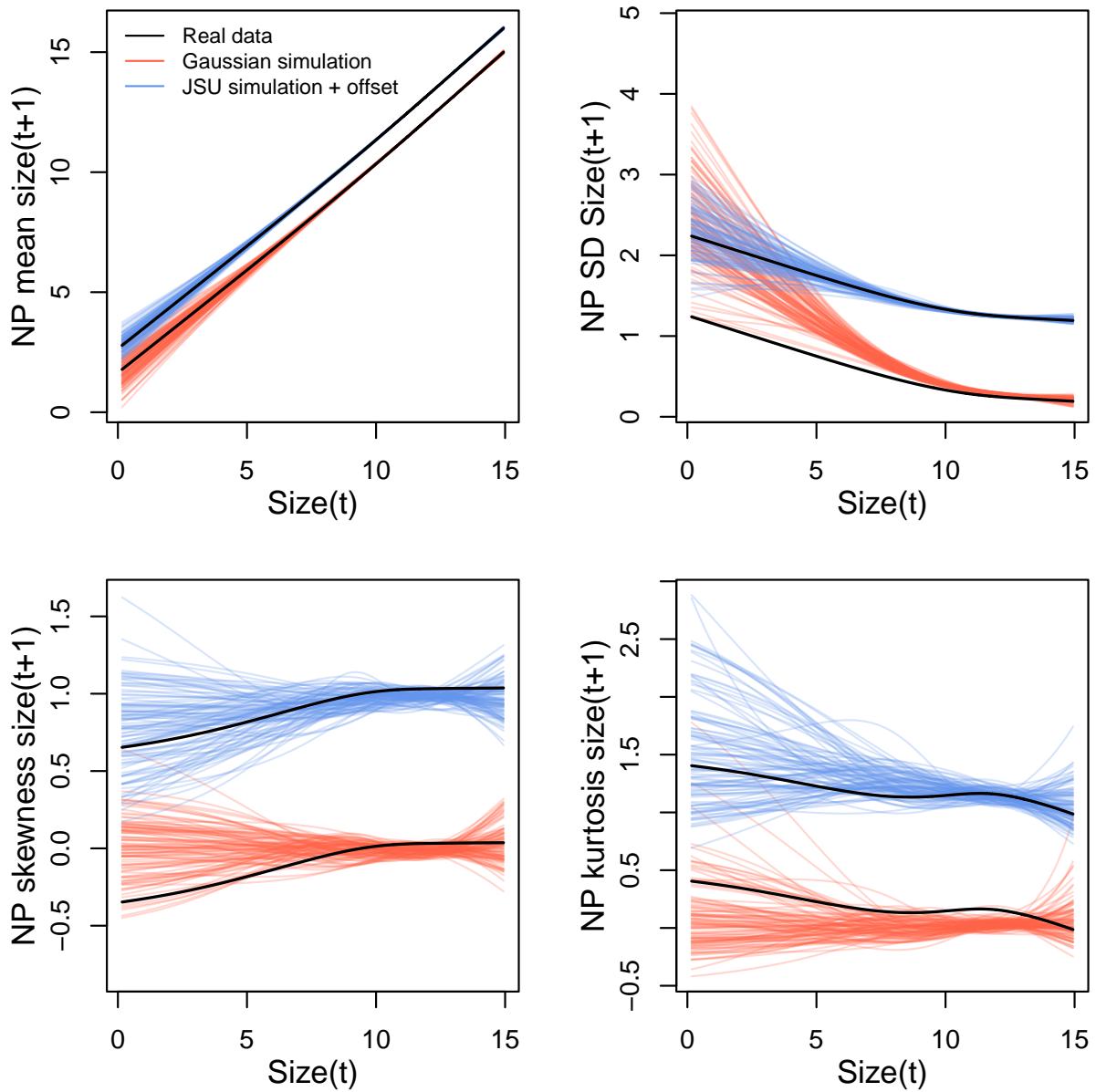


Figure S-5: Comparisons between real creosotebush data and data simulated from Gaussian and JSU growth models for nonparametric measures of mean, standard deviation, skewness, and excess kurtosis of future size conditional on current size. Moments of the future size distribution are plotted with respect to initial size; their distribution is also conditional on density but initial size is by far the stronger predictor of future size, so we chose this visualization. Values for the JSU model (and the corresponding “real data” values) are offset vertically by one unit for comparison. Figure made by script `creosote_growth_modeling.R`.

