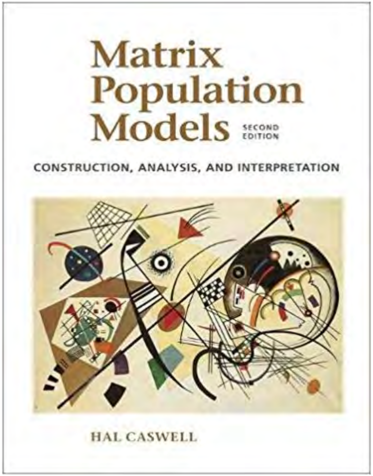
Matrix projection model $n(t+1) = \mathbf{A}n(t)$ divides a population into discrete ages or life-stages ("classes").

$$n_i(t+1) = \sum_j A_{ij} n_j(t)$$

$A_{ij}$ says:

*for each class-$j$ individual "now", how many class-$i$ individuals will be present "next year"?*

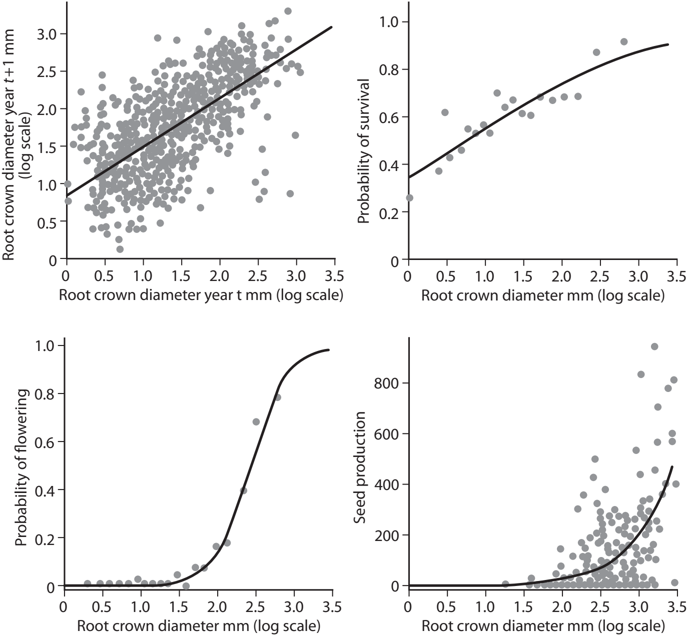
Class-membership must be good individual-level *state variable*: predicts individual fates.




Majority of empirical case studies are not actually age- or stage- structured: the "stages" are size classes.

For plants and many animals (esp. indeterminate growth), size is the best single predictor of demographic fates (survival, fecundity, growth).

Platte thistle,
*Cirsium canescens*
Rose et al. (2005)



# Making a matrix projection model

Modeler chooses divisions between size classes, or uses a hybrid size-stage classification (e.g., size $\times$ $\{$Vegetative, Flowering$\}$).

"Binning": matrix entries are observed transition frequencies. You count up:

- How many in class $j$ were still alive the next year and in size-class $i$?
- How many offspring did they put into each size class?

One cynic: the matrix is a complicated curve-fit that goes exactly through every data point...

# But what's wrong with that? or, how it all started.

Northern Monkshood, *Aconitum noveboracense*

Herbaceous perennial, listed as threatened.

Modeled by Philip Dixon and Bob Cook (then both at Cornell).

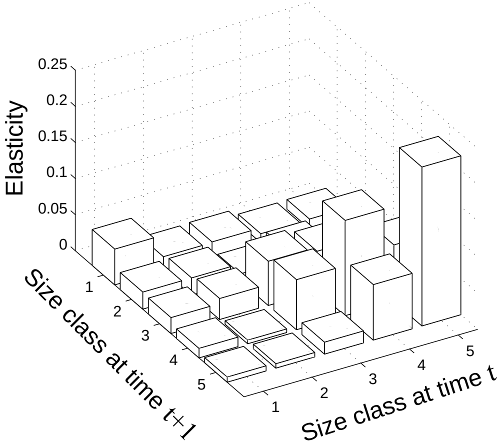
Size = stem diameter. Grow, shrink, or *split*.

