

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

*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-structure 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 (ecologically
¹⁰ important covariates and random effects) that Gaussian approaches typically pro-
¹¹ vide. Our approach emphasizes visual diagnostics of residuals from pilot Gaussian
¹² models and quantile-based metrics of skewness and kurtosis that vet the fit of the
¹³ Gaussian distribution and guide the selection of an alternative, if necessary. We
¹⁴ illustrate our methods by reanalyzing size transition data from our published IPM
¹⁵ studies, targeting a diversity of demographic quantities including population growth
¹⁶ rate, invasion wave velocity, and evolutionarily stable life history strategies.
- ¹⁷ 3. Across one coral and three plant case studies, skewness and excess kurtosis were
¹⁸ common features of size transition data and non-Gaussian growth models consis-
¹⁹ tently generated simulated data that were more consistent with the real data than
²⁰ pilot Gaussian models. However, in these case studies, the effects of “improved”
²¹ growth modeling on IPM results were generally modest, and differed in direction or
²² magnitude between different 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 Introduction

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

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

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

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

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

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

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

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

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

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

109 A workflow for growth modeling

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

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

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

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

130 stabilization. In particular log-transformation is often appropriate for size data (Ell-
131 ner et al., 2016), and it helps to avoid eviction at small sizes.

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

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

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

- 163 3. *Refit the growth model using the chosen distribution.*

164 In models with multiple covariates and/or random effects, each potentially affecting

several distribution parameters (location, scale, skew, kurtosis) in different ways, “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 or expected future 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 should be explored, or more complex functions relating distribution parameters to current size and other covariates. 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-

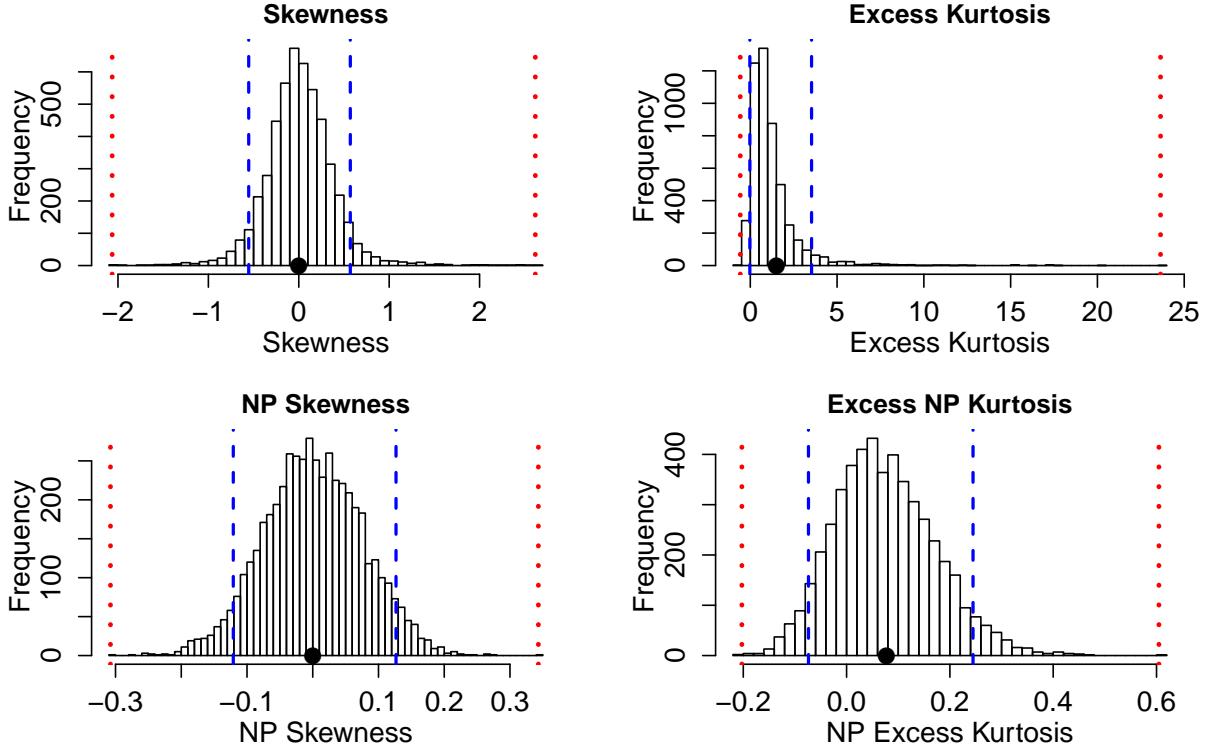


Figure 2: Histograms of skewness and kurtosis estimates using moment-based definitions (top two panels), compared with the nonparametric measures based on quantiles (bottom two panels). Note the very large differences in scale. Histograms are based on 5000 replicate draws of a sample of 200 independent 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
 229 its value for Gaussian distribution, and subtract 1 so that the value for a Gaussian is
 230 zero. We call this “NP Excess Kurtosis”. The value for a Gaussian distribution is zero. A
 231 value of ± 0.2 means that the tails are on average 20% heavier (or lighter) than those of
 232 a Gaussian with the same interquartile range. We calculate NP Kurtosis using $\alpha = 0.05$
 233 unless otherwise stated, to focus on the tail edges, 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
 242 as continuous functions of initial size or expected future size. In the examples below,
 243 we sometimes use the **qgam** package 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 $k = 4$ or $k = 6$ basis functions. For the gam-averse, other quantile
 248 regression models may be equally suitable, and we illustrate those, too. For consistency

249 with nonparametric skewness and kurtosis, in comparisons of real and simulated data
 250 below, we use quantile-based measures of location and scale, and use quantile regression
 251 to visualize these as functions of size. Specifically, following Wan et al. (2014),

$$252 \quad \text{NP Mean} = \frac{q_{0.25} + q_{0.5} + q_{0.75}}{3} \quad (5)$$

253 and

$$254 \quad \text{NP SD} = \frac{q_{0.75} - q_{0.25}}{1.35}. \quad (6)$$

255 1 Case study: lichen, *Vulpicida pinastri*

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

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

```
271 # XH is the data frame
272 # t0 and t1 are initial and final thallus area, respectively
273 fitGAU <- gam(list(t1 ~ t0 + I(t0^2), ~ t0 + I(t0^2)), data=XH, gamma=1.4, family=gaulss)
274 XH$fitted_mean = predict(fitGAU, type="response")[,1]
275 XH$fitted_sd <- 1/predict(fitGAU, type="response")[,2]
276 XH$scaledResids=residuals(fitGAU, type="response")/XH$fitted_sd
```

277 Quantile regression on the scaled residuals generates the diagnostics shown in Fig. 3
 278 (see script `Vulpicida_IPMS.R`). As expected based on previous analyses, visual analysis
 279 of the standardized residuals indicated negative skew, especially at larger sizes (Fig. 3B).
 280 We also find positive excess kurtosis for all sizes.

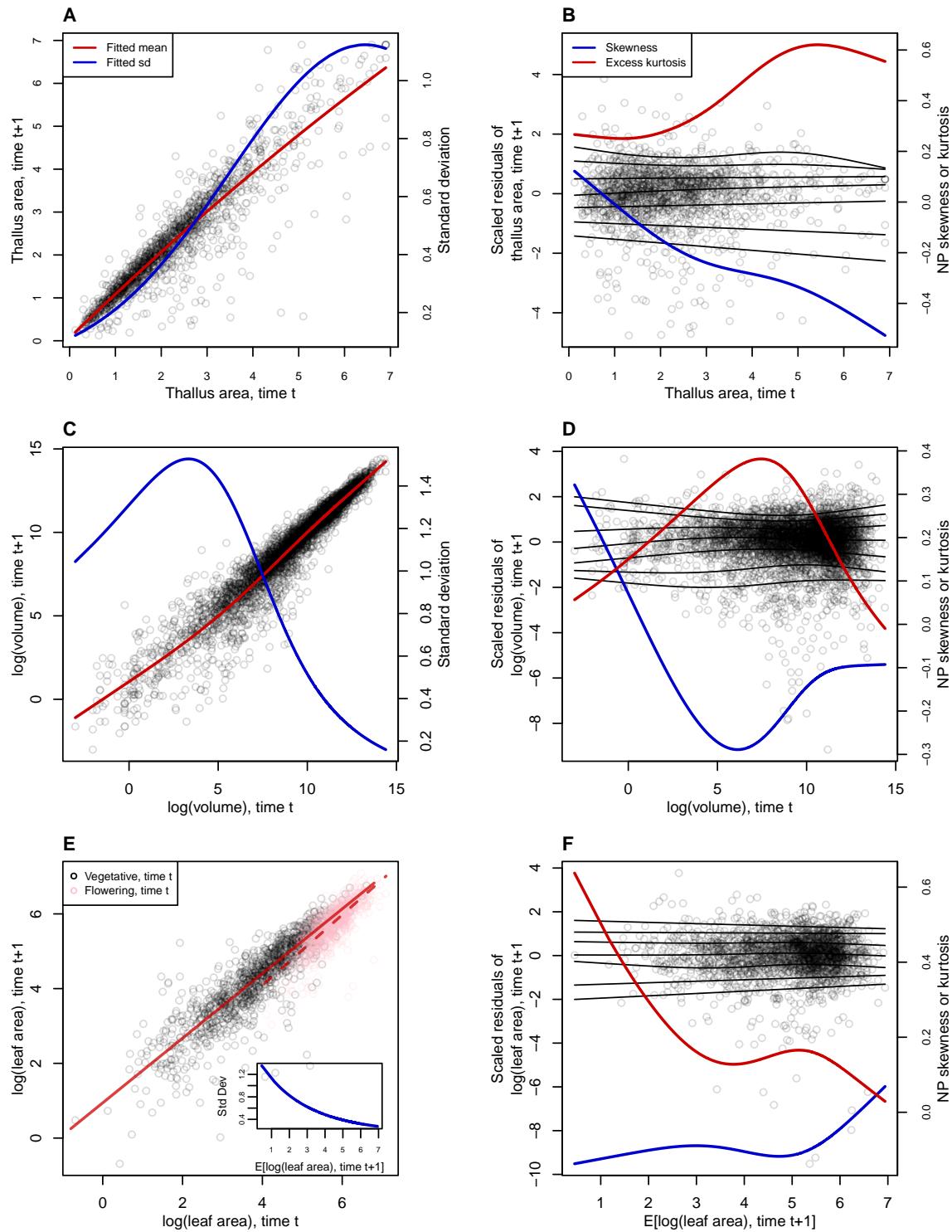


Figure 3: Best Gaussian models and diagnostics of standardized residuals for lichen (*Vulpicida pinastri*) **A,B**, cactus *Cylindriopuntia imbricata* **C,D**, and orchid *Orchis purpurea* **E,F** case studies. **A,C**, fitted mean and standard deviation of size at time $t+1$ conditional on initial size at time t . **E**, fitted means for plants that were vegetative or flowering 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 and excess kurtosis 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`.

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

```

296 ## define function that returns the JSU negative log-likelihood
297 LogLikJSU=function(pars){
298   dJSU(size_t1,
299     mu=pars[1]+pars[2]*size_t+pars[3]*size_t^2,
300     sigma=exp(pars[4]+pars[5]*size_t+pars[6]*size_t^2),
301     nu = pars[7]+pars[8]*size_t,
302     tau = exp(pars[9]), log=TRUE)
303 }
304 ## starting parameter values
305 p0<-c(coef(fitGAU22)[1:6],0,0,0)
306 ## fit with maxlik
307 outJSU=maxLik(logLik=LogLikJSU,start=p0*exp(0.2*rnorm(length(p0))),
308 method="BHHH",control=list(iterlim=5000,printLevel=2),finalHessian=FALSE);
  
```

309 Data simulation from the fitted JSU model indicates a compelling improvement over the
 310 best Gaussian model, particularly in skewness and kurtosis (Fig. 4).

