



This is a repository copy of *A standard protocol to report discrete stage-structured demographic information*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/200899/>

Version: Published Version

Article:

Gascoigne, S.J.L., Rolph, S., Sankey, D. et al. (37 more authors) (2023) A standard protocol to report discrete stage-structured demographic information. *Methods in Ecology and Evolution*. ISSN 2041-210X

<https://doi.org/10.1111/2041-210x.14164>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

A standard protocol to report discrete stage-structured demographic information

Samuel J. L. Gascoigne¹  | Simon Rolph²  | Daisy Sankey¹ | Nagalakshmi Nidadavolu¹ | Adrian S. Stell Pičman^{1,3} | Christina M. Hernández⁴  | Matthew E. R. Philpott¹ | Aiyla Salam¹ | Connor Bernard¹  | Erola Fenollosa¹  | Young Jun Lee¹ | Jessica McLean¹ | Shathuki Hetti Achchige Perera¹ | Oliver G. Spacey¹  | Maja Kajin^{1,3}  | Anna C. Vinton¹  | C. Ruth Archer⁵  | Jean H. Burns⁶  | Danielle L. Buss^{7,8}  | Hal Caswell⁹  | Judy P. Che-Castaldo¹⁰  | Dylan Z. Childs²  | Pol Capdevila¹¹  | Aldo Compagnoni^{1,12,13}  | Elizabeth Crone¹⁴  | Thomas H. G. Ezard¹⁵  | Dave Hodgson⁸  | Tiffany M. Knight^{12,13,16}  | Owen R. Jones¹⁷  | Eelke Jongejans^{18,19}  | Jenni McDonald^{20,21}  | Brigitte Tenhumberg²²  | Chelsea C. Thomas²³ | Andrew J. Tyre^{24,25}  | Satu Ramula²⁶  | Iain Stott²⁷  | Raymond L. Tremblay^{28,29,30}  | Phil Wilson⁸ | James W. Vaupel^{31,‡}  | Roberto Salguero-Gómez^{1,32,33} 

¹Department of Biology, University of Oxford, Oxford, UK; ²Department of Animal & Plant Sciences, University of Sheffield, Sheffield, UK; ³Department of Biology, University of Ljubljana, Ljubljana, Slovenia; ⁴Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York, USA; ⁵Institute of Evolutionary Ecology and Conservation Genomics, University of Ulm, Ulm, Germany; ⁶Department of Biology, Case Western Reserve University, Cleveland, Ohio, USA; ⁷Department of Archaeology, University of Cambridge, Cambridge, UK; ⁸College of Life and Environmental Sciences, University of Exeter, Penryn, UK; ⁹Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands; ¹⁰Branch of Species Status Assessment Science Support, U.S. Fish and Wildlife Service, Washington, DC, USA; ¹¹School of Biological Sciences, University of Bristol, Bristol, UK; ¹²Institute of Biology, Martin Luther University Halle-Wittenburg, Halle (Saale), Germany; ¹³German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany; ¹⁴Department of Biology, Tufts University, Medford, Massachusetts, USA; ¹⁵School of Ocean and Earth Science, University of Southampton, Southampton, UK; ¹⁶Department of Community Ecology, Helmholtz Centre for Environmental Research-UFZ, Halle (Saale), Germany; ¹⁷Department of Biology, University of Southern Denmark, Odense, Denmark; ¹⁸Animal Ecology and Physiology, Radboud University, Nijmegen, The Netherlands; ¹⁹NIOO-KNAW, Animal Ecology, Wageningen, The Netherlands; ²⁰Veterinary Department, Cats Protection, National Cat Centre, Haywards Heath, UK; ²¹Bristol Veterinary School, University of Bristol, Bristol, UK; ²²School of Biological Sciences and Department of Mathematics, University of Nebraska, Lincoln, Nebraska, USA; ²³Alexander Center for Applied Population Biology, Conservation & Science Department, Chicago, Illinois, USA; ²⁴School of Natural Resources, University of Nebraska-Lincoln, Lincoln, Nebraska, USA; ²⁵Global Resistance Management Team, Bayer U.S. – Crop Science, Chesterfield, Missouri, USA; ²⁶Department of Biology, University of Turku, Turku, Finland; ²⁷School of Life Sciences, University of Lincoln, Lincoln, UK; ²⁸Department of Biology, University of Puerto Rico, San Juan, Puerto Rico, USA; ²⁹Center for Applied Tropical Ecology and Conservation, University of Puerto Rico, San Juan, Puerto Rico, USA; ³⁰Department of Biology, University of Puerto Rico-Humacao, Puerto Rico, USA; ³¹Interdisciplinary Centre on Population Dynamics, University of Southern Denmark, Odense, Denmark; ³²Centre for Biodiversity and Conservation Science, University of Queensland, St Lucia, Queensland, Australia and ³³Evolutionary Demography Laboratory, Max Planck Institute for Demographic Research, Rostock, Germany