$$\begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{bmatrix}_{t+1} = \begin{bmatrix} 0.42 & 0.21 & 0.18 & 0.07 & 0.09 \\ 0.18 & 0.24 & 0.08 & 0.11 & 0.03 \\ 0.13 & 0.32 & 0.33 & 0.21 & 0.04 \\ 0.06 & 0.04 & 0.31 & 0.43 & 0.29 \\ 0.03 & 0.04 & 0.06 & 0.20 & 0.65 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{bmatrix}_{t}$$

Matrix entries are observed transition rates and births: 65% of sampled class-5 individuals were alive the next year, and again in class 5.

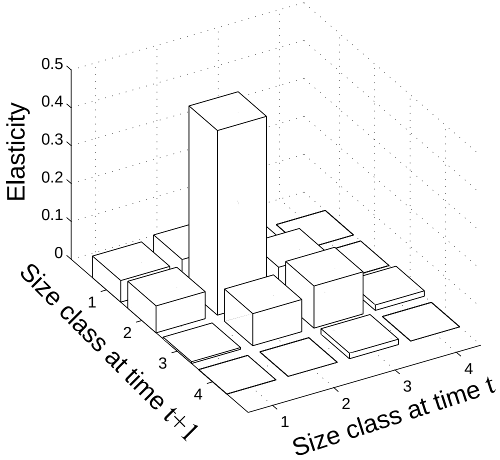
# Eigenvalue elasticity matrix for Monkshood

$$\text{Elasticity}(i,j) = \frac{\text{Percent change in } \lambda}{\text{Percent change in } A_{ij}}$$



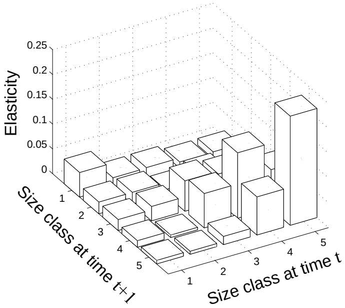

Survival of big plants is most important

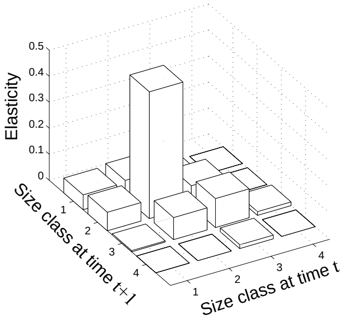
# Eigenvalue elasticity matrix for Monkshood



