

REPLY

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

Daniel F. Doak¹  | Ellen Waddle²  | Ryan E. Langendorf^{1,3} |
 Allison M. Louthan^{4,5}  | Nathalie Isabelle Chardon⁶  | Reilly Dibner⁷ |
 Robert K. Shriver⁸  | Cristina Linares⁹  | Maria Begoña Garcia¹⁰  |
 Sarah W. Fitzpatrick¹¹ | William F. Morris⁵ | Megan L. DeMarche¹² 

¹Department of Environmental Studies, University of Colorado, Boulder, Colorado, USA

²Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado, USA

³Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA

⁴Division of Biology, Kansas State University, Manhattan, Kansas, USA

⁵Department of Biology, Duke University, Durham, North Carolina, USA

⁶Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada

⁷Department of Zoology and Physiology, University of Wyoming, Laramie, Wyoming, USA

⁸Department of Natural Resources and Environmental Science, University of Nevada, Reno, Nevada, USA

⁹Department of Evolutionary Biology, Ecology and Environmental Sciences, Institut de Recerca de la Biodiversitat (IRBio), University of Barcelona, Barcelona, Spain

¹⁰Pyrenean Institute of Ecology (CSIC), Zaragoza, Spain

¹¹W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, Michigan, USA

¹²Plant Biology Department, University of Georgia, Athens, Georgia, USA

Correspondence

Daniel F. Doak

Email: daniel.doak@colorado.edu

Handling Editor: Gordon A. Fox

We thank Ellner et al. (2022; henceforth E22) for taking an interest in our recent paper on demographic modeling methods (Doak et al., 2021, henceforth referred to as D21). While the tone of E22's comment might imply otherwise, most of their comment emphasizes points with which we agree and made ourselves in D21, and, as we repeatedly acknowledged in D21, are similar to past advice and critiques of demographic modeling by these authors and by others. There are, however, some areas of disagreement or differences in emphasis, and we welcome this opportunity to engage in a constructive dialogue on these points. Overall, however, E22 and D21 offer differing advice on fairly few substantial points about how to build demographic models. This is good

news for ecologists and conservation managers who wish to use these tools to model population processes, estimate the vulnerability of target species, and provide effective management guidelines. Below, we outline our thoughts on E22's main points.

CONCEPTUALIZING THE STEPS IN BUILDING A DEMOGRAPHIC MODEL

E22 provide their Figure 1 as a counter to Figure 1 in D21, an outline of the steps in going from demographic data to a full demographic model. These two depictions

are both simplifications of a fairly involved set of steps, and they are largely the same. In our view, the only substantial difference is that E22's framework does not acknowledge that sometimes there are not enough data, data of the right kind, or a need to do statistical vital rate estimation, either continuous or discrete. For example, some analyses borrow information from past publications or other sources and cannot be statistically modeled or broken into fine size classes (e.g., Crouse et al., 1987; Omsted & Alvarez-Buylla, 1995), and in others the original data collection was done using broad size-based stages rather than continuous size measurements (e.g., Fortini et al., 2022; Kalisz et al., 2014). Yet demographic modeling may still be useful in such cases, especially so for conservation of rare, data-sparse species.

ESTIMATING VITAL RATES WITH STATISTICAL MODELS

E22's first main point is that when dealing with continuous state variables, such as measures of size, vital rates should always be modeled by fitting statistical functions to estimate vital rates. By and large, we agree with their views and arguments, which largely mirror points we made at some length in D21. As they and we have both stated, the careful fitting and evaluation of vital rate functions can have multiple benefits both in understanding demographic patterns and their drivers and in discounting the effects of nuisance variables when parameterizing demographic models.

One distinction that we believe is not well presented in E22 is the dichotomy they draw between discrete vital rate (DVR) estimation and continuous estimation via what E22 refer to as statistical vital rate estimation. As we noted in D21, there are multiple statistical approaches to DVR estimation that can, for example, account for and separate observation error versus process variance, test for demographic differences between size classes or for significant effects of drivers on vital rates, and discount nuisance variables. While we do not believe that we have any disagreement with E22 on this point, it is important to note that practitioners using categorical state variables can also leverage statistical modeling techniques to account for many of the same issues as they can when modeling continuous state variables. It is also possible to use these methods when analyzing discretized stages derived from a continuous state variable. In this sense, we believe that E22's division between "statistical demographic modeling versus discretizing continuous states into a set of contiguous discrete classes" in many cases represents a false dichotomy, although we acknowledge that most often DVR

estimation has been done without the use of these statistical methods.