Correspondence

Samuel J. L. Gascoigne
Email: samuel.gascoigne@pmb.ox.ac.uk

Roberto Salguero-Gómez
Email: rob.salguero@biology.ox.ac.uk

Funding information

Maria Zambrano Next Generation

Abstract

1. Stage-based demographic methods, such as matrix population models (MPMs), are powerful tools used to address a broad range of fundamental questions in ecology, evolutionary biology and conservation science. Accordingly, MPMs now exist for over 3000 species worldwide. These data are being digitised as

[‡]Unfortunately, our dear friend and colleague passed before the submission of this paper.

EU Fellowship; Natural Environment Research Council, Grant/Award Number: NE/L007770/1, NE/M018458/1 and NE/N006798/1; NSF, Grant/Award Number: DEB-1933497; NSF Advances in Bioinformatics Development Award, Grant/Award Number: DBI-1661342; European Research Council, Grant/Award Number: 788195

Handling Editor: Res Altwegg

an ongoing process and periodically released into two large open-access online repositories: the COMPADRE Plant Matrix Database and the COMADRE Animal Matrix Database. During the last decade, data archiving and curation of COMPADRE and COMADRE, and subsequent comparative research, have revealed pronounced variation in how MPMs are parameterized and reported.

2. Here, we summarise current issues related to the parameterisation and reporting of MPMs that arise most frequently and outline how they affect MPM construction, analysis, and interpretation. To quantify variation in how MPMs are reported, we present results from a survey identifying key aspects of MPMs that are frequently unreported in manuscripts. We then screen COMPADRE and COMADRE to quantify how often key pieces of information are omitted from manuscripts using MPMs.
3. Over 80% of surveyed researchers ($n = 60$) state a clear benefit to adopting more standardised methodologies for reporting MPMs. Furthermore, over 85% of the 300 MPMs assessed from COMPADRE and COMADRE omitted one or more elements that are key to their accurate interpretation. Based on these insights, we identify fundamental issues that can arise from MPM construction and communication and provide suggestions to improve clarity, reproducibility and future research utilising MPMs and their required metadata. To fortify reproducibility and empower researchers to take full advantage of their demographic data, we introduce a standardised protocol to present MPMs in publications. This standard is linked to www.compadre-db.org, so that authors wishing to archive their MPMs can do so prior to submission of publications, following examples from other open-access repositories such as DRYAD, Figshare and Zenodo.
4. Combining and standardising MPMs parameterized from populations around the globe and across the tree of life opens up powerful research opportunities in evolutionary biology, ecology and conservation research. However, this potential can only be fully realised by adopting standardised methods to ensure reproducibility.

KEYWORDS

comparative demography, matrix population models, open access, reproducibility

1 | INTRODUCTION

Population ecology has come of age. The development of theories, experimental approaches and statistical methodologies have resulted in the publication of demographic information for an increasingly representative sample of the world's biodiversity (De Magalhães & Costa, 2009; Levin et al., 2022; Salguero-Gómez et al., 2015, 2016). These data span the taxonomic tree from microbes (Jouvet et al., 2018) to macro-vertebrates (Fujiwara & Caswell, 2001), and cover virtually all continents and biomes—though with important taxonomic biases (Conde et al., 2019; Römer et al., 2021). The potential of this impressive and rapidly increasing amount of information is starting to be realised. Indeed, through combining these demographic models, researchers have identified functional traits that explain variation in plant life history strategies (Adler et al., 2014; also see Bernard et al., 2023), short-term (transient) characteristics

that drive the demographic dynamics of plant populations in variable environments (McDonald et al., 2016), and ways in which life history strategies allow species to persist alongside a changing climate (Jelbert et al., 2019; Paniw et al., 2019).