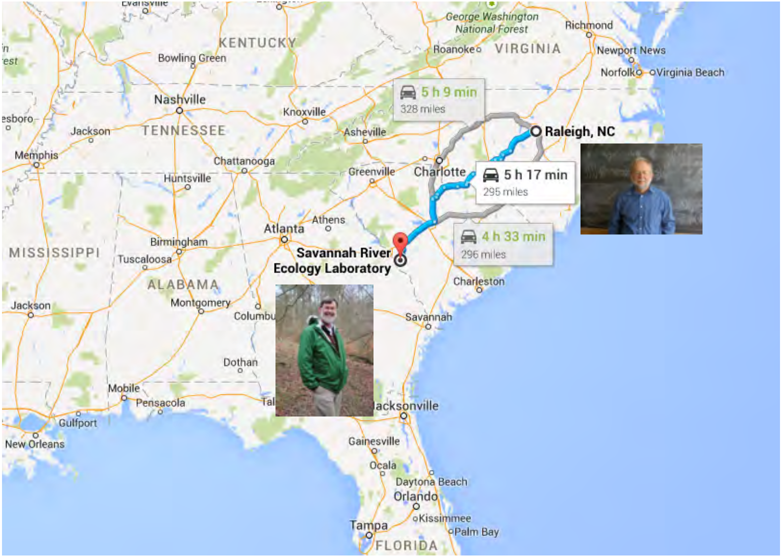

Survival of small plants is most important

It all depends on how you choose your size categories



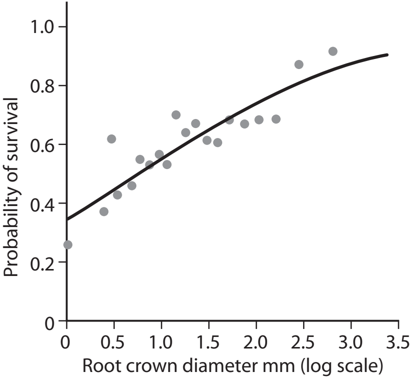



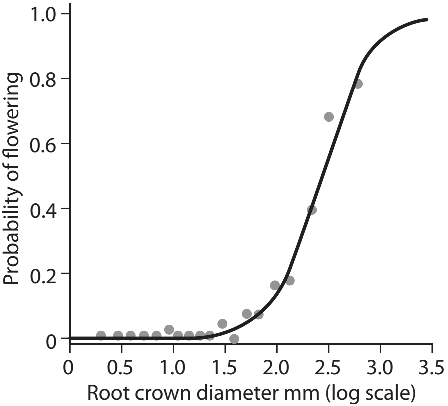
# By the early 1990's...



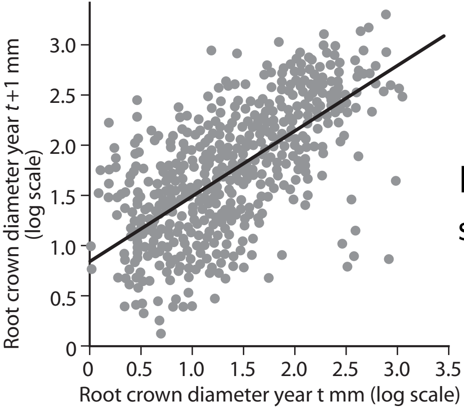
# How should we choose size categories?

# Don't.


Probability of survival
0
0.2
0.4
0.6
0.8
1.0
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Root crown diameter mm (log scale)


Probability of flowering
0
0.2
0.4
0.6
0.8
1.0
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Root crown diameter mm (log scale)

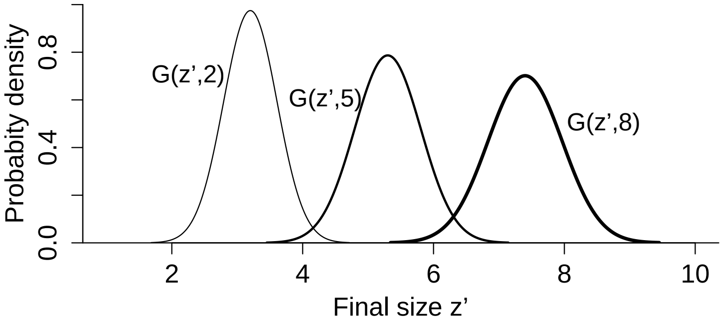
# Same idea for growth




Fitted linear regression of size$(t+1)$ on size$(t)$

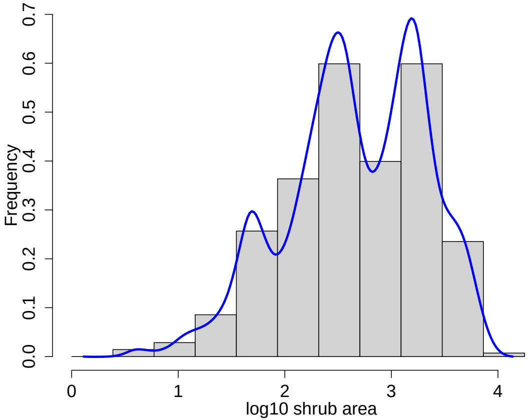
# Size regression equation = dynamic model for size $z$

Size "now" (e.g., $z = 2, 5$ or $8$) determines the probability distribution of size "next year"



# Integral Projection Model (Easterling,Ellner,Dixon 2000)

$n(z,t) =$ distribution of individual size $z$, $L \le z \le U$.