311 To understand the practical consequences of improved growth modeling, we as-
 312 sembled the remainder of the lichen IPM following Shriver et al. 2012. The asymptotic
 313 population growth rate based on Gaussian growth ($\lambda_{GAU} = 1.001$) differs from the JSU
 314 growth model ($\lambda_{JSU} = 0.991$) by about 1% annual population growth, in line with re-
 315 sults of Peterson et al. 2019. However, even this modest difference can lead to strongly

³¹⁶ biased estimates of extinction risk from the Gaussian model, particularly over longer
³¹⁷ time horizons (Fig. 5). We also explored differences in other life history metrics (Table).²
³¹⁸ For example, the JSU growth model predicts values for mean lifespan, mean lifetime
³¹⁹ reproductive success, and mean age at reproduction that are 19%, 25%, and 14% lower
³²⁰ than the Gaussian growth model. In this case study, properly modeling non-normal size
³²¹ transitions – which was easy to do with a few extra lines of code – can have important
³²² effects on ecological inferences.

³²³ One could argue that the lichen data set was a convenient “straw man” to disqualify
³²⁴ Gaussian growth, since it was recognized by the original and subsequent IPM analysts
³²⁵ that this species requires a skewed distribution of size transitions (Peterson et al., 2019;
³²⁶ Shriner et al., 2012). In all remaining case studies, including those in the Appendix,
³²⁷ we re-examine growth data that were modeled as Gaussian by the data originators in
³²⁸ published IPM studies.

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

³³⁰ The next case study, focusing on the tree cholla cactus *Cylindriopuntia imbricata* at the
³³¹ Sevilleta Long-Term Ecological Research site in central New Mexico, adds a new feature
³³² on top of the simple size-dependent regressions in the previous study: random effects
³³³ associated with temporal (year) and spatial (plot) environmental heterogeneity. This
³³⁴ long-term study of cactus demography was initiated in 2004 and different subsets of
³³⁵ the data have been analyzed in various IPM studies, all using Gaussian growth kernels
³³⁶ (Compagnoni et al., 2016; Czachura and Miller, 2020; Elder and Miller, 2016; Miller
³³⁷ et al., 2009; Ohm and Miller, 2014). In fact, (Elder and Miller, 2016) presented a Gaus-
³³⁸ sian growth model fit to the cactus data as an example of a well fit growth function,
³³⁹ based on a marginal distribution of residuals that appeared approximately Gaussian
³⁴⁰ and posterior predictive checks (PPCs) of a Bayesian model that suggested consistency
³⁴¹ between the real data and data simulated from the fitted model (Fig. 4 in (Elder and
³⁴² Miller, 2016)). While PPCs and the associated “Bayesian P-value” are popular diagnostic
³⁴³ tools, they are often considered to be too conservative (Conn et al., 2018; Zhang, 2014),
³⁴⁴ failing to reject marginally bad models even though they are very effective in rejecting
³⁴⁵ models that are terrible. The choice of discrepancy function (the statistic used to com-
³⁴⁶ pare real and simulated data) can also be limiting: in our previous work, we used a
³⁴⁷ discrepancy function focused on variance (the sum of the squared residuals), so we had

²What is the best way to cite Chrissy Hernandez' life history functions?

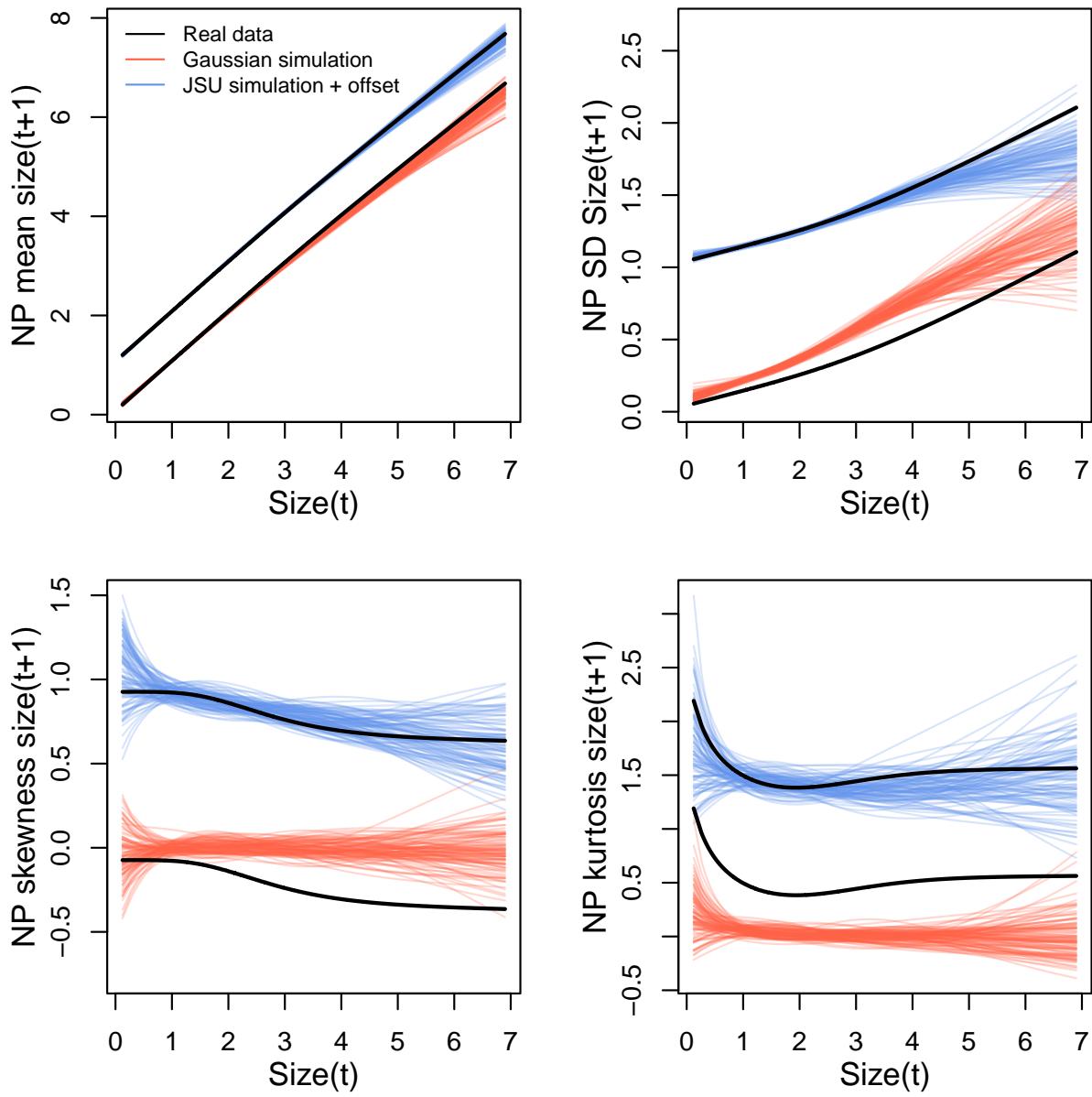


Figure 4: Comparisons among real lichen data and data simulated from Gaussian and JSU growth models for NP mean, NP standard deviation, NP skewness, and NP excess kurtosis of future size conditional on current size. Figure made by script `Vuplicida_IPMs.R`.

348 a built-in blind-spot for mismatches in higher moments. In the clarity of hindsight, the
 349 PPC gave a false sense of security; the Gaussian was a poor choice all along.

350 The data for this new analysis include 4844 size transition observations from 929 in-
 351 dividuals spanning 13 transition years (2004–2018) and 11 spatial replicates (three spatial
 352 blocks in years 2004–2008 and eight 30m-by-30m plots in years 2009–2018). The data are

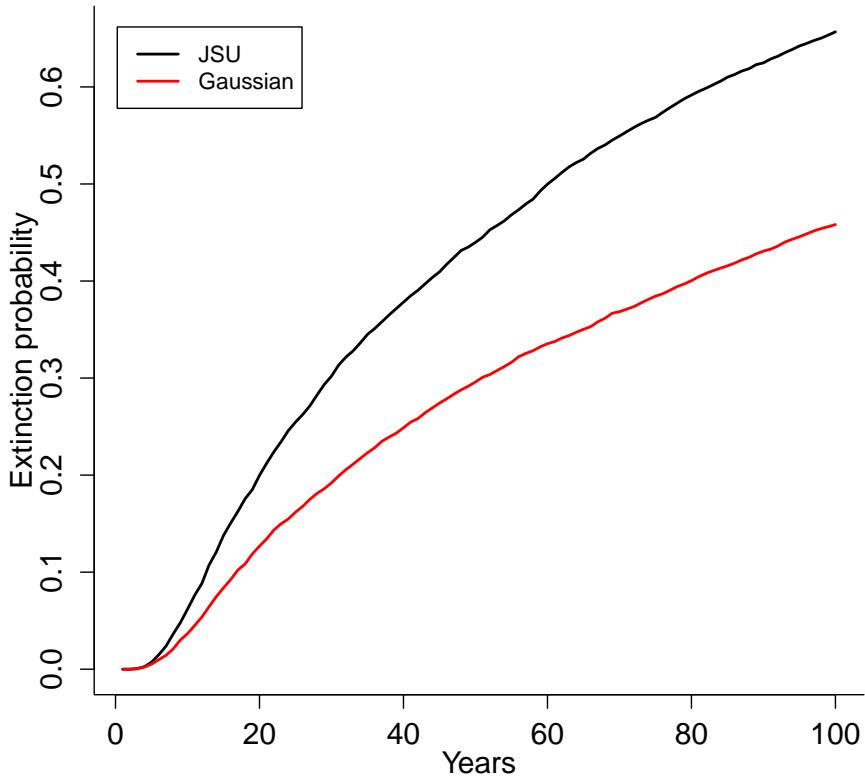


Figure 5: Extinction risk estimated from individual-based simulation of IPMs based on Gaussian and Johnson's S-U (JSU) growth distributions. Figure made by script Vuplicida_IPMs.R.

353 provided in Miller (2020). Following previous studies, we quantified size as the natural
 354 logarithm of plant volume (cm^3), derived from height and width measurements.

355 We begin the growth modeling workflow, as above, with a generalized additive
 356 model with the mean and standard deviation of size in year $t + 1$ modeled as function
 357 of size in year t , with random intercepts for year and plot and assuming normally dis-
 358 tributed residuals (`family=gaulss()`). The standardized residuals, accounting for size-
 359 dependent residual variance (Fig. 3C), show clear signals of negative skew and positive
 360 excess kurtosis across most of the size distribution but strongest in the middle of the size
 361 distribution (Fig. 3D).

362 To better capture size transitions, we need a distribution with negative skew and
 363 positive excess kurtosis, but both of which may be negligible at some sizes. We first tried
 364 Johnson's S_U and then the skewed t distributions, both of which are limited to positive

365 excess kurtosis. Both distributions provided some improvement over the Gaussian, but
 366 were not happy with the fit of either. Iterating through the workflow (Fig. 1), we arrived
 367 at the SHASH distribution, which is more flexible than either the JSU or skewed t , capa-
 368 ble of capturing a greater range of kurtosis for a given amount of skew, and vice versa
 369 (Jones and Pewsey (2009); Appendix S.1). Furthermore, fitting the SHASH as a general-
 370 ized additive model with **mgcv** allowed for flexible, non-monotonic size-dependence in
 371 skewness and kurtosis without the need for model selection on specific size-dependent
 372 functions; through iterations of trial and error, we found this flexibility was necessary
 373 to generate simulated data that compared favorably to the real data. The other distri-
 374 butions that we tried are not available as **mgcv** families, so we fit these with custom
 375 maximum likelihood functions, an approach we illustrate in the next case study. The
 376 SHASH gam for the cactus data included random intercepts for the location parameter,
 377 representing spatial and temporal heterogeneity:

```

378   fit_shash <- gam(list(logvol_t1 ~ s(logvol_t,k=4) +
379     s(plot,bs="re") + s(year_t,bs="re"), # location
380     ~ s(logvol_t,k=4), # log-scale
381     ~ s(logvol_t,k=4), # skewness
382     ~ s(logvol_t,k=4)), # log-kurtosis
383     data = CYIM_grow,
384     family = shash,
385     optimizer = "efs")
  
```

386 The SHASH model provided good correspondence between simulated and real
 387 data, and provided more compelling improvement over the Gaussian model than we
 388 saw in the coral case study (Fig. S-7). The SHASH model over-estimated negative skew
 389 at some sizes relative to the signal of skewness in the data (Fig. S-7C), but the nature of
 390 size-dependent skew in the data is not very biologically plausible and may instead be
 391 driven by the tail-wagging tendency of gams. As in the coral case study, we see that cor-
 392 rectly modeling skewness and kurtosis improved estimation of the mean and standard
 393 deviation (Fig. S-7A,B), yielding a growth model that is clearly truer to the data than the
 394 pilot Gaussian fit.