One of the most widely used tools for describing and analysing species' complex life histories is the matrix population model (MPM, hereafter). Briefly, in an MPM, individuals of a population are classified by discrete stages and/or ages (st/age hereafter) according to some biological (Caswell, 2001, p. 31) or statistical/sampling criteria (Salguero-Gómez & Plotkin, 2010). These individuals are followed in discrete time steps, typically adjusted by the generation time of the species. Indeed, time steps can vary from 12 to 24 h as in nematode worms *Caenorhabditis elegans* and aphids *Myzus persicae* (Bruijning et al., 2019; Li et al., 2014), to monthly/annual periods in mammals and plants (Coulson et al., 2001; Ferreira et al., 2016), all the way to 50 years in slow-growing red woods (Namkoong & Roberds, 1974).

From these data, researchers estimate losses through mortality, transition probabilities among st/ages and their per-capita a/sexual contributions via reproduction (Nordstrom et al., 2021; Omeyer et al., 2021).

A single MPM can be used to calculate a vast repertoire of biologically meaningful outputs. These outputs include proxies for the performance and viability of populations, such as deterministic (λ) or stochastic population growth rates (λ_s) (Doak et al., 2005), quasi-extinction risk (Davis, 2022), population response to perturbations of underlying vital rates such as survival or reproduction (Caswell, 2001, p. 206), transient dynamics (Capdevila et al., 2020; Ezard et al., 2010; Stott et al., 2011) and life history traits, such as rates of senescence (Baudisch et al., 2013), degree of iteroparity (Salguero-Gómez et al., 2017) and age at maturity (Caswell, 2001, p. 124). This wealth of demographic inference highlights why many advances in demography and life history theory utilise MPMs (Franco & Silvertown, 1996; Pfister, 1998; Sæther et al., 2013; Tuljapurkar, 1989).

MPMs for plants and animals have been archived, error-checked, complemented with additional information (e.g. GPS coordinates, IUCN conservation status), and released open-access in the COMPADRE Plant Matrix Database (Salguero-Gómez et al., 2015) and the COMADRE Animal Matrix Database (Salguero-Gómez et al., 2016). In the latest data release, COMPADRE v. 6.22.5 [COMADRE v. 4.21.8] contains 8851 [3317] MPMs from 760 [415] unique species published in 643 [395] studies. At the time of writing, a further 1307 species are pending digitization in the COMPADRE network, at a rate of 4.5 new works containing MPMs being screened, digitised, and quality checked every week (S. Gascoigne, pers. obs.). However, one of the challenges of the digitization process is the tremendous variation in how data are collected, presented and used to parameterize MPMs.

Data standardisation improves reproducibility and promotes data sharing across research disciplines (Reichman et al., 2011). Data standardisation, and the associated detailed metadata, is therefore key for research to be replicated, validated, openly discussed and ultimately for science to advance (Powers & Hampton, 2019; Reichman et al., 2011; Salguero-Gómez et al., 2021). Examples of these standards include reporting sample size and variance of estimates and detailing the full list of original sources of data (Gerstner et al., 2017). In this context, standards can be used as checklist items to improve publications quality and reproducibility and to aid the peer-review process (Reichman et al., 2011). Furthermore, meta-analyses (Gurevitch et al., 2018) and phylogenetic comparative analyses (Healy et al., 2019; Salguero-Gómez et al., 2017), which offer valuable opportunities to examine general patterns and identify gaps in knowledge, rely on data conforming to certain standards.