Instead of

$$n_i(t+1) = \sum_j A_{ij} n_j(t)$$

we have

$$n(z', t+1) = \int_L^U K(z', z) n(z, t)\, dz$$

$$K(z', z) = \underbrace{s(z)G(z', z)}_{Survival \ \& \ growth} + \underbrace{F(z', z)}_{Reproduction}$$

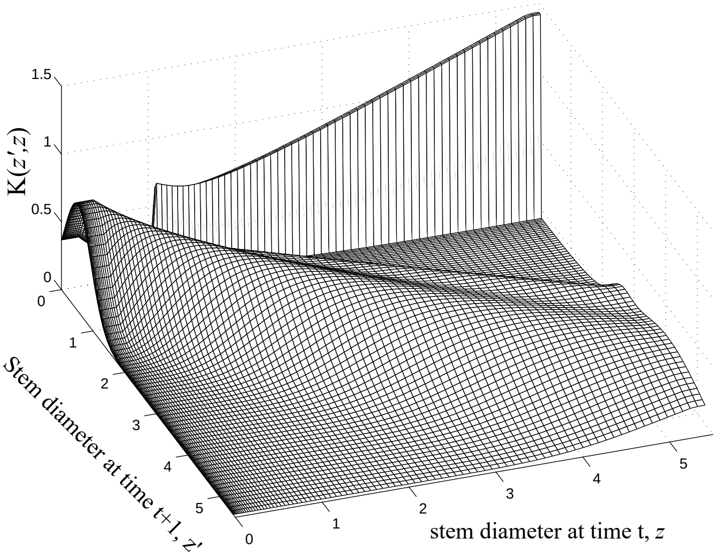
Over the last 10 years, 75% of published demographic models for size-structured populations have been IPMs (D. Doak et al. (2021), Ecological Monographs 91: e01447)

# Monkshood demographic models (Easterling et al. 2000)

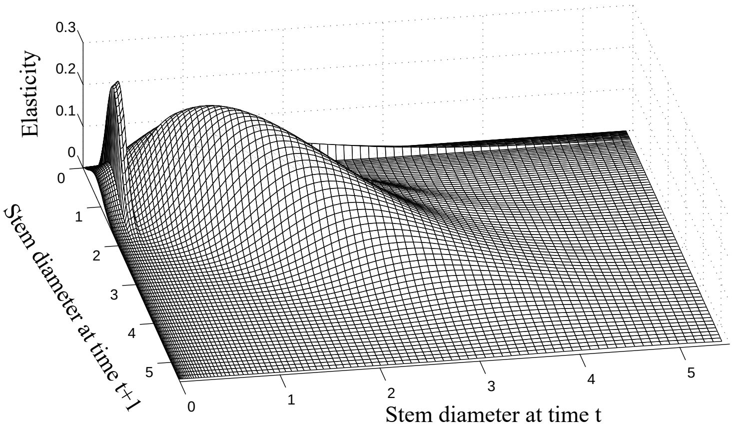
$z$ is plant size (stem diameter in mm), $z'$ is subsequent size.

| Demographic process | Equation |
|---|---|
| Adult Survival | logit $p_s = 1.34 + 0.92z$ |
| Growth | $z' \sim$ Gaussian$(\mu = 0.37 + 0.73z, \sigma^2 = 0.13 + 0.23z)$ |
| Offspring number | $0.034 + 0.38z$ |
| Fraction clonal offspring | $0.39$ |
| Size distribution, clonal | Gaussian$(\mu = 0.3 + 0.57z, \sigma^2 = -0.005 + 0.192z)$ |
| Size distribution, seedlings | Uniform$[0.15, 0.25]$ |

# Projection kernel for Monkshood



# Elasticity surface for Monkshood



# The two key ideas in IPMs

1. If an individual state variable varies continuously, model it as continuous. Conceptual model then aligns with the biology, avoids the never-resolved problem of how to choose size categories.

2. Instead of binning, the model is based on equations describing state-fate relationships – usually statistical models fitted to empirical demographic data. (In principle could be mechanistic models — "DEB meets IPM" — but this is rare).

Demographic data analysis is more work than binning: more time, and a lot more thought. Why should we bother?