ADVANTAGES OF DIFFERENT APPROACHES FOR DATA SETS WITH LIMITED SAMPLE SIZES

The most significant disagreement between our groups is over the extent to which moderate to small sample sizes limit the utility of DVR estimation, relative to continuous vital rate estimation (CVR). E22 show in their Figure 2 that DVR can result in nonsensical models where all or no individuals in a class survive or none progress out of the class, due to limited sample size. We do not disagree, and we discussed this problem at length in D21 (see *Results: Simulation results: Sample size effects* in D21). Instead we argue this is an easily diagnosed issue, particularly compared to other potential problems that can arise in demographic modeling, and that this issue should and typically does cause researchers to alter their model structure rather than interpret or publish results from such a model.

E22 state that "Stott et al. (2010) showed that ca. 25% of published DVR-based matrix models include biologically implausible discontinuities in the life cycle (i.e., are reducible and possibly non-ergodic), a consequence of some life-history transitions that actually occur going undetected in a finite sample". While this is in fact what Stott et al. claim, we reviewed the examples of reducible matrices in Stott et al.'s Appendix S1 (see our Appendix S1). The great majority of these matrices do not contain "biologically implausible discontinuities in the life cycle" and the small number that do are typically recognized and discussed by the authors. When these cases are eliminated, the percentage of published projection matrices that are unwittingly or erroneously reducible represent only 0.3% of the cases Stott et al. reviewed, nowhere close to 25%. In Appendix S1, we break down the legitimate reasons that reducible matrices are produced and how they are discussed by the authors reviewed by Stott et al., 2010. Thus we conclude that the rate at which demographers have produced biologically unrealistic matrices without recognizing that they have done so has been greatly exaggerated by both Stott et al. (2010) and E22. Our careful examination of Stott et al.'s results supports our claim that most demographers recognize the problem of small sample sizes in some life stages and adjust the structure of their matrices to avoid problems, just as we did in the analyses in D21.

In D21, we focused on the more insidious problem of a model that appears to give reasonable answers but may be biased in some way. Specifically, we showed that

previously underappreciated aspects of a demographic study, including total sample size, even versus proportional bins, and integration method, often have more important effects on model accuracy than does discrete versus continuous estimation, which has, we believed, received disproportionate emphasis in the demographic literature. For this reason, we contend that the analyses and results in D21 are highly applicable to the question of how to construct future models and in interpreting the published demographic literature. Although DVR models are sometimes published with serious flaws (Kendall et al., 2019), we note that there has been no comparable evaluation of the frequency of potentially equally serious problems in published CVR models. It seems highly likely that such errors are as or more common with these models. For example, the problem of eviction from integral projection models (IPMs) that was highlighted by Williams et al. (2012) was undiagnosed in at least several earlier published IPMs, with important implications for biological inference (see Table 1 in Williams et al., 2012). In addition, many of the errors highlighted in Kendall et al., such as an unnecessary lag in reproduction, are as easy to commit for IPMs as for matrix models, involving in both cases the translation of biological understanding into the correct model structure.

Two related points are also worth clarifying. First, E22 suggest that our simulation procedures are unrealistic in the ways that they subsample data, misrepresenting how much small sample sizes will actually limit the utility of DVR approaches. We disagree with this criticism, as we feel that the procedures we used are a reasonable approach to the creation of rarified data sets in an automated and replicated fashion. While it is certainly true that other procedures would create more data gaps that could cause failure of DVR estimates, as we discussed in D21 and in the preceding paragraphs, most of these problems would be avoided in most empirical studies by shifting bins or in the field sampling itself (e.g., targeting individuals in data-deficient bins). E22 certainly disagree with us on this point, and we feel that readers, especially empirical demographers, will have to decide for themselves which viewpoint is more valid and useful.