MPMs are being adapted, extended and applied beyond their original, species-specific context in comparative demography. However, not all MPMs are built and reported equally. The current presentation of MPMs in COMPADRE and COMADRE may give the false impression that all MPMs are published in a homogeneous format, despite differences in how and why the MPMs are produced (Caswell, 2001). This impression may have emerged from the amount of verification the COMADRE and COMPADRE digitisation team does behind the scenes (e.g. validating model outputs, author

correspondence for additional information). While verification is an inevitable aspect of database curation, most of our efforts are spent communicating with authors rather than digitising data. Our goal here is to (i) present the current standard of MPM communication in the literature, (ii) identify common issues in MPM communication and their impacts, (iii) suggest ways to support the clear communication of MPM data and metadata, (iv) highlight advantages for authors and the scientific community at large and (v) introduce a standard method for sharing MPM data and metadata.

2 | MPM COMMUNICATION: CURRENT STATE OF AFFAIRS

To present the current practices in MPM data and metadata communication, with the ultimate goal to evaluate the need for standardised data and metadata reporting, we performed a survey of researchers and screened a subset of papers that have been used to generate MPMs stored in COMPADRE and COMADRE.

2.1 | A survey on matrix communication

We surveyed expert population ecologists, who we identified as having published peer-reviewed papers that include MPMs, regarding our current ability to communicate MPM data and metadata for reproducibility purposes. Specifically, we asked how well peer-reviewed publications relay the attributes of MPMs necessary for reproducibility. Additionally, we asked if researchers thought a standardised method of matrix communication is 'necessary for the coherent communication of MPMs in the literature' (the full list of 11 questions can be found in Supporting Information). The survey was distributed using Google Forms. We identified 1390 potential participants based on the criterion of being the lead and/or corresponding author from a publication containing at least one MPM. Over 50% of corresponding email addresses were outdated and not contacted further. Of the remaining approximately 650 researchers, that were contacted, 60 participants completed the survey. As expected, researchers report a great deal of heterogeneity in components of MPM communication (Figure 1). The best communicated attributes according to these survey participants are trait names (i.e. the phenotype by which the MPM was structured—stage/age/size classes), census duration and projection interval while the worst communicated attributes are life cycle graphs, formulae defining the vital rates and population vectors (i.e. number/frequency of individuals in each st/age). Importantly, 83% of survey participants agreed that the discipline needs a standardised method for MPM communication.

2.2 | A screen of papers in COMPADRE and COMADRE

To quantify how well MPM data and metadata are communicated in peer-reviewed publications, we screened 300 randomly