# Why model

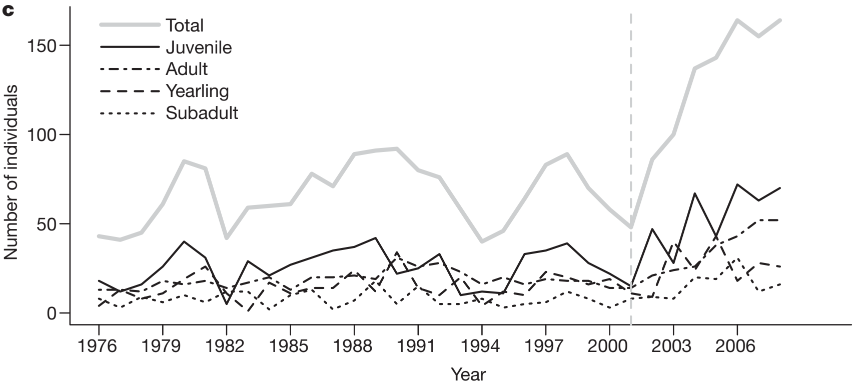
We can use everything statisticians know about fitting smooth functions – which is a lot.

1. Good (and improving) ways to choose model complexity appropriate for sample size and "noise" level: AIC, WAIC, cross-validation, etc. This is a *feature*, not a *bug*. Binning: seat-of-the-pants. Proposed algorithms (Vandermeer, Moloney) are *ad hoc*, never used.

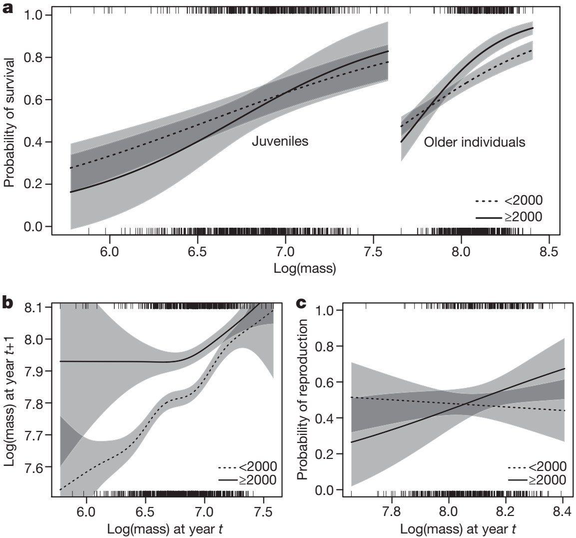
2. Can estimate multiple sources of relevant variation (treatment, age, sex, etc.), and include random effects to increase precision by controlling for unmeasured heterogeneity.

3. Good ways to fit state-fate relationships without assuming a functional form (e.g., splines) — highly under-exploited.


The yellow-bellied marmot (Ozgul et al. 2010)




Marmot population trends, from: Ozgul et al. (2010) Coupled dynamics of body mass and population growth in response to environmental change. Nature 466: 482 - 485.




Model predicts:

$\lambda = 1.02 \rightarrow 1.18$

Larger juveniles and adults (as observed)

Increase in $\lambda$ after 2000 results from:

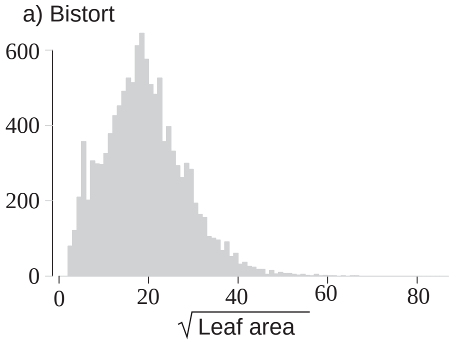
- Adults are larger
- Large adults have higher survival & reproduction.