1158 kernels had very similar monotonic decreases in λ with increasing local density, and
 1159 nearly identical wave velocities (Fig. S-6). This species has very low mortality risk once

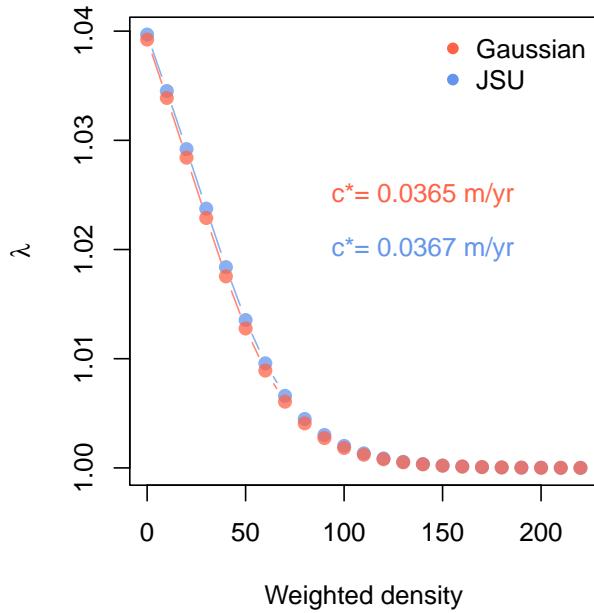


Figure S-6: Density dependence in fitness (λ) and asymptotic velocity of the creosote encroachment wave (c^*) for Gaussian and JSU growth kernels. Weighted density is the sum of sizes ($\log(cm^3)$) of all conspecifics within a five-meter transect “window”. Figure made by script `creosote_growth_modeling_qgam.R`.

established (mean remaining life expectancy of a median-sized shrub is 24,408 years) and its population growth and wave expansion are limited by very low seedling recruitment ((Drees et al., 2023)). Weak size-dependence in survival likely explains why the improvement in growth modeling had little influence on SIPM predictions.

1164 S.4 Additional results

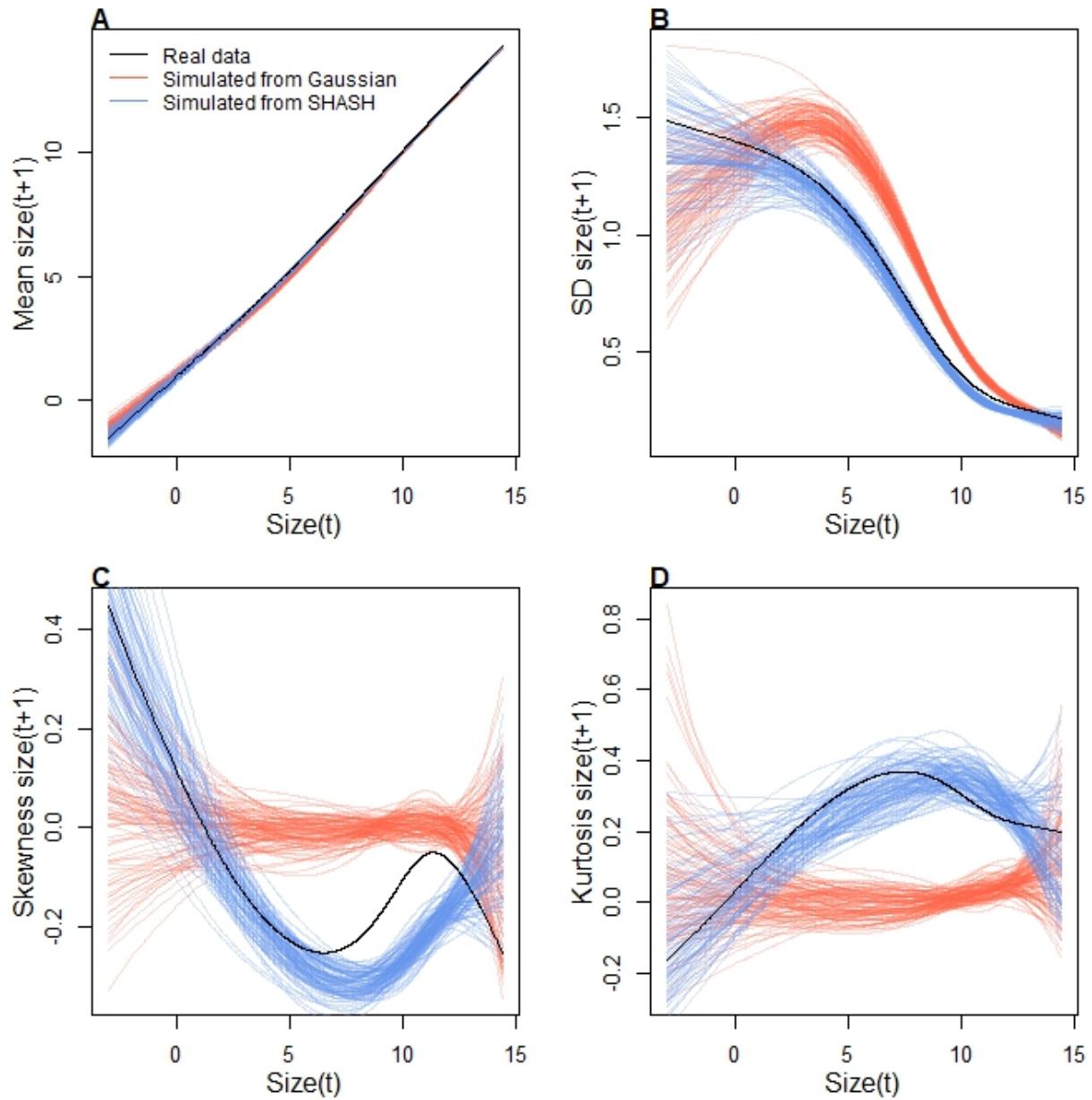


Figure S-7: Comparisons among real cactus data and data simulated from Gaussian and SHASH growth models for mean, standard deviation, NP skewness, and NP excess kurtosis of future size conditional on current size. Figure made by script `cactus_growth_modeling_qgam.R`.