Second, E22 argue that the simulations used in D21 are unrealistically simplistic in that we focused on demographic models that use only size as a classifier, whereas many demographic analyses also include other sources of variation (e.g., time periods, plots/transects/sites, and other cross-classified state variables). As we discuss in D21, we agree that taking a strictly DVR approach in these situations will generally create more problems more quickly, due to limited sample sizes, than when using CVR functions. However, when there are strongly nonlinear or threshold responses of vital rates to size,

CVR models have their own disadvantages. While DVR-based models will suffer from reduced sample sizes for vital rate estimation, they also offer more flexibility than the most flexible CVR fits, such as nonlinear smoothing functions, because patterns of variation in vital rates across stages are not bound to any assumptions about smooth functional forms. While CVR-based models are able to “borrow strength” to estimate fewer parameters, in practice demographers have often relied on quite simple parametric functional forms (e.g., fecundity as a linear or quadratic function of size) and particular distributional assumptions for vital rates (e.g., that individuals of a given size this year will follow a Normal size distribution next year). Further, as is well-known, with limited sample sizes model selection procedures will tend to select simpler functional forms. While this is statistically justified, it can also be biologically unrealistic, essentially using statistical criteria to make a decision that could be improved with often-informal biological understanding. If demographers, some of whom are mathematically seasoned and statistically savvy but others of whom are closer to novices, may be susceptible to fitting illogical DVR-based models with limited sample sizes, we should also acknowledge that we as a group are just as susceptible to parameterizing CVR-based models that are poor approximations of biologically complex patterns.

MODEL COMPLEXITY, ASSUMPTIONS, AND THE PROBABILITY OF ERRORS

A core disagreement between our groups is the assertion by E22 that CVR approaches require fewer simplifying assumptions and are less error prone than DVR approaches. We continue to believe that this is highly debatable and that, in many cases, DVR models are more straightforward to implement than are CVR models. E22 seem to imply that fitting generalized linear (mixed) models and using advanced numerical integration methods are easier to do correctly and less prone to mismatches with underlying biology than DVR models. We strongly suspect, based on our collective experiences as reviewers, advisors, and collaborators, that the rate of serious errors in IPMs is at least as high as for DVR models. Relevant to this point, we found in D21 that published descriptions of IPMs frequently lack important details regarding their structure, discretization, and analysis, making it hard to assess whether mistakes have been made in their construction. To be absolutely clear, we are not criticizing CVR models as a useful tool nor questioning the fact that in many cases they will be the best approach to modeling demographic data. We are,

however, arguing that they are not a panacea, and that there are circumstances in which DVR approaches will work extremely well and CVR approaches are probably not required.

METHODS FOR REPRESENTING GROWTH IN CVR-BASED MODELS

E22 state that one of their key pieces of advice is to “construct the population model that is implied by the vital rate models, and implement it numerically using accurate methods.” We certainly agree with this general advice. One goal of our analyses in D21 was to show the importance of thinking carefully about how to represent growth probabilities when discretizing models, as this is an area where there can easily be improvements in the construction and presentation of these models. The two most common methods are the mesh point method (that E22 refer to as the midpoint rule) and the cumulative distribution function (CDF) difference method. Our literature review showed that very few IPM studies even reported which integration method was used. In their Figure 3A,B, E22 show the midpoint rule performs well at large class numbers (as we say in D21), but many IPMs use 100 classes or less and the default for *IPMpack*, a widely used R package for IPM modeling, is 50. Further, as E22 point out in their Appendix S2, the use of a large number of classes with multiple continuous state variables or covariates (“dimensions”) can quickly result in astronomical model dimensions (e.g., E22 state that “with as few as 30 mesh points in each of three dimensions, the discretized kernel has dimension 27,000, and over 700 million entries”). Incorporating parameter uncertainty via bootstrapping and/or stochastic variation in vital rates into such a model would quickly exceed the capacity of the laptop or desktop computers used by the vast majority of ecologists and conservation biologists. Given these considerations, we continue to advocate for the CDF-difference approach, which provides more accurate results with lower numbers of classes. There may be better approaches on the horizon, such as the ones developed in the unpublished dissertation that E22 mention in their appendix. However, perhaps not surprisingly, we focused on the methods that are currently available and accessible for most demographers. Finally, while we agree that the best practice is to test how many size classes are needed for results to converge, as we stated in D21, our literature review shows that authors of demographic papers rarely say whether they have done so, either for models deemed Matrix Projection Models (MPMs) or IPMs by their creators. In short, we are glad that E22 agree with the advice we give in D21 to test how

class or mesh point number influences results, advice that we acknowledged has been given before by others, including the E22 authors, but still is rarely followed. We also re-emphasize that demographic studies using CVR functions should explicitly outline how they treat growth probabilities for analysis.