# Why model: cross-classification and gaps in the data

*Ecological Monographs*, 91(2), 2021, e01447


A critical comparison of integral projection and matrix projection models for demographic analysis

Daniel F. Doak,[1,18] Ellen Waddle,[2] Ryan E. Langendorf,[3] Allison M. Louthan,[4,5] Nathalie Isabelle Chardon,[6] Reilly R. Dibner,[7] Douglas A. Keinath,[7,8] Elizabeth Lombardi,[9] Christopher Steenbock,[10] Robert K. Shriver,[11] Cristina Linares,[12] Maria Begoña Garcia,[13] W. Chris Funk,[14] Sarah W. Fitzpatrick,[15] William F. Morris,[16] and Megan L. Peterson[17]




If your sample data look like this ($n = 11,882$, no gaps), "binning" to make a $20 \times 20$ size-structured matrix is equivalent to making an IPM.

# Cross-classification and gaps in the data

Most data sets are (lots) smaller, but Doak et al. (2021) "showed" that the same is true with hundreds of observations instead of thousands. Who needs an IPM?

# Cross-classification and gaps in the data

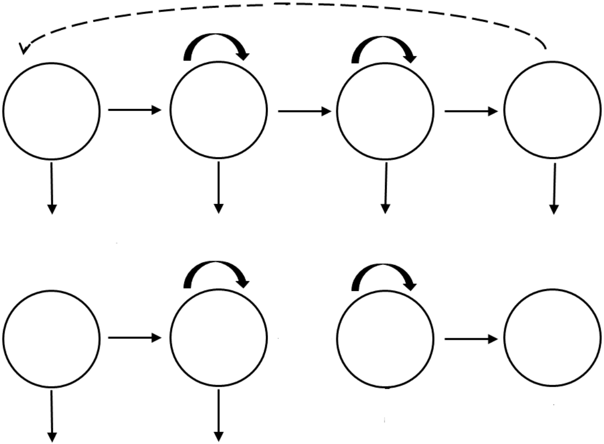
Most data sets are (lots) smaller, but Doak et al. (2021) "showed" that the same is true with hundreds of observations instead of thousands. Who needs an IPM?

Doak et al. (2021) used stratified sub-sampling to guarantee that regardless of sample size, *each size class was represented in proportion to its abundance in the full sample.* Because if not...

Published Leslie matrix for *S. jarrovi*:

100% survival at ages 6, 7

0% survival at age 8



$\approx 25\%$ of published matrix models have biologically implausible discontinuities; many have $\lambda = 1$ exactly (as above).

I. Stott, S. Townley, D. Carslake, and D. J. Hodgson. 2010. On reducibility and ergodicity of population projection matrix models. Methods in Ecology and Evolution, 1:242–252.

Subsample bistort data: each quintile of size distribution represented in proportion to abundance in full sample. Compare IPM with 10-size-class “binning” MPM.

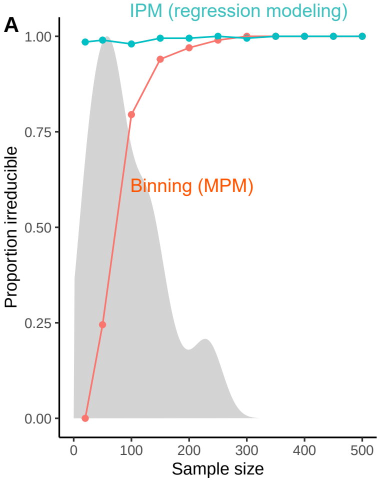


**Table S1** Number of individuals observed annually in four populations of the perennial herb *Astragalus scaphoides*

| | No. individuals observed | | | |
|---|---|---|---|---|
| Year | Sheep | McDevitt | Haynes | Reservoir |
| 2003 | 126 | 73 | 178 | 184 |
| 2004 | 116 | 69 | 147 | 236 |
| 2005 | 126 | 74 | 126 | 254 |
| 2006 | 110 | 65 | 77 | 235 |
| 2007 | 130 | 61 | 83 | 239 |
| 2008 | 121 | 63 | 77 | 228 |
| 2009 | 134 | 55 | 104 | 223 |
| 2010 | 150 | 49 | 132 | 220 |
| 2011 | 135 | 42 | 153 | 194 |
| 2012 | 61 | 29 | 144 | 122 |
| 2013 | 45 | 11 | 114 | 110 |
| 2014 | 19 | 7 | 84 | 63 |

Satu Ramula, Natalie Z. Kerr, Elizabeth E. Crone (2020). Using statistics to design and estimate vital rates in matrix population models for a perennial herb. Population Ecology 62:53–63.