395 We explored how improved growth modeling influenced IPM results, leveraging
 396 the plot and year structure of the study design to quantify spatial and temporal variance
 397 in fitness. We used the fitted random effects from the vital rate models to estimate the
 398 asymptotic growth rate for each year (λ_t), centered on the average plot, and for each plot
 399 (λ_p), centered on the average year. This allowed us to quantify demographic variance

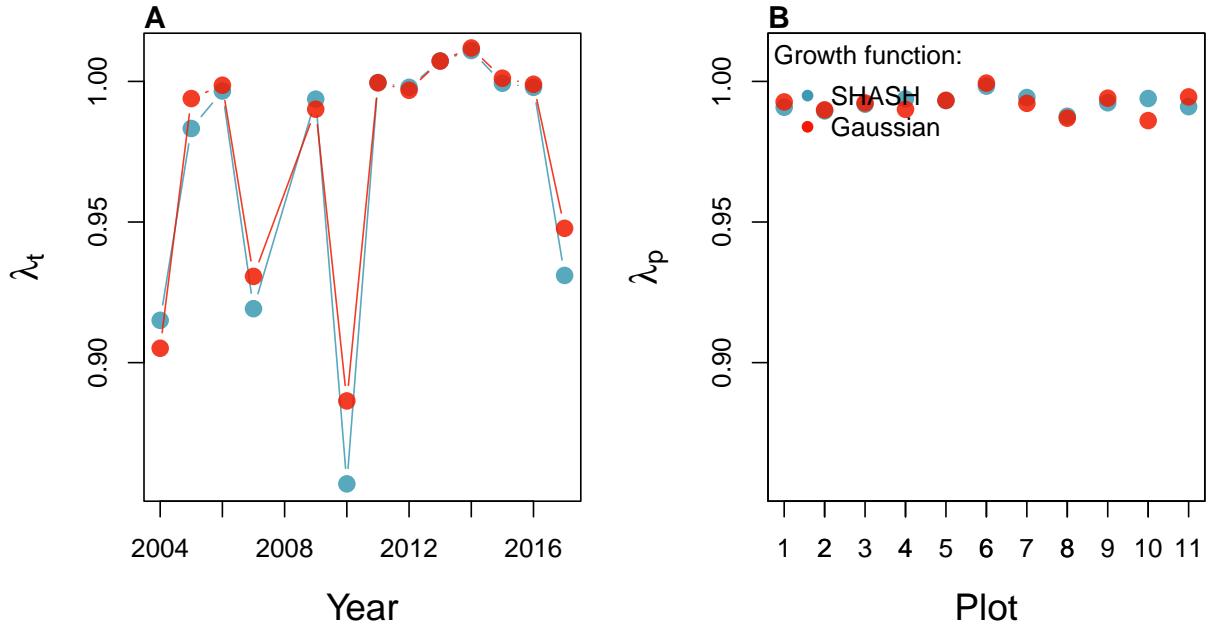


Figure 6: Temporal (A) and spatial (B) heterogeneity in fitness for the tree cholla cactus (*Cylindriopuntia imbricata*) predicted by IPMs using Gaussian or SHASH growth models. Figure made by script `cactus_growth_modeling_qgam.R`.

400 associated with temporal and spatial heterogeneity. We found that the Gaussian growth
 401 model tended to over-estimate λ_t , particularly in the harshest years (Fig. 6A), and thus
 402 under-estimated temporal variance in fitness ($SD(\lambda_{t(Gaussian)}) = 0.042$, $SD(\lambda_{t(SHASH)}) =$
 403 0.048). The opposite was true for plot-to-plot variation ($SD(\lambda_{p(Gaussian)}) = 0.0037$,
 404 $SD(\lambda_{p(SHASH)}) = 0.0028$), although spatial variation in fitness was much lower than
 405 temporal variation (Fig. 6B). The difference in temporal variance would suggest that
 406 Gaussian growth modeling would lead to over-estimation of the stochastic growth rate
 407 λ_S , since temporal variance has a negative influence on λ_S . However, this was not the
 408 case: stochastic IPMs based on Gaussian and SHASH growth models had nearly identi-
 409 cal stochastic growth rates ($\lambda_S(Gaussian) = 0.9907$, $\lambda_S(Gaussian) = 0.991$). This is
 410 likely because temporal fluctuations in vital rates, which is where the SHASH growth
 411 model would make a difference, have a weaker influence on λ_S than the temporal fluc-
 412 tuations in size structure that they generate (Compagnoni et al., 2016; Ellis and Crone,
 413 2013). Thus, depending on the target of one's analysis, modeling non-Gaussian size
 414 transitions with a Gaussian growth model could bias results in either direction, or make
 415 no difference at all.

416 **2.1 Case study: lady orchid, *Orchis purpurea***

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

429 The data, originated by Dr. Hans Jacquemyn and used here with permission, come
430 from 368 plants in a Belgian population that was censused annually from 2003 through
431 2011 (for this reanalysis we are using data only from the “light” habitat). Size was mea-
432 sured as leaf area (cm^2) summed over all leaves, and we analyzed the natural logarithm
433 of total leaf area as the size variable of the IPM.

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

449 For Gaussian models, weights w_i can be used to indicate that the observations y_i
450 vary in their dispersion around the mean. In general, the iterative steps are as follows,
451 and code to execute these steps may be found in SCRIPT.

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:

$$y_i = \mu_i + \epsilon_i$$

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

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

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

465 As improvements, we explored the skewed t and Johnson's SU distributions, both
466 leptokurtic distributions with flexible skewness. We were happier with the skewed t ,
467 which we fit with a custom likelihood function similar to the JSU growth model fit to
468 the lichen data. However, rather than re-fitting all the parameters of the skewed t model,

469 here we build a “hybrid” likelihood function that uses the fitted mean and standard
 470 deviation from the best Gaussian model, and estimates parameters that control skew-
 471 ness and kurtosis as linear functions of predicted future size. This is easy because the
 472 **gamlss.dist** package provides a parameterization of the skewed t in which the location
 473 parameter μ is the mean and scale parameter σ is the standard deviation (Rigby et al.,
 474 2019). The hybrid likelihood looks like this:

```

 475     ## log_area_t1 and log_area_t are the size obervations
 476     ## GAU_fitted and GAU_sd are mean and standard deviation from best Gaussian
 477     ## pars is a vector of free parameters to be estimated
 478 SSTLogLik=function(pars){
 479   dSST(log_area_t1,
 480     mu=GAU_fitted,
 481     sigma=GAU_sd,
 482     nu = exp(pars[1] + pars[2]*GAU_fitted),
 483     tau = exp(pars[3] + pars[4]*GAU_fitted)+2,
 484     log=TRUE)
 485   }

```

486 Based on diagnostics of the standardized residuals, parameters that control skew-
 487 ness and kurtosis are defined as linear functions of the mean, and those coefficients
 488 are estimated by maximum likelihood (note that the tau parameter uses a $\log(x - 2)$
 489 link function). This approach relies on the robustness of Gaussian models to deviations
 490 from normality, which implies that the fitted mean and variance from a Gaussian model
 491 are good approximations for the mean and variance of a corresponding non-Gaussian
 492 model. If one is skeptical of this approach, it is possible to simultaneously re-fit all pa-
 493 rameters of the skewed t . However, recall that unlike the lichen case study, our pilot
 494 Gaussian approach included random effects for year, and therefore the expected values
 495 getting passed into dSST account for this source of variation. Coding random effects
 496 “from scratch” into a custom likelihood model is possible (we provide guidance on one
 497 way to do this, using the “shrinkage” approach, in Appendix XX) but should gener-
 498 ally not be necessary. Instead, a key advantage of the hybrid approach is convenient
 499 retention of the fitted random effects and associated variance components, which get
 500 shuttled from the Gaussian model into the non-Gaussian model without any fuss (it was
 501 critical that we used a parameterization of the skewed t for which `mu` is the mean and
 502 `sigma` is the standard deviation). And, if this approach does not “work” (i.e., deviations
 503 from normality biased the fitted values of the Gaussian model) one would quickly find

504 out through the simulation step of the workflow. In this case, size transition data sim-
505 ulated from this model corresponded favorably to the real data, much better than the
506 pilot Gaussian model, including improvements in the standard deviation, skewness, and
507 kurtosis of future size (Fig. S-8).

508 Finally, we used the improved growth model to revisit key results of the original
509 study. Miller et al. (2012) used the orchid IPM to estimate the evolutionarily stable
510 strategy (ESS) as the mean size at flowering that maximizes lifetime reproductive success
511 (R_0), given the constraint that flowering when small reduces growth and thus elevates
512 mortality risk. Repeating that analysis here, we found that improved growth modeling
513 has virtually no influence on predictions for optimal life history strategies (Fig. 7). ESS
514 flowering sizes were nearly identical between IPMs with Gaussian vs skewed t growth
515 models, and both aligned well with the observed mean flowering size (dashed vertical
516 line in Fig. 7).

517 3 Discussion

518 Much of the appeal of integral projection models has stemmed from their embrace of
519 continuous size structure through reliance on regression-based approaches, and the po-
520 tentially complex fixed- and random-effect structures that these approaches allow. Using
521 familiar statistical tools and with relatively few parameters to estimate, IPM users can
522 incorporate important sources of variation in demography and interrogate their influ-
523 ence on ecological and evolutionary dynamics. With this opportunity comes the burden
524 of getting it right: IPMs are good models of the populations they are intended repre-
525 sent only insofar as the statistical models provide good fits to the underlying data. The
526 growth sub-model is the trickiest part of “getting it right” because it defines a distri-
527 bution of future size conditional on current size. Distributions have many properties –
528 “moments” – and a good growth model should recapitulate the properties of real size
529 transitions. The default assumption of normally distributed size transitions, employed
530 overwhelmingly across 20+ years of IPM studies, is an arbitrary historical precedent.
531 In our case studies (chosen simply because we had the data at our fingertips) and, we
532 suspect, more broadly, skewness and excess kurtosis were common features of size tran-
533 sition data: shrinking was more common than growing, and large changes in size were
534 more common than a Gaussian model would predict. Our most important message is
535 that the standard assumption of normally-distributed size transitions should be aban-
536 doned and a more inquisitive process of growth modeling should take its place.