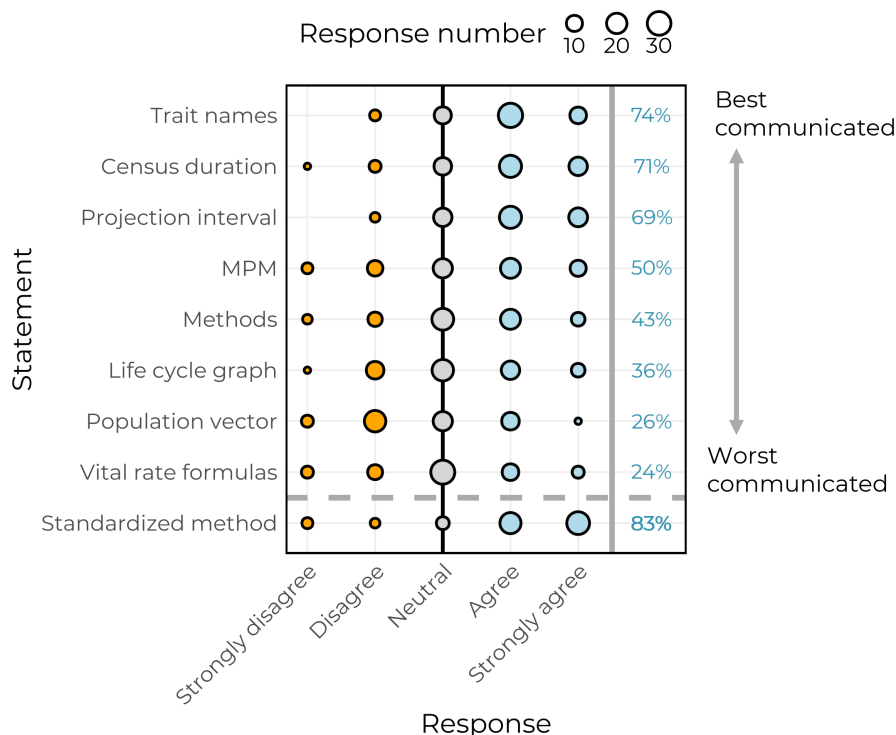


FIGURE 1 Survey results from experts in population ecology that participated ($n=60$). Participants ranked their confidence in the appropriate communication of components of matrix population models (MPMs) in peer-reviewed papers. Each component of MPM communication on the y-axis represents a statement shown in the survey (see SOM for the full survey). For all statements above the dashed line, participants were asked if that attribute (e.g. projection interval) is sufficiently well-reported in peer-reviewed publications. The statement 'standardised method' indicates participants' response to whether the field of population ecology would benefit from a standardised method of MPM reporting. The size of the dots indicates the number of respondents with that response and are coloured (i.e. orange = disagreement; grey = neutral; blue = agreement). For ease, percent agreement (i.e. the percentage of participants that either agreed or strongly agreed with the statement) is shown on the right-hand side of the plot.

sampled papers containing MPMs already digitised in COMPADRE and COMADRE (150 papers each). Across the different key attributes of MPMs that we examined, there was considerable variation in how reliably authors provided the data and metadata necessary for digitising, archiving, and performing comparative analysis (Figure 2). For instance, the generic location of the examined population (i.e. province/city/landmark; COMPADRE: 95.1%, COMADRE: 86.2% of papers reported it), the fully parameterized MPM (93.3%, 88.9%), and the census date (89.6%, 77.7%) were frequently explicitly stated in the papers, while latitude-longitude of the examined population (52.4%, 39.9%), its life-cycle diagram (44.5%, 40.1%), and population vector (i.e. st/age distribution of individuals at time t) (33.2%, 32.6%) were not. Interestingly, plant studies using MPMs (COMPADRE) contain overall more explicit data and metadata than animal studies (COMADRE; Figure 2). Furthermore, we used this information to categorise the quality of each of the examined 300 papers according to their reproducibility—defined as their inclusion of components of MPM communication (Figure 3). The distribution of component communication across kingdoms is similar. Crucially, only 13.9% of papers in COMADRE and 15.8% of papers in COMPADRE contain all the information necessary for comparative analyses and accurate projections (Figure 3). Thus, approximately 85% of papers require emailing authors to request undisclosed information.

3 | COMMON ISSUES IN MATRIX CONSTRUCTION

Here, we identify key issues in the parameterization of MPMs to illustrate the impact of methodology on demographic inference. To do so, we draw from the findings from the previous section and our experience curating COMPADRE and COMADRE. We outline the following issues for two reasons: (i) to advise demographers in how to identify them in the literature and (ii) to prevent these issues persisting in future publications. We note that a comprehensive list was recently made available by Kendall et al. (2019, see also Che-Castaldo et al., 2020). Here, we add to these previous papers by outlining steps for researchers to avoid/mitigate these issues in their own research. A summary of these issues, from occurrence to impact, is detailed in Figure S1.

3.1 | Census type, timing and frequency

MPMs are discrete-time demographic models parameterized by the tracking of individuals across censuses. Thus, the type (e.g. longitudinal, cross-sectional), timing and frequency of sampling needs to be carefully planned. These criteria are particularly important as census