# Multiple sources of variation: random year effects

```
require(mgcv);
mixed.surv = gam(surv~ logsize + s(year,bs="re"),
      family=binomial, data=X);

fdata =X[X$surv==1,];
mixed.flow = gam(flower~logsize + s(year, bs="re"),
      family=binomial, data=fdata);

gdata =X[cdata$surv==1,];
mixed.grow = gam(logsize1~ s(logsize) + s(year,bs="re"),
      data=gdata);
```

# Fitting a stochastic MPM

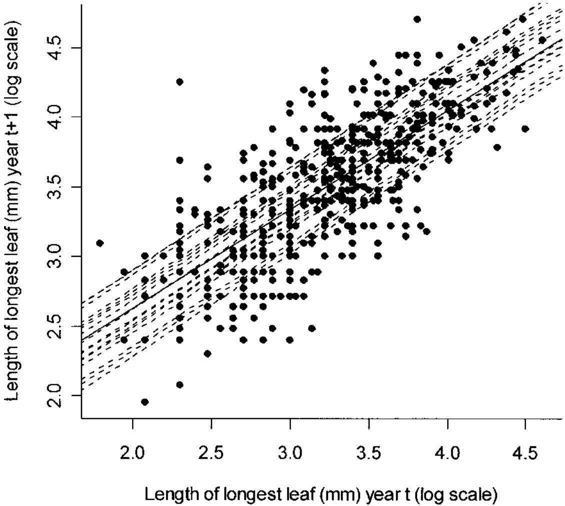
Binning needs BIG samples: good representation for (# size classes $\times$ # years) matrix entries.

Statistical modeling has to include many correlations: good year for growth is probably good for everyone, and "bad" for shrinkage

*Astragalus tyghensis* (Kaye & Pyke 2003, site 25), 9 annual matrices

- For same-direction pairs of growth transitions, 80% of significant correlations were positive.
- For opposite-direction pairs, 100% of significant were negative.

$4 \times 4$ matrix (survival, growth, flowering) had 234 parameters for means, variances, covariances. Fitted IPM had 11 parameters.




Four parameters in IPM: slope, intercept mean and variance, error variance.

KE Rose, M Rees, PJ Grubb (2002) Evolution in the real world: stochastic variation and the determinants of fitness in Carlina vulgaris. Evolution 56:1416–1430.

# This is NOT the model.

$$n(z', t+1) = \int_L^U K(z', z) n(z, t) \, dz$$

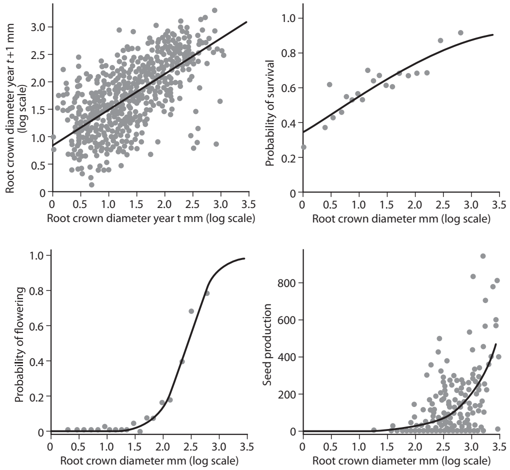
# THIS is the model.

Summary of Platte thistle demography. $z, z'$ are current and subsequent sizes; $D$ is damage by herbivores.

| Demographic process | Equation |
|---|---|
| Growth | $z' \sim \text{Gaussian}(\mu = 0.83 + 0.69z, \sigma^2 = 0.19)$ |
| Adult Survival | $\text{logit } p_s = -0.62 + 0.85z$ |
| Flowering probability | $\text{logit } p_f = -10.22 + 4.25z$ |
| Mean number of seeds | $\log f_n = 0.37 + 2.02z - 1.96D$ |
| Establishment probability | $p_e = 0.067$ |
| Seedling size | $z \sim \text{Gaussian}(\mu=0.75,\ \sigma^2 = 0.17)$ |

K.E. Rose, S. M. Louda, and M. Rees. 2005. Demographic and evolutionary impacts of native and invasive insect herbivores: A case study with Platte thistle, *Cirsium canescens.* Ecology 86: 453–465.