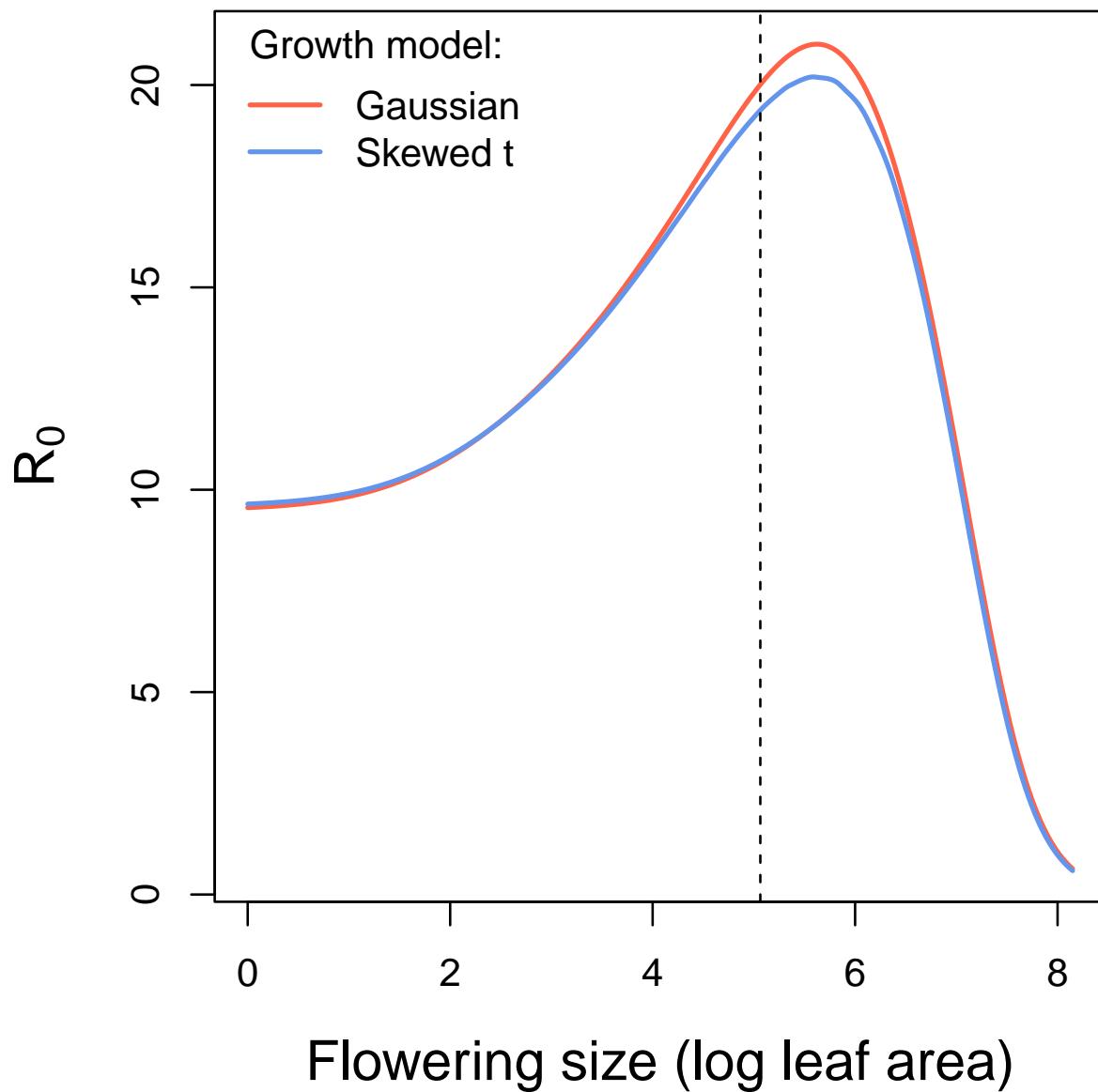


Figure 7: Orchid life history results from IPMs using Gaussian or skewed t growth models. Lifetime reproductive success (R_0) is shown as a function of mean size of flowering. Dashed vertical line shows the observed mean flowering size.

537 We have attempted to lay out a general workflow for what that process should look
 538 like, guided by visual diagnostics of standardized residuals. One implication of relying
 539 on visual diagnostics is that goodness of fit is in the eye of the beholder. This approach
 540 can empower IPM users to make informed choices, but it is not very prescriptive: we
 541 have not suggested any hard rules for when one or another distribution should be used,
 542 only that a good growth model should generate data that look like the real thing. Al-

543 ternatively, model selection could be used to identify best-fitting growth distributions
544 and best-fitting functions for higher moments. However, model selection among growth
545 distributions with 3-5 parameters, each of which may be functions of state variables or
546 fitted values, can quickly explode in complexity, and we are not convinced it is worth
547 the trouble. It should be possible to find a good growth model without worrying about
548 which one is “best”.

549 Our work follows the important contribution of Peterson et al. 2019, who were
550 motivated by a similar problem (inadequacy of the Gaussian distribution for skewed
551 size transitions) but arrived at different recommendations for dealing with the problem.
552 These authors developed a creative approach in which size data are transformed onto a
553 [0, 1] scale and then size transitions are modeled using beta regression. The beta distri-
554 bution can accommodate positive, negative, or zero skew; as Peterson et al. demonstrate,
555 this approach is a viable option for skewed growth data. However, beta regression also
556 has some important downsides: common beta regression packages do not fit random
557 effects(e.g., **betareg** (Cribari-Neto and Zeileis, 2010)) or do not do so reliably (in our ex-
558 perience **gamlss** is numerically unstable); the two-parameter beta distribution has rigid
559 mean-variance and skewness-kurtosis relationships that may not describe the data well;
560 and³. Rather than shoe-horn size transition data into a default beta distribution, we find
561 it more natural and appropriate to leverage the vast arsenal of real-valued probability
562 distributions – all of them at one’s fingertips with a few lines of code – and let the data
563 suggest appropriate alternatives based on the residuals of a pilot Gaussian model. The
564 workflow that we have envisioned requires no sacrifice of complexity in random effects
565 or covariates for the sake of modeling skewness and kurtosis.

566 While the arsenal of candidate distributions is indeed vast, in our analyses for this
567 paper, we found ourselves coming back time and again to a few usual suspects. The
568 four-parameter SHASH distribution, for example, is able to flexibly accommodate inde-
569 pendent, size-dependent variation in variance, skewness, and kurtosis, and it is available
570 as a distribution family in the well-developed **mgcv** package. In our case study analy-
571 ses it was consistently among the top non-Gaussian candidates and was our model of
572 choice for the coral, cactus, and pike data sets. While we have emphasized the im-
573 portance of moving away from a single default distribution of size transitions, if one
574 were to want or need a default distribution (e.g., in software packages for IPM con-
575 struction [cite IPMpack]) one could do worse than the SHASH. In cases where size
576 transitions are leptokurtic but consistently so across the size distribution, the Johnson’s
577 S-U (used for creosotebush) and skewed *t* (used for lady orchid) distributions were easy

³ Steve, I recall you have a beef with transformation to [0,1], so that would go here.

578 to fit with custom likelihood functions. All of the distributions we have used (and
579 the specific parameterizations we have implemented) share the property that their lo-
580 cation and scale parameters correspond to the mean and standard deviation, which
581 is not essential but it facilitates interpretation and an intuitive connection to the pilot
582 Gaussian model. The five-parameter skewed generalized t (sgt) generalizes many other
583 real-valued distributions (cite) and is therefore another flexible option, but it does not
584 share the location=mean and scale=sd property, and in our experience can be hard to fit.
585 Finally, finding an appropriate non-Gaussian alternative does not solve all the problems
586 of growth modeling. “Eviction” from the approximating matrix of the IPM kernel is an
587 ever-present danger and requires vigilance to detect and correct (Williams et al., 2012).

588 In all of our case studies, non-Gaussian growth models always yielded more sat-
589 isfying fits to size transition data than the Gaussian models published in those papers.
590 However, much to our relief, none of these re-analyses yielded a “gotcha” result that
591 overturned results of the original study. In this small sampling of case studies, improved
592 growth modeling had weak to modest effects on IPM results, similar in magnitude to
593 the results of Peterson et al. (2019). We caution against taking too much comfort in this
594 outcome; we can imagine other scenarios in which the choice of the growth distribution
595 could be more consequential. It is worth noting that most of our case studies focused on
596 perennial life histories (perennial plants, corals, lichens), characterized by relatively slow
597 growth, heavy losses during recruitment, and high survival once established. Life cycles
598 such as these may be relatively robust to subtle features of the growth kernel. More
599 systematic comparative analyses may provide insight into which types of species and
600 life histories are more likely to exhibit strong skewness and kurtosis, and the conditions
601 under which demographic analysis is more or less sensitive to these features of size tran-
602 sition. It is also worth noting, as we saw in several case studies, that different outputs
603 from the same model can be more or less sensitive to the choice of growth distribution.

604 Across our case studies we have attempted to illustrate a diversity of software pack-
605 ages and computational approaches to model fitting, to reflect the diversity of prefer-
606 ences and habits that the community of IPM users bring to their own analyses. We like
607 generalized additive models (gams) for their flexibility and for **mgcv**’s numerous op-
608 tions for distribution families and overall speed and reliability. However, there are some
609 applications for which classical parametric regression would be preferable because the
610 coefficients carry biological meaning. For example, regression coefficients may be tar-
611 gets of natural selection (Rees and Ellner, 2016) and may combine to influence traits of
612 interest such as the expected size at flowering (e.g. in Fig. 7A), a function of the intercept
613 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 (Elderd 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.

3.1 Conclusion

Acknowledgements

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

4 Authorship statement

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

5 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

⁶⁴² the end of the article. During peer review, our data and code are available at https://github.com/texmiller/IPM_size_transitions.
⁶⁴³

644 Literature Cited

- 645 Anscombe, F. J. and Glynn, W. J. (1983). Distribution of the kurtosis statistic b_2 for
646 normal samples. *Biometrika*, 70(1):227–234.
- 647 Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2007). The lme4 package. *R package
648 version*, 2(1):74.
- 649 Bruno, J. F., Ellner, S. P., Vu, I., Kim, K., and Harvell, C. D. (2011). Impacts of aspergillosis
650 on sea fan coral demography: modeling a moving target. *Ecological Monographs*,
651 81(1):123–139.
- 652 Compagnoni, A., Bibian, A. J., Ochocki, B. M., Rogers, H. S., Schultz, E. L., Sneck, M. E.,
653 Elderd, B. D., Iler, A. M., Inouye, D. W., Jacquemyn, H., et al. (2016). The effect of
654 demographic correlations on the stochastic population dynamics of perennial plants.
655 *Ecological Monographs*, 86(4):480–494.
- 656 Conn, P. B., Johnson, D. S., Williams, P. J., Melin, S. R., and Hooten, M. B. (2018). A guide
657 to bayesian model checking for ecologists. *Ecological Monographs*, 88(4):526–542.
- 658 Cooch, E. G. and White, G. C. (2020, accessed 5/17/2020). *Program MARK - a 'gentle
659 introduction'*. Available at phidot.org.
- 660 Coulson, T. (2012). Integral projections models, their construction and use in posing
661 hypotheses in ecology. *Oikos*, 121(9):1337–1350.
- 662 Cribari-Neto, F. and Zeileis, A. (2010). Beta regression in r. *Journal of statistical software*,
663 34:1–24.
- 664 Crone, E. E. (2016). Contrasting effects of spatial heterogeneity and environmental
665 stochasticity on population dynamics of a perennial wildflower. *Journal of Ecology*,
666 104(2):281–291.
- 667 Czachura, K. and Miller, T. E. (2020). Demographic back-casting reveals that subtle
668 dimensions of climate change have strong effects on population viability. *Journal of
669 Ecology*.
- 670 D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g_1 .
671 *Biometrika*, pages 679–681.
- 672 Davis, C. (2015). *sgt: Skewed Generalized T Distribution Tree*. R package version 2.0.