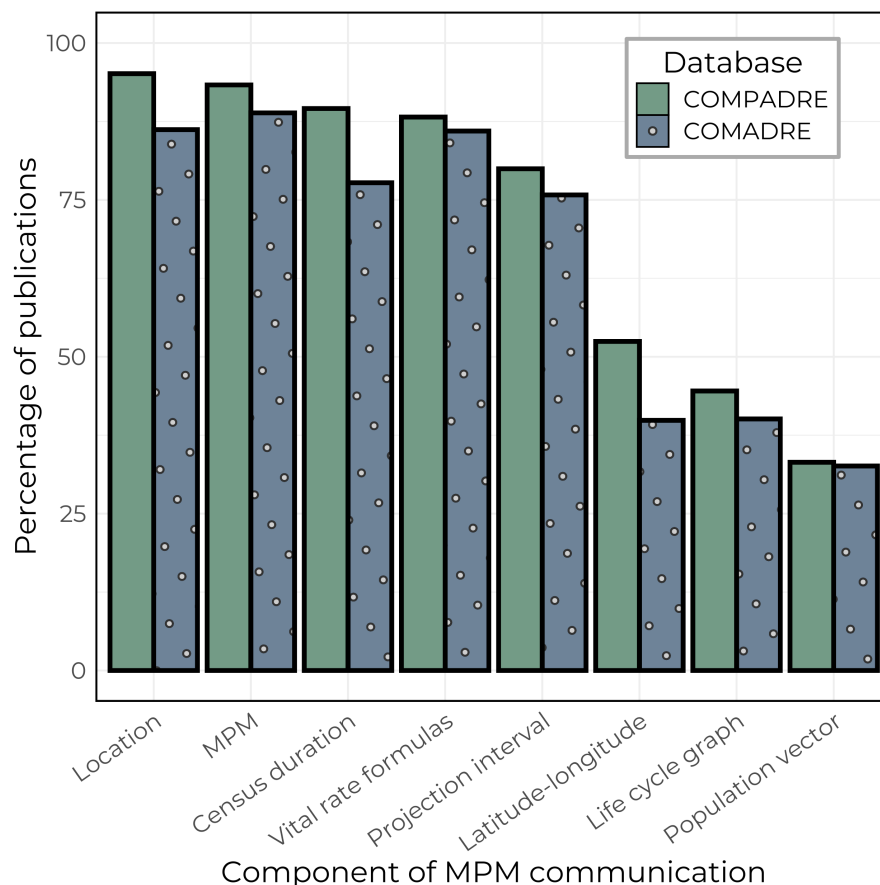


FIGURE 2 Both plant and animal MPM papers show similar patterns as to which presented components of MPM are communicated. The y-axis indicates the percentage of peer-reviewed publications in COMPADRE and COMADRE that contain a given attribute necessary for the clear communication of MPM information and its reproducibility from a random subset of 150 papers of the 643 total papers from COMPADRE and 150 out of the 395 total papers from COMADRE (300 papers total). The attributes are: Location: province/city/landmark; MPM: was the MPM included in the manuscript; Census duration: start and end dates for data collection; Vital rate formulas: decomposition of matrix elements into their underlying components (i.e. contributions from survival, growth, and reproduction); Projection interval: the time period between observations; Latitude-longitude: spatial coordinates; Life cycle graph: the visual representation of demographic transitions and a /sexual per-capita contributions; Population vector: st /age distribution of individuals at time t associated with reported MPMs.

type directly affects matrix construction, and census timing and frequency can inadvertently influence demographic outputs (Emery & Gross, 2005).

Typically, an MPM comes in two forms regarding the spread of reproduction between censuses: birth-flow or birth-pulse (see Caswell, 2001, p. 22). The distinction is based on whether reproduction occurs continuously (i.e. birth-flow) or in a narrow temporal window (i.e. birth-pulse). Birth-pulse MPMs are further categorised into pre- versus post-reproductive census. Although both pre- and post-reproductive censuses often lead to similar demographic inference (see Cooch et al., 2003), their difference lies in when populations are censused relative to the position of the narrow reproductive window. In the former, populations are censused immediately before a reproductive window, while post-reproductive censuses follow on from a reproductive window. A pre-reproductive census requires the inclusion of offspring survival in reproductive matrix elements, while a post-reproductive census requires the inclusion of parent survival in reproductive