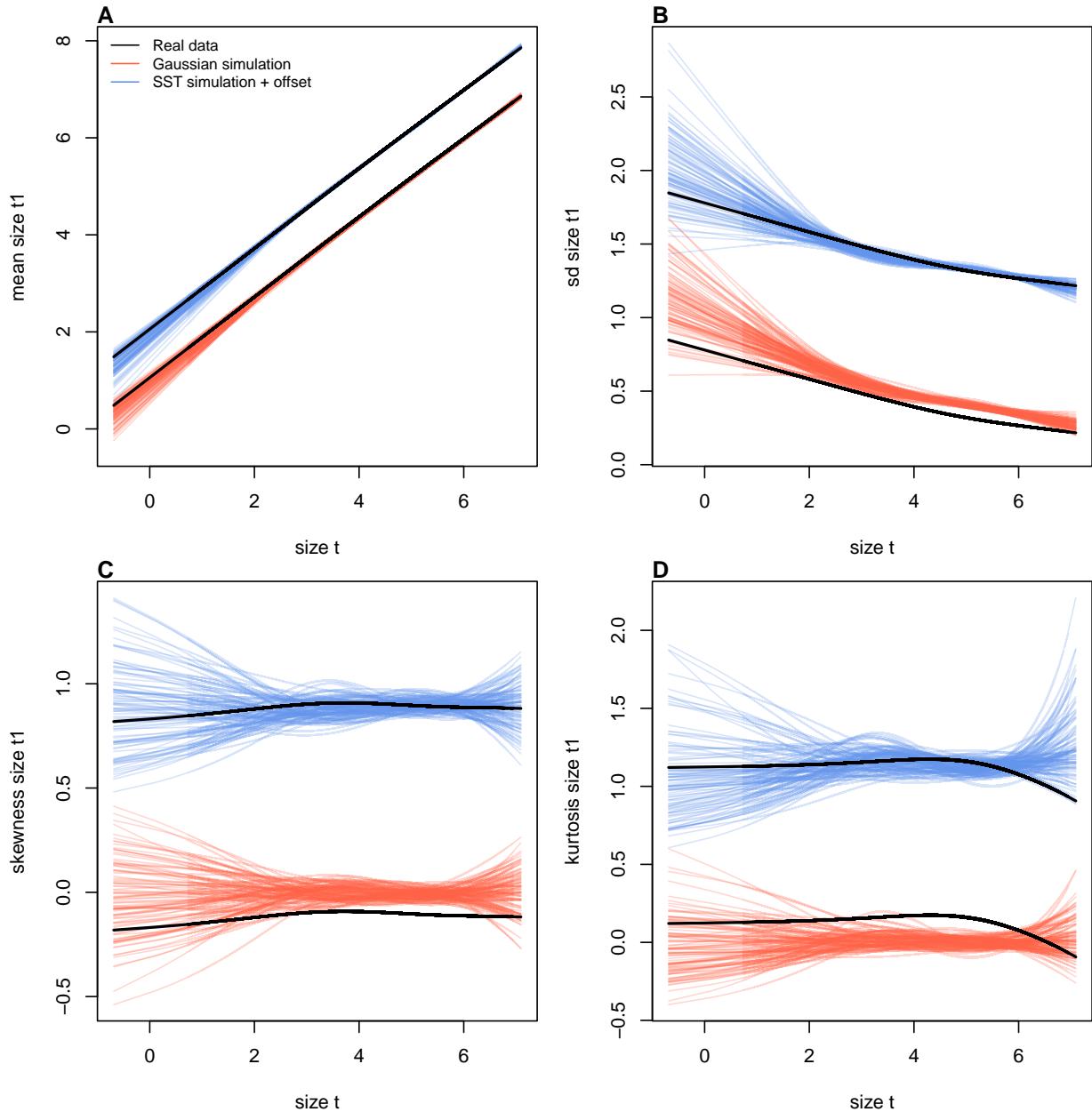


Figure S-8: Comparisons between real orchid data and data simulated from Gaussian and skewed t growth models for mean, standard deviation, NP skewness, and NP kurtosis of future size conditional on current size. Top row (A-D) shows plants that were vegetative at the start of the transition year and bottom row (E-H) shows plants that were flowering at the start of the transition year. Figure made by script `orchid_growth_modeling_rq.R`.