- 673 Drees, T., Ochocki, B. M., Collins, S. L., and Miller, T. E. (2023). Demography and
674 dispersal at a grass-shrub ecotone: a spatial integral projection model for woody plant
675 encroachment. *Ecological Monographs*, page e1574.
- 676 Easterling, M. R., Ellner, S. P., and Dixon, P. M. (2000). Size-specific sensitivity: applying
677 a new structured population model. *Ecology*, 81(3):694–708.
- 678 Elderd, B. D. and Miller, T. E. (2016). Quantifying demographic uncertainty: Bayesian
679 methods for integral projection models. *Ecological Monographs*, 86(1):125–144.
- 680 Ellis, M. M. and Crone, E. E. (2013). The role of transient dynamics in stochastic popula-
681 tion growth for nine perennial plants. *Ecology*, 94(8):1681–1686.
- 682 Ellner, S. P., Adler, P. B., Childs, D. Z., Hooker, G., Miller, T. E., and Rees, M. (2022).
683 A critical comparison of integral projection and matrix projection models for demo-
684 graphic analysis: Comment. *Ecology*, 103(10):e3605.
- 685 Ellner, S. P., Childs, D. Z., and Rees, M. (2016). *Data-driven Modeling of Structured Popula-*
686 *tions: A Practical Guide to the Integral Projection Model*. Springer, New York.
- 687 Gould, W. R. and Nichols, J. D. (1998). Estimation of temporal variability of survival in
688 animal populations. *Ecology*, 79:2531 – 2538.
- 689 Gu, C. (2013). *Smoothing Spline ANOVA Models*. Springer Science+Business Media, New
690 York, 2 edition.
- 691 Hadfield, J. D. et al. (2010). Mcmc methods for multi-response generalized linear mixed
692 models: the mcmcglmm r package. *Journal of Statistical Software*, 33(2):1–22.
- 693 Hérault, B., Bachelot, B., Poorter, L., Rossi, V., Bongers, F., Chave, J., Paine, C. T., Wagner,
694 F., and Baraloto, C. (2011). Functional traits shape ontogenetic growth trajectories of
695 rain forest tree species. *Journal of ecology*, 99(6):1431–1440.
- 696 Jones, M. and Pewsey, A. (2009). Sinh-arcsinh distributions. *Biometrika*, 96:761 – 780.
- 697 Jones, M. C., Rosco, J. F., and Pewsey, A. (2011). Skewness-invariant measures of kurtosis.
698 *The American Statistician*, 65(2):89 – 95.
- 699 Jones, O. R., Barks, P., Stott, I., James, T. D., Levin, S., Petry, W. K., Capdevila, P., Che-
700 Castaldo, J., Jackson, J., Römer, G., et al. (2022). Rcompadre and rage—two r pack-
701 ages to facilitate the use of the comadre and comadre databases and calculation of

- 702 life-history traits from matrix population models. *Methods in Ecology and Evolution*,
703 13(4):770–781.
- 704 Komsta, L. and Novomestky, F. (2015). Moments, cumulants, skewness, kurtosis and
705 related tests. *R package version*, 14(1).
- 706 Link, W. A. and Nichols, J. D. (1994). On the importance of sampling variance to inves-
707 tigations of temporal variation in animal population size. *Oikos*, 69(3):539 – 544.
- 708 Louthan, A. M., Keighron, M., Kiekebusch, E., Cayton, H., Terando, A., and Morris, W. F.
709 (2022). Climate change weakens the impact of disturbance interval on the growth rate
710 of natural populations of venus flytrap. *Ecological Monographs*, 92(4):e1528.
- 711 McGillivray, H. (1986). Skewness and asymmetry: measures and orderings. *Annals of*
712 *Statistics*, 14:994–1011.
- 713 Metcalf, C. J. E., Ellner, S. P., Childs, D. Z., Salguero-Gómez, R., Merow, C., McMahon,
714 S. M., Jongejans, E., and Rees, M. (2015). Statistical modelling of annual variation for
715 inference on stochastic population dynamics using Integral Projection Models. *Methods*
716 *in Ecology and Evolution*, 6:1007–1017.
- 717 Metcalf, C. J. E., Graham, A. L., Martinez-Bakker, M., and Childs, D. Z. (2016). Oppor-
718 tunities and challenges of i ntegral p rojection m odels for modelling host-parasite
719 dynamics. *Journal of Animal Ecology*, 85(2):343–355.
- 720 Metcalf, J. C., Rose, K. E., and Rees, M. (2003). Evolutionary demography of monocarpic
721 perennials. *Trends in Ecology & Evolution*, 18(9):471–480.
- 722 Miller, T. E. (2020). Long-term study of tree cholla demography in the los pinos
723 mountains, sevilleta national wildlife refuge. [https://doi.org/10.6073/pasta/
724 dd06df3f950afe4a4642306182237d13](https://doi.org/10.6073/pasta/dd06df3f950afe4a4642306182237d13).
- 725 Miller, T. E., Louda, S. M., Rose, K. A., and Eckberg, J. O. (2009). Impacts of insect
726 herbivory on cactus population dynamics: experimental demography across an envi-
727 ronmental gradient. *Ecological Monographs*, 79(1):155–172.
- 728 Miller, T. E., Williams, J. L., Jongejans, E., Brys, R., and Jacquemyn, H. (2012). Evolution-
729 ary demography of iteroparous plants: incorporating non-lethal costs of reproduction
730 into integral projection models. *Proceedings of the Royal Society B: Biological Sciences*,
731 279(1739):2831–2840.

- 732 Ochocki, B. M., Drees, T., and Miller, T. E. (2023). Density-dependent demography of
733 creosote bush (*larrea tridentata*) along grass-shrub ecotones. <https://doi.org/10.6073/pasta/ca53c16f16dcf9fb11f3ee99ea5445ac>.
- 735 Ohm, J. R. and Miller, T. E. (2014). Balancing anti-herbivore benefits and anti-pollinator
736 costs of defensive mutualists. *Ecology*, 95(10):2924–2935.
- 737 Ozgul, A., Childs, D. Z., Oli, M. K., Armitage, K. B., Blumstein, D. T., Olson, L. E., Tul-
738 japurkar, S., and Coulson, T. (2010). Coupled dynamics of body mass and population
739 growth in response to environmental change. *Nature*, 466(7305):482–485.
- 740 Peterson, M. L., Morris, W., Linares, C., and Doak, D. (2019). Improving structured
741 population models with more realistic representations of non-normal growth. *Methods
742 in Ecology and Evolution*, 10(9):1431–1444.
- 743 Plard, F., Schindler, S., Arlettaz, R., and Schaub, M. (2018). Sex-specific heterogeneity
744 in fixed morphological traits influences individual fitness in a monogamous bird
745 population. *The American Naturalist*, 191(1):106–119.
- 746 Rees, M., Childs, D. Z., and Ellner, S. P. (2014). Building integral projection models: a
747 user's guide. *Journal of Animal Ecology*, 83(3):528–545.
- 748 Rees, M. and Ellner, S. P. (2016). Evolving integral projection models: evolutionary
749 demography meets eco-evolutionary dynamics. *Methods in Ecology and Evolution*,
750 7(2):157–170.
- 751 Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z., and De Bastiani, F. (2019). *Distributions
752 for modeling location, scale, and shape: Using GAMLS in R*. CRC press.
- 753 Salguero-Gómez, R. and Casper, B. B. (2010). Keeping plant shrinkage in the demo-
754 graphic loop. *Journal of Ecology*, 98(2):312–323.
- 755 Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky,
756 C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., and Araya-Ajoy, Y. G. (2020).
757 Robustness of linear mixed-effects models to violations of distributional assumptions.
758 *Methods in ecology and evolution*, 11(9):1141–1152.
- 759 Schultz, E. L., Eckberg, J. O., Berg, S. S., Louda, S. M., and Miller, T. E. (2017). Native
760 insect herbivory overwhelms context dependence to limit complex invasion dynamics
761 of exotic weeds. *Ecology letters*, 20(11):1374–1384.

- 762 Shriver, R. K., Cutler, K., and Doak, D. F. (2012). Comparative demography of an epi-
763 phytic lichen: support for general life history patterns and solutions to common prob-
764 lems in demographic parameter estimation. *Oecologia*, 170:137–146.
- 765 Stasinopoulos, D. M., Rigby, R. A., et al. (2007). Generalized additive models for location
766 scale and shape (gamlss) in r. *Journal of Statistical Software*, 23(7):1–46.
- 767 Stubberud, M. W., Vindenes, Y., Vøllestad, L. A., Winfield, I. J., Stenseth, N. C., and Lan-
768 gangen, Ø. (2019). Effects of size-and sex-selective harvesting: An integral projection
769 model approach. *Ecology and Evolution*, 9(22):12556–12570.
- 770 Wan, X., Wang, W., Liu, J., and Tong, T. (2014). Estimating the sample mean and stan-
771 dard deviation from the sample size, median, range and/or interquartile range. *BMC*
772 *medical research methodology*, 14:1–13.
- 773 Wilber, M. Q., Langwig, K. E., Kilpatrick, A. M., McCallum, H. I., and Briggs, C. J. (2016).
774 Integral projection models for host–parasite systems with an application to amphibian
775 chytrid fungus. *Methods in Ecology and Evolution*, 7(10):1182–1194.
- 776 Williams, J. L., Miller, T. E., and Ellner, S. P. (2012). Avoiding unintentional eviction from
777 integral projection models. *Ecology*, 93(9):2008–2014.
- 778 Wood, S. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and
779 Hall/CRC, 2 edition.
- 780 Zhang, J. L. (2014). Comparative investigation of three bayesian p values. *Computational*
781 *Statistics & Data Analysis*, 79:277–291.

Appendices

782 S.1 The Jones-Pewsey distribution

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

789

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

790 and Z has a $\text{Normal}(0,1)$ distribution. Equivalently,

791

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

792 Parameters $\delta = 1, \epsilon = 0$ give the $\text{Normal}(0,1)$ distribution. Skewness has the sign of ϵ ,
793 and $\delta > 0$ controls tail weight, with heavier than Gaussian tails for $\delta < 1$ and lighter
794 than Gaussian tails for $\delta > 1$. A formula for the density $f_{\epsilon,\delta}$ is given by Jones and Pewsey
795 (2009, eqn. 2). The general four-parameter family with location parameter μ and scale
796 parameter σ is defined as the probability densities of $\mu + \sigma X_{\epsilon,\delta}$. We refer to this as the
797 JP distribution family.

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

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

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

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

$$820 \quad X_{\lambda, \tau} = \sinh(e^{-\tau} \operatorname{asinh}(Z) + \lambda). \quad (\text{S.3})$$

821 Thus, the two parameterizations are related by

$$822 \quad \delta = e^\tau, \epsilon = \delta\lambda = e^\tau\lambda. \quad (\text{S.4})$$

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

830 S.2 Estimating mixed-effects models using shrinkage

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

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

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

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

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

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

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

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

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

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

```

877 pars1 = pars[1:ncol(U)]; pars2=pars[-(1:ncol(U))];
878 mu = U%*%pars1;
879 sigma = exp(pars2[1]+pars2[2]*mu);
880 dSN1(new.size, mu=mu, sigma=sigma, nu=pars2[3], log=TRUE)
881 }

```

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

```

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

```

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

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

$$894 \quad \hat{a}_t = a_t + \varepsilon_t \quad (S.5)$$

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

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

$$905 \quad \mathbb{E}(\varepsilon_t | a_t) = 0. \quad (S.6)$$

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

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

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

$$913 \quad \mathbb{E}(S^2) = \sigma^2 + \frac{1}{T} \sum_{t=1}^T \mathbb{E}Var(\varepsilon_t) - \frac{1}{T(T-1)} \sum_{i=1}^{j-1} \sum_{j=1}^T \mathbb{E}Cov(\varepsilon_i, \varepsilon_j). \quad (\text{S.7})$$

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

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

$$921 \quad \hat{\sigma}^2 = S^2 - \frac{1}{T} \sum_{t=1}^T \hat{V}_{t,t} + \frac{1}{T(T-1)} \sum_{i=1}^{j-1} \sum_{j=1}^T \hat{V}_{i,j}. \quad (\text{S.8})$$

922 $\hat{\sigma}^2$ estimates the variance of the distribution from which the a_t are assumed to be drawn.