In a related point, E22 use Figure 3C to suggest that the midpoint rule can sometimes out-perform the CDF difference method. As their Figure 3A,B shows, and as we presented in D21, the CDF difference approach is more robust at smaller class numbers for some of the most commonly published outputs and processes of demographic models. In contrast, their Figure 3C suggests that the midpoint approach can provide better answers for estimation of the variance of future reproductive output. In Appendix S1, we outline why E22’s analysis does not show what they claim, and that the CDF difference method actually performs better than the midpoint method for calculating variation in lifetime reproductive output, using both the simplified model E22 analyze and a full demographic model analysis. While we agree with their most general point, that sometimes different methods are needed for different questions, they do not provide any well-supported evidence that the midpoint method outperforms the CDF difference approach for any demographic output.

DIFFERENCES IN EMPHASIS

E22 state that “with few exceptions, we agree with D21 on the factual content of their paper; our disagreements center on interpretations of that material and the consequent recommendations for ‘best practices’.” We agree that many of the differences between E22 and D21 seem largely to be a matter of emphasis on different aspects of demographic modeling. Here we discuss three of these aspects.

First, we feel that E22 largely ignores a key finding of our work: with declining sample sizes, the variance in key model outputs increases just as rapidly for CVR (that is, IPM-type) models as for DVR ones (D21: Figure 14). The simple models we use to generate these results do not feature stochasticity or other complications, and, as we and E22 describe, with these complications, CVR estimation may offer many advantages for analysis over simple DVR estimation. Nonetheless, the view that CVR methods will invariably provide more reliable demographic results, even when data are extremely sparse, is not supported by our analyses. We are also unaware of any other attempt to clearly assess if CVR models generally perform better than DVR models, and E22 provide no such analysis. This result is, we feel, one of the primary lessons of our study: the quantity and quality of

data obtained from field studies have far more important effects on the reliability of demographic results than do most details of modeling approaches.

Second, we argue that the claim that discretizing continuous state variables is “artificial” has received disproportionate emphasis in the demographic literature relative to the many other approximations made in all demographic models. While DVR approaches do indeed categorize continuous state variables, this is only a problem if it matters for the results. As we note in D21, and many others have said more eloquently before us (Box, 1976), the goal of model building is to achieve useful results, not to reflect as accurately as possible all aspects of nature.

Finally, we continue to advocate for an important role of a range of demographic modeling methods, given that they are shown, as in D21, to perform well. Many of us, the authors of D21, work with a diversity of conservation practitioners and on species that range from some of the most demographically well-studied taxa to ones with very limited data. While E22 argue that “a well-rounded population ecologist will eventually need to learn linear algebra and calculus and overcome their initial fears of derivatives, eigenvalues, and integrals,” we remain cognizant of the importance of making demographic models accessible to conservation practitioners and managers who may not have had this educational background or training and do not have the time or monetary resources to pursue such training now. Of course, we recognize that different tools are more or less appropriate for different situations or skill sets, and that results need to be displayed or described to different audiences. Particularly since we do not find evidence that CVR modeling methods deliver strikingly different predictions than do DVR models, we do not see a justification for elevating or denigrating either.

This view, that the accuracy of the predictions of models is the basis to judge their use, was inherent in the approach we took in D21. Whether or not it was the intention of E22’s authors, we are concerned that conservation practitioners might take from E22 the message that, unless vital rates are modeled using advanced statistical analysis, a demographic analysis is not worth doing. We are aware of numerous examples in which very limited data on rare species (e.g., maybe a single year’s data from one site for relatively few individuals that may have been placed into classes from the start or where limited information is due to species inhabiting environments where monitoring is highly challenging) have been used to make simple DVR models that yielded important management insights. A classic example is the matrix model for the loggerhead sea turtle (Crouse et al., 1987), which was based on crude point estimates of fecundity and survival for a few size classes plus a curve of size versus age estimated from a single site, all taken from the literature.