# Better yet: THIS is the model you show to managers




We can visualize the model and interrogate it like any other statistical model: residual plots, test against simpler/more complex alternatives, etc.

$n(z', t+1) = \int_L^U K(z', z) n(z, t) dz$ is *one way* to implement the model numerically.

# Another way: as an ABM/IBM

- An IPM is stochastic at the level of the individual: survival probability; probability distribution for size at time $t+1$.

- But it is deterministic at the level of the population (or deterministic conditional on covariates): no *demographic stochasticity*, the randomness due to unpredictability in the fate of any one individual.

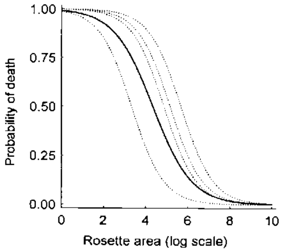
- The regression equations that define an IPM also define an "equivalent" ABM that includes demographic stochasticity.

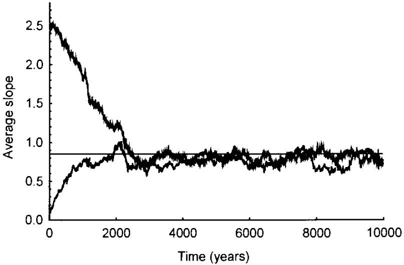
The ABM is more realistic; the IPM runs a lot faster.

# This is where Mark and Dylan came on board....

While we were thinking about Monkshood conservation. . .

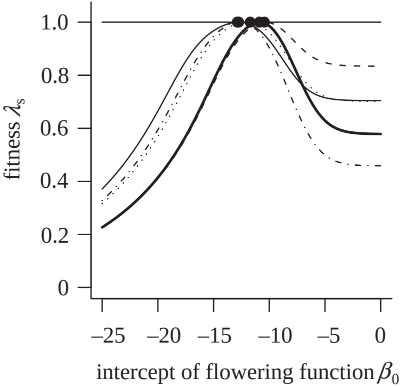
# Mark was simulating life history evolution, using ABMs defined by fitted regression equations



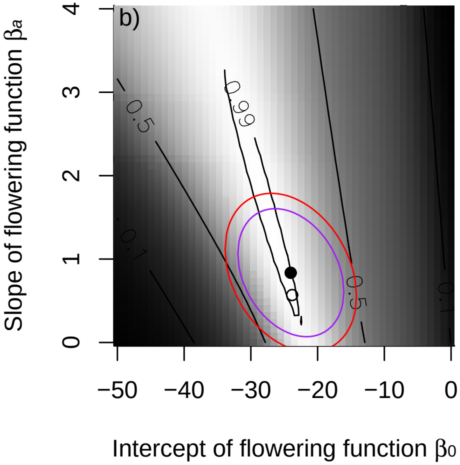

Mark Rees, Andy Sheppard, David Briese, and Marc Mangel (1999). Evolution of Size-Dependent Flowering in *Onopordum illyricum*: A Quantitative Assessment of the Role of Stochastic Selection Pressures. American Naturalist 154: 628–651.

# Finding ESS's using IPMs defined by the same regression equations (ESS $(\beta_0, \beta_1)$ in logit $p_f(z) = \beta_0 + \beta_1 z$.)




D. Z. Childs, M. Rees, K.E. Rose, P.J. Grubb, and S.P. Ellner (2004). Evolution of size-dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model. Proceedings of the Royal Society of London Series B 271: 425-434.

# Sometimes it works




●: Estimated parameters from field data (with 95%, 99% confidence ellipses)

○: Predicted ESS parameters

Contour lines: strategy fitness when invading the ESS.

What do you gain by doing calculations with an IPM, instead of doing simulations of an ABM?

And what do you lose?

When is an IPM really NOT the best choice?