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

$$927 \quad \tilde{a}_t = \bar{a}_t + \sqrt{\frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \hat{V}_{t,t}}} (\hat{a}_t - \bar{a}_t). \quad (\text{S.9})$$

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

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

$$935 \quad \tilde{a}_t = \bar{a}_t + \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \hat{V}_{t,t}} (\hat{a}_t - \bar{a}_t). \quad (\text{S.10})$$

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

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

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

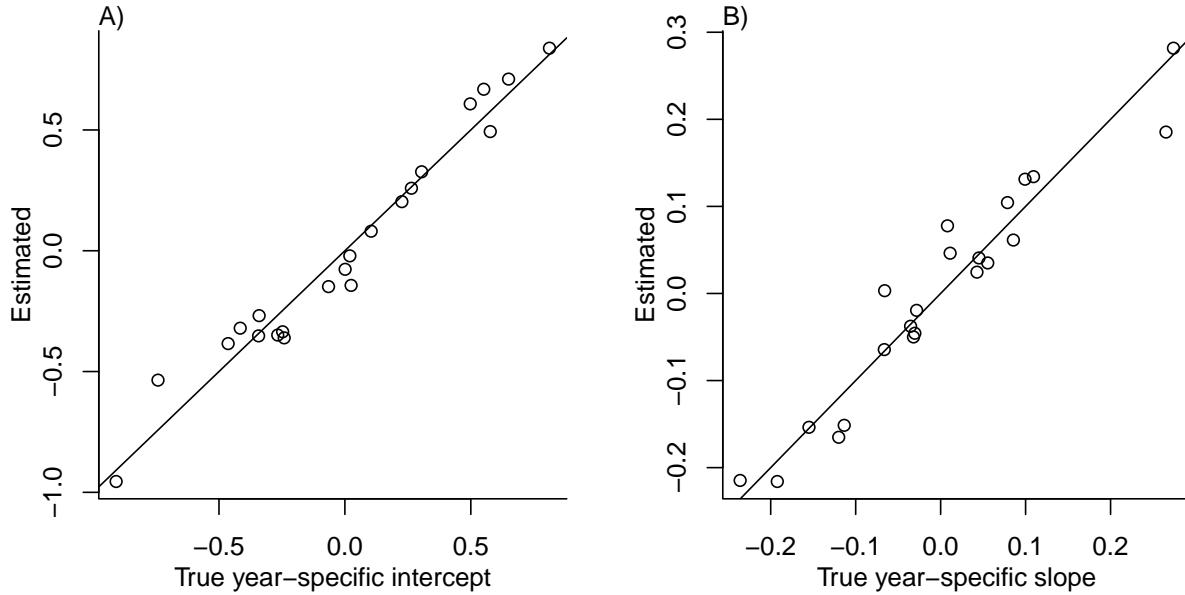
954 The take-home message is that estimating random effects from the regression coef-
 955 ficients is very simple:

```

956 # Variance-covariance matrices for intercepts and slopes
957 V1 = V[1:T,1:T]; V2 = V[(T+1):(2*T),(T+1):(2*T)];
958 # Extract year-specific intercepts, center them to zero
959 fixed.fx = coefs[1:T]; fixed.fx = fixed.fx-mean(fixed.fx);
960
961 # Estimate sigma^2
962 var.hat = mean(fixed.fx^2) - mean(diag(V1)) +
963           (sum(V1)-sum(diag(V1)))/(2*T*(T-1));
964
965 # Shrink deviations from the mean
966 shrinkRanIntercept = fixed.fx*sqrt(var.hat/(var.hat + diag(V1)));
967
968 # Do it all again for the slopes
969 fixed.fx2 = coefs[(T+1):(2*T)]; fixed.fx2 = fixed.fx2-mean(fixed.fx2);
970 var2.hat = mean(fixed.fx2^2) - mean(diag(V2)) +
971           (sum(V2)-sum(diag(V2)))/(2*T*(T-1));
  
```

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

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



979 S.3 Additional case studies

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

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

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

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

```
1000 # XH is a data frame holding the data
1001 # logarea.t0, .t1 denote initial and final values of log-transformed area
1002 fitGAU <- gam(list(logarea.t1~ s(logarea.t0), ~ s(logarea.t0)),
1003   data=XH, gamma=1.4, family=gaulss())
1004 fitted_all = predict(fitGAU,type="response");
1005 fitted_sd = 1/fitted_all[,2];
1006 scaledResids = residuals(fitGAU,type='response')/fitted_sd;
```

1007 Fig. S-1A shows the log-transformed data and Gaussian model. The mean function
1008 (solid red curve) is visually nearly linear, but the fitted spline is strongly favored over a
1009 linear model for the mean ($\Delta AIC \approx 9$). The spline for standard deviation σ versus initial
1010 size reflects the evident greater variability in growth at smaller sizes.

1011 There are no blatant signs of trouble in the pilot Gaussian model, but quantile re-
1012 gressions on the scaled residuals, and the NP Skewness and Kurtosis metrics derived
1013 from them (Eq. 3 and 4), suggest deviations from normality (Fig. S-1B). Specifically,
1014 skewness switches from negative to positive across the size range, with smaller corals
1015 more prone to extreme shrinkage and larger corals more prone to extreme growth. Kur-
1016 tosis also changes direction over the size distribution, with thinner tails than Gaussian
1017 at small sizes and fatter tails at large sizes. The fitted nonparametric moments suggest
1018 that the upper and lower tails of size transition probabilities may differ by up to 20%,
1019 and the weight of the tails may be 20% greater or less than Gaussian, depending on
1020 initial size – not overwhelming deficiencies, but not trivial either. Are these deviations
1021 from normality severe enough to warrant a second, non-Gaussian iteration of growth
1022 modeling? To answer that question, we simulated data from the fitted Gaussian model
1023 and examined whether key properties of the simulated data are consistent with those
1024 of the real data – this is the ultimate litmus test for a growth model's adequacy and
1025 should be a standard element of IPM construction, in our opinion. If the simulated data

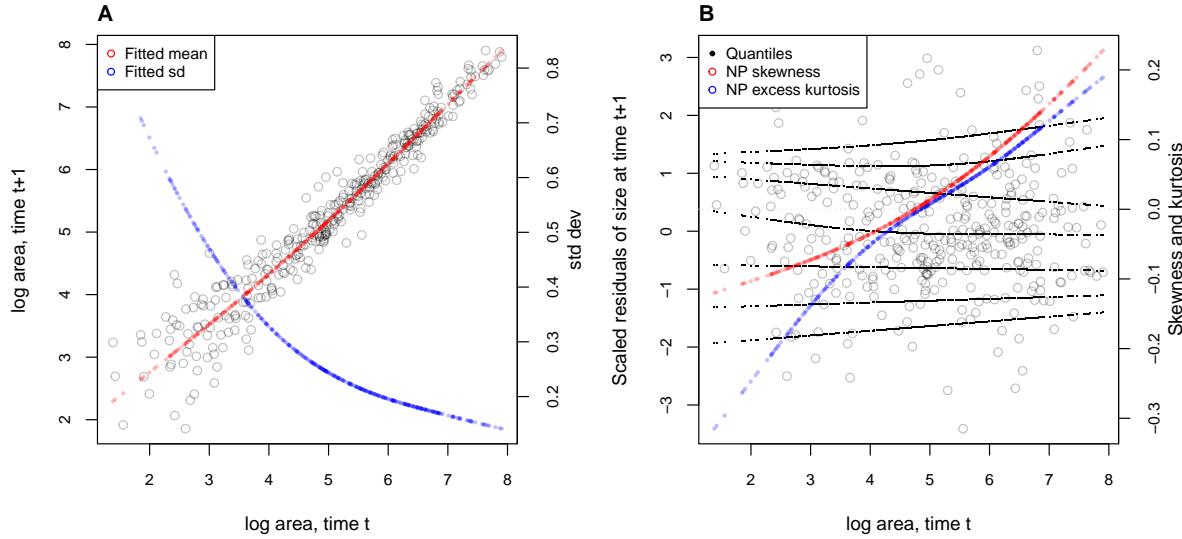


Figure S-1: **A**, Size transition data for sea fan corals, *Gorgonia ventalina*, and fitted gam with mean (red) and standard deviation (blue) of future size conditional on current size. **B**, Quantile regressions of scaled residuals on size and nonparametric estimates of skewness (red) and excess kurtosis (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.

1026 are not consistent with the real data, it is time to choose a better distribution (Fig. 1).
 1027 In this case, most of 100 Gaussian model simulations are out of line with the skew at
 1028 smallest and largest sizes, and excess kurtosis observed at moderately large sizes (Fig.
 1029 S-2 CD). For at least some parts of the size distribution, a non-Gaussian model would
 1030 better capture size transitions.

1031 We sought a distribution that could accommodate the observed changes in the sign
 1032 of skewness and excess kurtosis. We chose the sinh-arcsinh (SHASH) distribution, a
 1033 four-parameter distribution that, conveniently, is included in **mgcv**'s **gam()** function.
 1034 For consistency with the Gaussian for location and scale, specification of basis functions
 1035 ($k = 4$) is limited to parameters for skewness and kurtosis:

```
1036 fitSHASH <- gam(list(logarea.t1 ~ s(logarea.t0), # <- location
1037   ~ s(logarea.t0), # <- log-scale
1038   ~ s(logarea.t0,k=4), # <- skewness
1039   ~ s(logarea.t0,k=4)), # <- log-kurtosis
1040   data = XH, gamma = 1.4, family = shash, optimizer = "efs")
```

1041 The fitted model's mean and variance are nearly identical to the Gaussian (Fig. S-2AB),
 1042 and the fitted trends in skewness and kurtosis are much less "wiggly" than the estimate

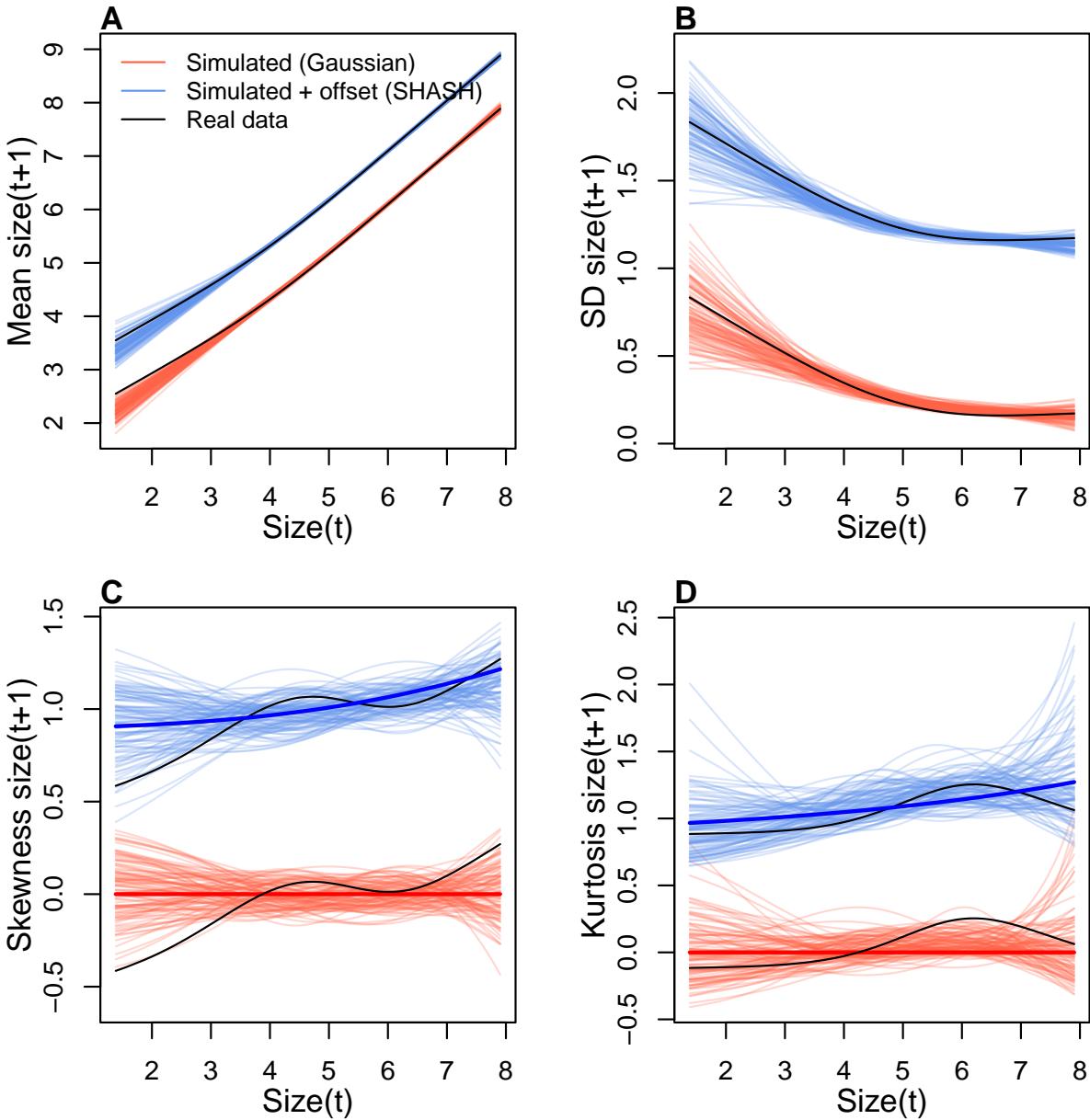


Figure S-2: Comparisons among real coral data and data simulated from Gaussian and SHASH growth models for mean, standard deviation, NP skewness, and NP excess kurtosis of future size conditional on current size. Note that plotted values for the SHASH are offset by one unit to allow comparisons. In the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