Despite sparse data and a lack of statistical analysis, this study has been cited 779 times as of this writing (according to Web of Science) and, more importantly, the model’s conclusions played an important role in backing legislation regulating fishing activities that threatened to drive the turtles extinct (Crowder et al., 1994). Similarly basic demographic descriptions have also been used to test life history theory and address other conceptual issues in our field (e.g., Young, 1990). It is possible that these and similar analyses were erroneous due to their model construction. That possibility is nearly impossible to assess for either DVR or CVR models, but in our view, the bulk of evidence suggests that these simple matrix models gave valuable insights. Furthermore, use of simpler modeling approaches, where they are appropriate or necessary, can help bridge the gap between those who better understand modeling and those who best understand the complexities of their study species and the constraints on data collection. In these situations, we would hate to see the perfect become the enemy of the good.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Daniel F. Doak  <https://orcid.org/0000-0001-8980-3392>

Ellen Waddle  <https://orcid.org/0000-0001-7084-9040>

Allison M. Louthan  <https://orcid.org/0000-0002-2735-6539>

Nathalie Isabelle Chardon  <https://orcid.org/0000-0001-9120-4778>

Robert K. Shriver  <https://orcid.org/0000-0002-4590-4834>

Cristina Linares  <https://orcid.org/0000-0003-3855-2743>

Maria Begoña Garcia  <https://orcid.org/0000-0003-4231-6006>

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

REFERENCES

- Box, G. E. P. 1976. “Science and Statistics.” *Journal of the American Statistical Association* 71: 791–9.
- Crouse, D. T., L. B. Crowder, and H. Caswell. 1987. “A Stage-Based Population Model for Loggerhead Sea Turtles and Implications for Conservation.” *Ecology* 68: 1412–23.
- Crowder, L. B., D. T. Crouse, S. S. Heppell, and T. H. Martin. 1994. “Predicting the Impact of Turtle Excluder Devices on Loggerhead Sea Turtle Populations.” *Ecological Applications* 4: 437–45.
- Doak, D. F., E. Waddle, R. E. Langendorf, A. M. Louthan, N. I. Chardon, R. Dibner, D. Keinath, et al. 2021. “A Critical Comparison of Integral Projection and Matrix Projection Models for Demographic Analysis.” *Ecological Monographs* 91: e01447. <https://doi.org/10.1002/ecm.1447>.

- Ellner, S., P. Adler, D. Childs, G. Hooker, T. Miller, and M. Rees. 2022. "A Critical Comparison of Integral Projection and Matrix Projection Models for Demographic Analysis: Comment." *Ecology*: e3605. <https://doi.org/10.1002/ecy.3605>.
- Fortini, L. B., P. D. Krushelnycky, D. R. Drake, F. Starr, K. Starr, and C. G. Chimera. 2022. "Complex Demographic Responses to Contrasting Climate Drivers Lead to Divergent Population Trends across the Range of a Threatened Alpine Plant." *Global Ecology and Conservation* 33: e01954.
- Kalisz, S., R. B. Spigler, and C. C. Horvitz. 2014. "In a Long-term Experimental Demography Study, Excluding Ungulates Reversed Invader's Explosive Population Growth Rate and Restored Natives." *Proceedings of the National Academy of Sciences USA* 111: 4501–6.
- Kendall, B. E., M. Fujiwara, J. Diaz-Lopez, S. Schneider, J. Voigt, and S. Wiesner. 2019. "Persistent Problems in the Construction of Matrix Population Models." *Ecological Modelling* 406: 33–43.
- Omsted, I., and E. R. Alvarez-Buylla. 1995. "Sustainable Harvesting of Tropical Trees: Demography and Matrix Models of Two Palm Species in Mexico." *Ecological Applications* 5: 484–500.
- Stott, I., S. Townley, D. Carslake, and D. J. Hodgson. 2010. "On Reducibility and Ergodicity of Population Projection Matrix Models." *Methods in Ecology and Evolution* 1: 242–52.
- Williams, J. L., T. E. Miller, and S. P. Ellner. 2012. "Avoiding Unintentional Eviction from Integral Projection Models." *Ecology* 93: 2008–14.
- Young, T. P. 1990. "Evolution of Semelparity in Mount Kenya Lobelias." *Evolutionary Ecology* 4: 157–71.

SUPPORTING INFORMATION

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

How to cite this article: Doak, Daniel F., Ellen Waddle, Ryan E. Langendorf, Allison M. Louthan, Nathalie Isabelle Chardon, Reilly Dibner, Robert K. Shriver, et al. 2022. "A Critical Comparison of Integral Projection and Matrix Projection Models for Demographic Analysis: Reply." *Ecology* 103(10): e3822. <https://doi.org/10.1002/ecy.3822>