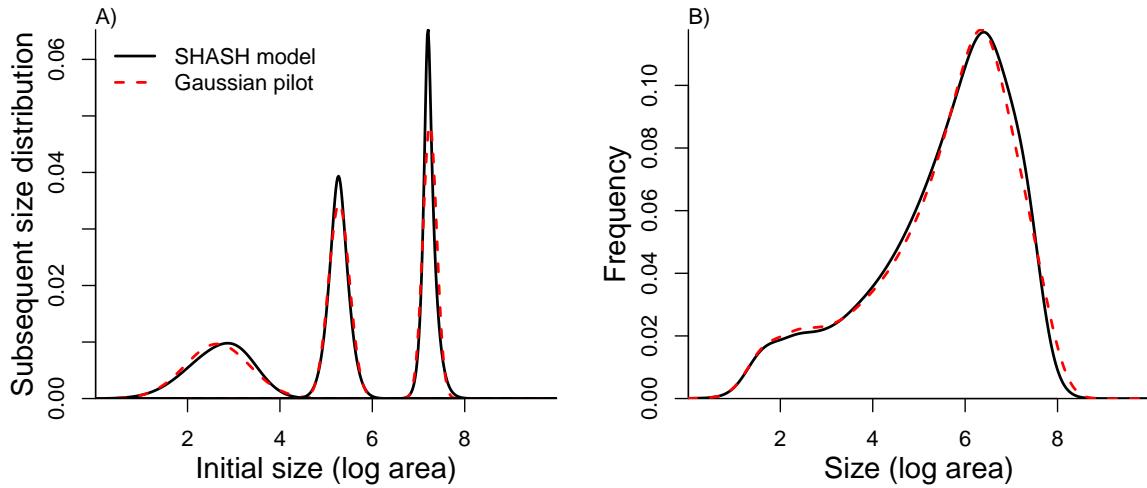


Figure S-3: Comparisons between the fitted SHASH growth model (solid black curves) and the Gaussian pilot model (dashed red curves) for sea fans *G. ventalina*. A) Predicted frequency distributions of size in year $t + 1$ for three different values of size in year t . The leftmost pair of curves are 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.

1046 the difference between models, the SHASH model is clearly favored by AIC ($\Delta AIC =$
 1047 5.45) despite having twice as many parameters to fit.

1048 What, then, have we gained by fitting a better growth model? Fig. S-3A compares
 1049 the predicted distributions of subsequent size in the fitted model and Gaussian pilot
 1050 models, for the median size of a new recruit (leftmost pair of curves), the median initial
 1051 size (central curves), and the 95th percentile of initial size in the data (rightmost
 1052 curves). The differences are small, and most pronounced for the smallest size, where recruits
 1053 are predicted to grow slightly larger under the SHASH model than the Gaussian
 1054 model. The direction of this difference was surprising, because the SHASH has negative
 1055 skew at small sizes in the data. However, the SHASH model also gives a better prediction
 1056 of mean growth at small sizes than the Gaussian model. At intermediate sizes the
 1057 predictions are nearly identical; at large sizes the SHASH has slightly lower standard
 1058 deviation, but fatter tails (excess kurtosis). Fig. S-3B shows the predicted steady-state
 1059 size distributions resulting from a constant unit input of recruits. Again, the differences

1060 are very subtle. Finally, the Gaussian and SHASH growth models predict very similar
1061 mean life span (17.7 and 17.9 years, respectively).

1062 From these outputs, there is little evidence that improved modeling of coral growth
1063 meaningfully improved biological inferences from the IPM. One could argue that it was
1064 not worth the trouble, even though it was almost no trouble at all. But before fitting
1065 the SHASH model, we could not have known whether or not it would have made a
1066 difference.

1067 In this case study we used `gam` to fit both the Gaussian and SHASH models because
1068 that obviated model selection on functions for mean, variance, and higher moments.
1069 However, `gam` should be used with caution. Nonparametric regression models notori-
1070 ously “wag their tails” because the ends of the fitted curve can be pulled close to the
1071 outermost data points. This is especially problematic for growth modeling, because data
1072 are typically sparse near the bounds of the size distribution. To minimize the risk of
1073 overfitting we specified the number of “knots” ($k=4$) and used $\text{gamma}=1.4$ to overweight
1074 model degrees of freedom as suggested by Gu (2013, sec. 3.2). But it is always impor-
1075 tant to plot the fitted splines and make sure they do not wag unrealistically. If they do,
1076 parametric regression may be a better choice.

1077 S.3.2 Case study: creosotebush, *Larrea tridentata*

1078 Our next case study comes from our studies of the woody shrub creosotebush (*Larrea tri-*
1079 *dentata*) at the Sevilleta Long-Term Ecological Research (LTER) site in central New Mex-
1080 ico, US. At this site as elsewhere in the Southwest US, creosotebush is encroaching into
1081 desert grassland habitats. The data described here were collected along transects span-
1082 ning grass-shrub ecotones to understand patterns of density dependence in creosotebush
1083 demography. Specifically, we asked whether fitness is maximized approaching zero den-
1084 sity at the leading edge of the expansion front (consistent with ‘pulled’ expansion), or
1085 whether there is a demographic advantage for shrubs at higher density due to positive
1086 feedbacks expected for ecosystem engineers (leading to ‘pushed’ expansion). Our pub-
1087 lished study (Drees et al., 2023) used a spatial integral projection model (SIPM) to predict
1088 the speed of shrub encroachment, assuming normally-distributed size transitions. Here
1089 we step through our suggested workflow to ask whether a non-Gaussian model would
1090 have been more faithful to the data, and how such an improvement would influence
1091 predictions for the speed of encroachment.

1092 Growth data come from 522 shrubs censused longitudinally over four years (2013-
1093 2017). Census individuals occurred along 12 replicate transects (200 to 600 m in length)

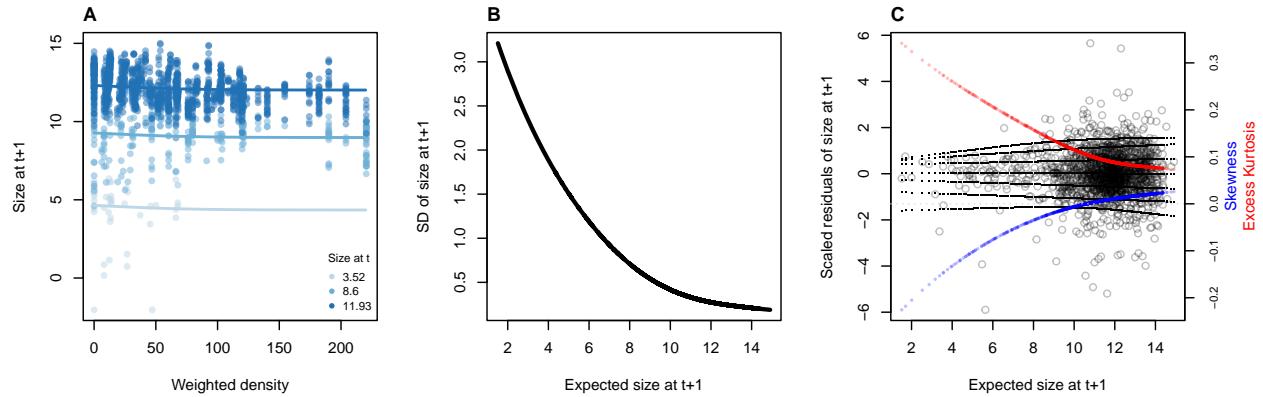


Figure S-4: **A**, Creosotebush size transition data with respect to initial size (colors) and local weighted density (sum of sizes of all plants within a five-meter transect window). Size is quantified as the natural logarithm of plant volume (cm^3). **B**, Standard deviation of size at time $t + 1$ 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.

```
1094 LATR_GAU <- gam(list(log_volume_t1~ s(log_volume_t,k=4) +
1095   s(dens_scaled,k=4) + s(unique.transect,bs="re"), ~ 1),
1096   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).

1113 The model accounts for non-constant variance, which indicate greater dispersion for
1114 smaller values of expected size (Fig. S-4B). Quantiles of the standardized residuals indi-
1115 cate that skew and excess kurtosis are both greater at smaller sizes (Fig. S-4C). Skewness
1116 is close to zero for larger plants (the best-sampled size range) but excess kurtosis re-
1117 mains positive for large plants (ca. 10% heavier tails than Gaussian). As a candidate
1118 for improvement, we turned to the Johnson's S_U (JSU) distribution, a four-parameter,
1119 leptokurtic distribution capable of skew in either direction.

1120 Following our suggested workflow, rather than re-fitting a JSU model from scratch,
1121 we parameterize a model where the residuals from the Gaussian model are fitted by
1122 a JSU distribution. This is relatively easy because the **gamlss.dist** package provides a
1123 parameterization of the JSU in which the location parameter μ is the mean and scale
1124 parameter σ is the standard deviation (Rigby et al., 2019). We fit the "hybrid" model by
1125 writing a likelihood function that uses the fitted mean and standard deviation functions
1126 from Gaussian pilot model, and estimates the parameters that control skewness and
1127 kurtosis as linear functions of predicted future size. The "hybrid" likelihood looks like
1128 this:

```
1129 JSULogLik=function(pars){  
1130   dJSU(LATR_grow$log_volume_t1,  
1131     mu=LATR_grow$GAU_mean,  
1132     sigma=LATR_grow$GAU_sd,  
1133     nu = pars[1]+pars[2]*LATR_grow$GAU_mean,  
1134     tau = exp(pars[3]+pars[4]*LATR_grow$GAU_mean), log=TRUE)  
1135 }
```

1136 The mean and standard deviation of the JSU are set to those of the best Gaussian
1137 model and parameters controlling skewness and kurtosis were fit independently, follow-
1138 ing our approach to the orchid data. The hybrid JSU model performed well, generating
1139 simulated data that aligned with the real data better than the best Gaussian model, par-
1140 ticularly in the standard deviation and kurtosis (Fig. S-5). The JSU model has exactly the
1141 same mean and standard deviation of future size as the Gaussian, but Fig. S-4 uses the
1142 quantile-based nonparametric mean and standard deviation. The results show that even
1143 though the JSU was not fitted to match those, it comes closer than the Gaussian model
1144 as a result of accounting for the skew and kurtosis.

1145 The improvement of the JSU over the Gaussian growth model, while visually sat-
1146 isfying, had only weak influence on SIPM results. The Gaussian model slightly over-
1147 estimated the low-density growth rate, but models using either Gaussian or JSU growth

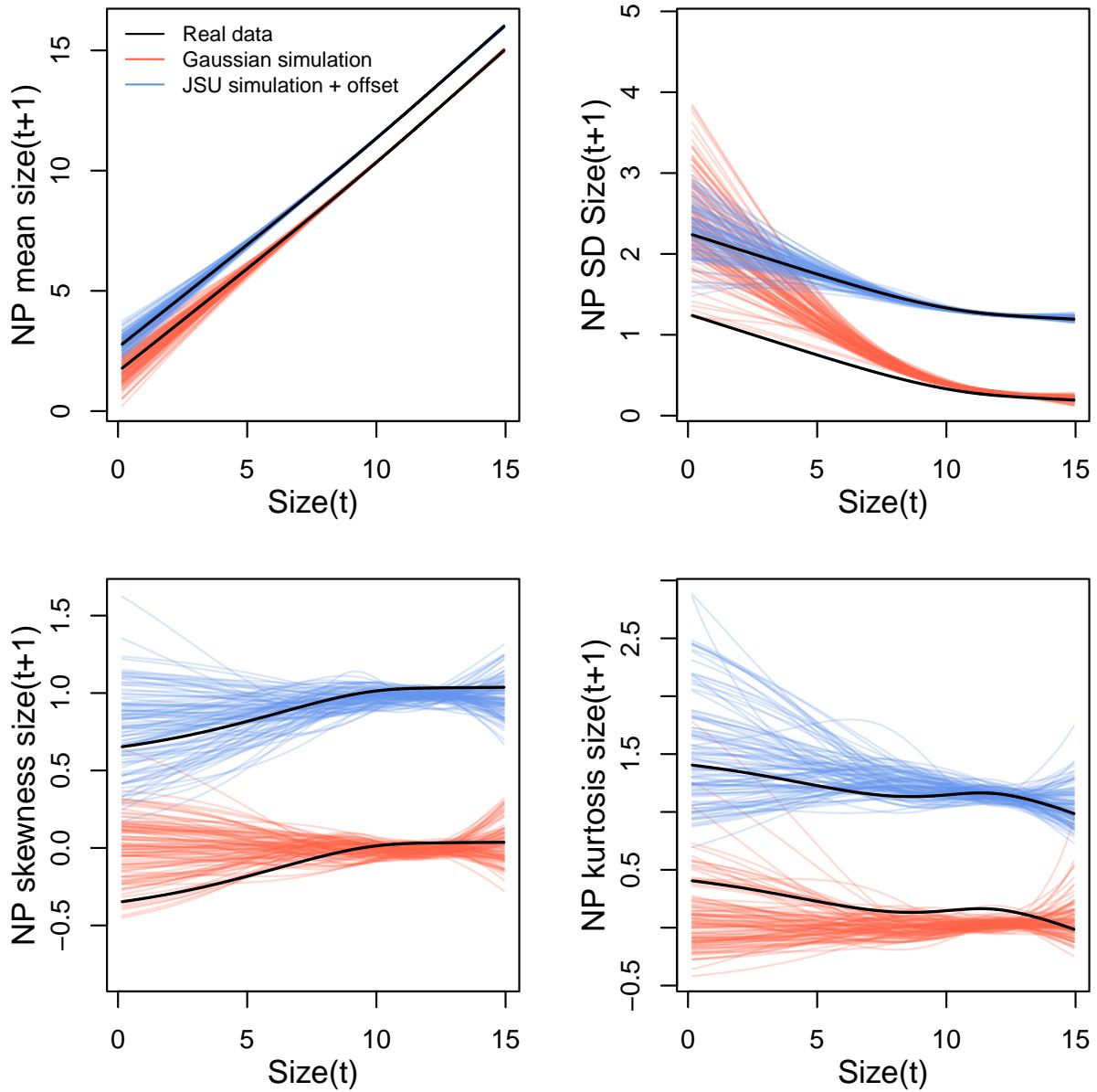


Figure S-5: Comparisons between real creosotebush data and data simulated from Gaussian and JSU growth models for nonparametric measures of mean, standard deviation, skewness, and excess kurtosis of future size conditional on current size. Moments of the future size distribution are plotted with respect 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`.

¹¹⁴⁸ kernels had very similar monotonic decreases in λ with increasing local density, and
¹¹⁴⁹ nearly identical wave velocities (Fig. S-6). This species has very low mortality risk once

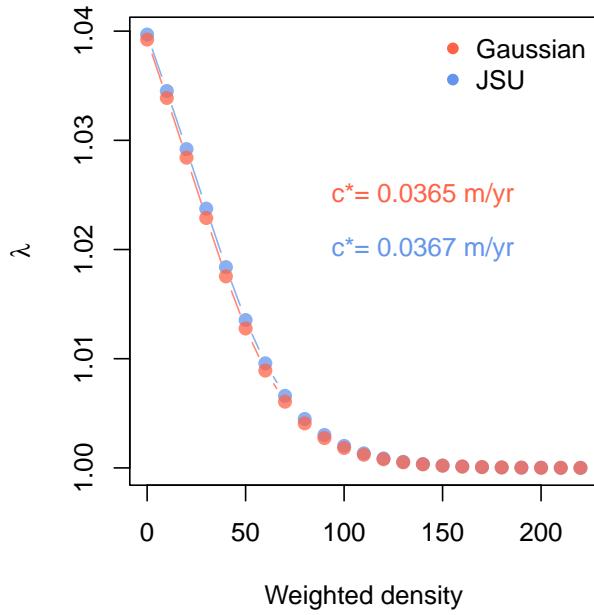


Figure S-6: Density dependence in fitness (λ) and asymptotic velocity of the creosote encroachment wave (c^*) for Gaussian and JSU growth kernels. Weighted density is the sum of sizes ($\log(cm^3)$) of all conspecifics within a five-meter transect “window”. Figure made by script `creosote_growth_modeling_qgam.R`.

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.

1154 S.4 Additional results

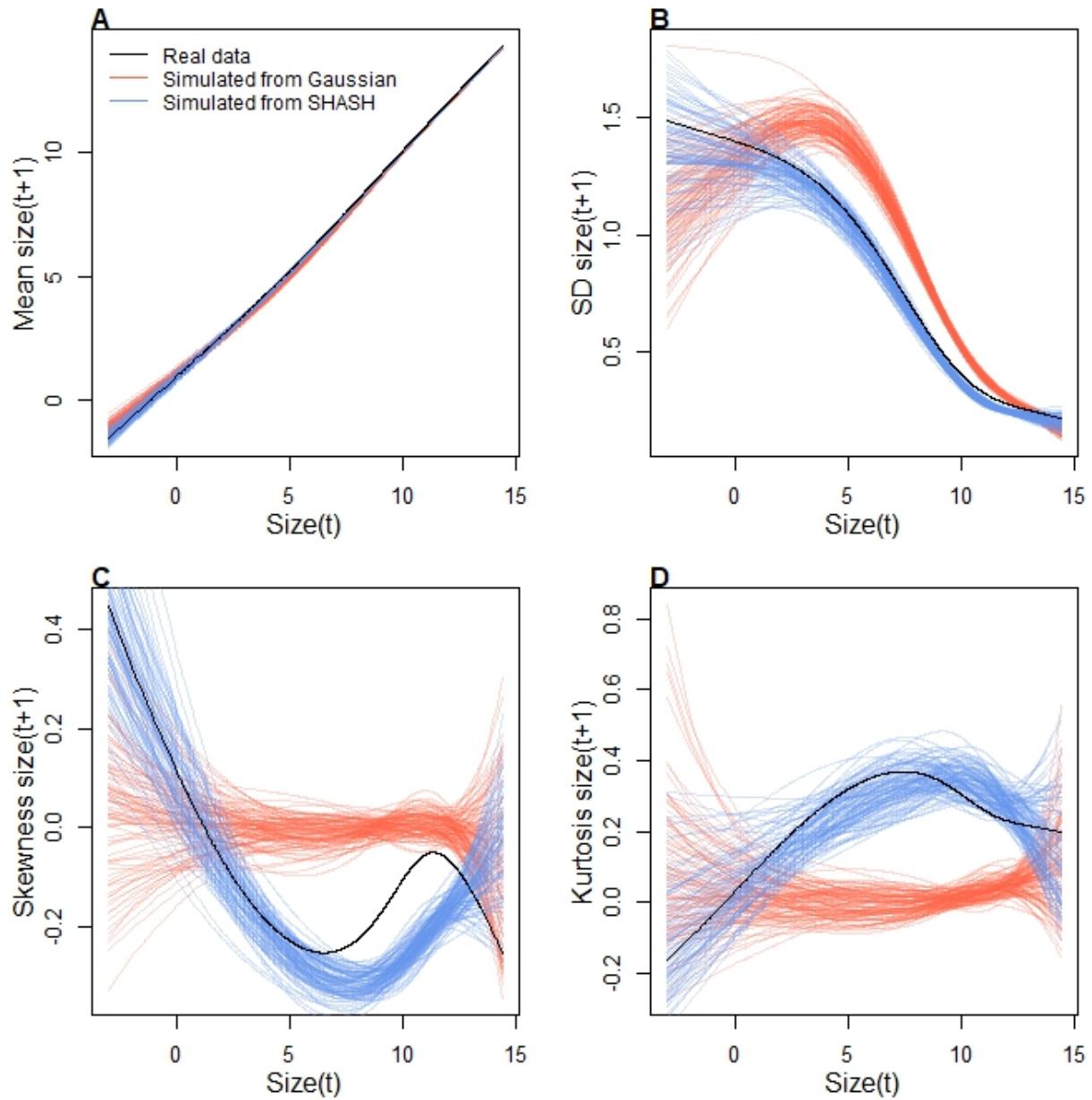


Figure S-7: Comparisons among real cactus data and data simulated from Gaussian and SHASH growth models for mean, standard deviation, NP skewness, and NP excess kurtosis of future size conditional on current size. Figure made by script `cactus_growth_modeling_qgam.R`.

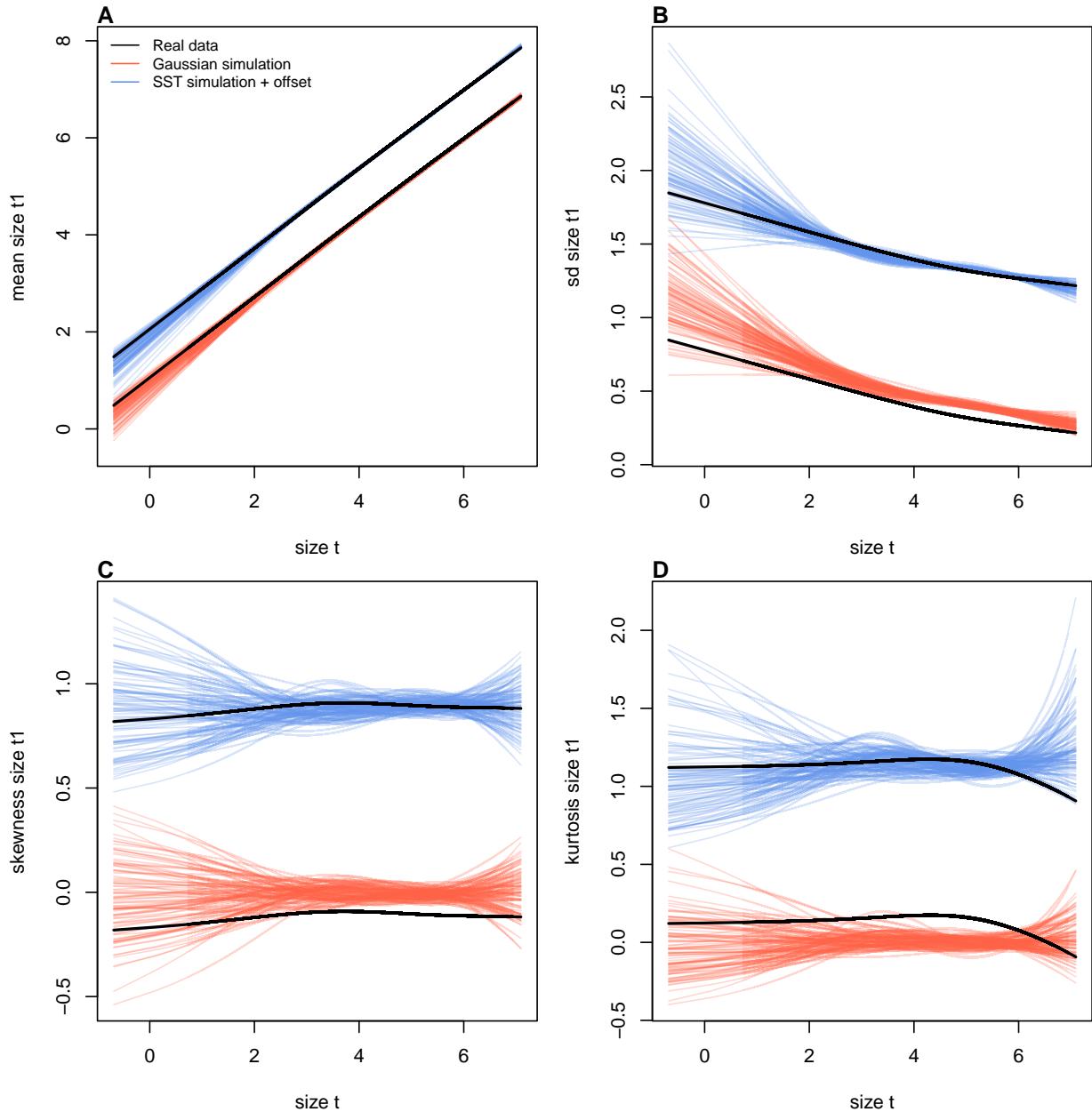


Figure S-8: Comparisons between real orchid data and data simulated from Gaussian and skewed t growth models for mean, standard deviation, NP skewness, and NP kurtosis of future size conditional on current size. Top row (A-D) shows plants that were vegetative at the start of the transition year and bottom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