matrix elements. We often encounter mistakes in the accommodation of offspring or parent survival in reproductive matrix elements (see also Kendall et al., 2019). A key step in matrix construction that can prevent the incorrect accommodation of survival is drawing the life cycle graph (as per Ebert, 1999, p. 61) with respect to census timing (demonstrated in Ellner et al., 2016, p. 13), as well as explicitly detailing the census type used to parameterize the MPM. However, sometimes drawing the life cycle graph may be unfeasible or uninformative. For example, the graph for an age classified model with 100 age classes is too large to draw and too redundant to be useful; but, they can be simplified with a dashed line if multiple adjacent classes have the same demographic rates (e.g. Ebert, 1999, p. 2). Models with many stages and highly connected transitions are not feasible to draw the life cycle graph (e.g. the graph for *Calathea ovandensis* in Neubert and Caswell (2000)). But even in complex situations (e.g. the series of seasonal graphs for the emperor penguin in Jenouvrier et al. (2010)) the graph may be helpful in organising the structure of the model.

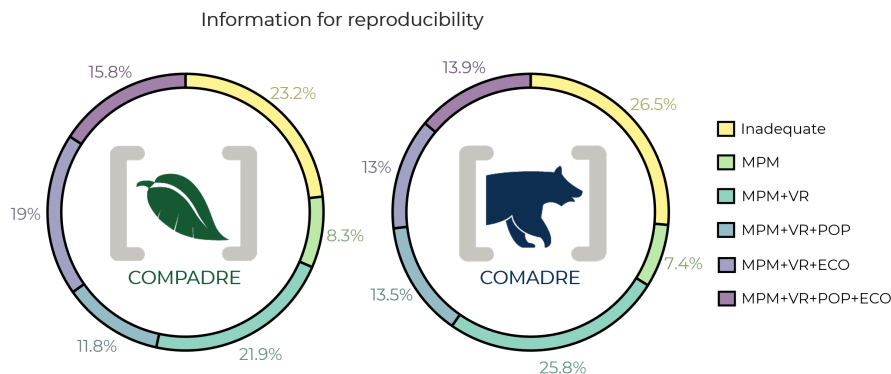


FIGURE 3 Across plant and animal MPM papers, most publications do not contain sufficient information for reproducibility. Proportion of papers in COMPADRE and COMADRE grouped by their open-access information in peer-review publications regarding matrix population model (MPM) data and metadata. Following the same scheme as in Figure 2, papers were ranked into six groups from 'inadequate' to 'MPM+VR+POP+ECO' (i.e. fully reproducible). 'Inadequate' refers to papers missing the MPM and/or projection interval (i.e. an MPM specific time interval necessary for projection), without which most demographic outputs cannot be calculated. 'MPM': paper contains the MPM and projection interval but no vital rate formulas describing the matrix elements. 'MPM+VR': contains all of the information for 'MPM' along with vital rate formulas for the matrix elements. 'MPM+VR+POP': contains all of the information for 'MPM+VR' along with the population vector. 'MPM+VR+ECO': contains all of the information for 'MPM+VR' along with latitude-longitude coordinates and census duration of the examined population. 'MPM+VR+POP+ECO': contains all of the information for 'MPM+VR' along with population vector/distribution, latitude-longitude coordinates and census duration.

Census timing and frequency affects model construction, making a constructed MPM impractical for demographic inference if the life history of the examined organism is not considered. Consider a researcher comparing the demographic processes of fruit flies and fruit trees. The researcher first notices that there are four discrete stages to the fruit flies' life history: three juvenile stages encompassing the development from egg to instar to pupae, and one adult stage where individuals disperse and reproduce. Since development from egg to adult takes ~10 days in this species, the researcher decides to perform the census every 10 days for both the fruit fly and the fruit trees over a 3-month period. However, because neither mortality nor reproduction occur across such a short census in the fruit tree population, the resulting fruit tree MPM, when projected forward, will persist forever, neither increasing nor declining. This same issue would occur the other way around. If 5-year intervals were deemed sufficient for the fruit trees, then individually measured fruit flies would never survive across time steps. A solution to this problem exists, using periodic matrix models to include periods much shorter or longer than other periods. For example, Hunter and Caswell (2005) analysed the Sooty Shearwater *Puffinus griseus* including two harvesting periods of several weeks in duration and then an annual interval for the species, with a lifespan of decades. Smith et al. (2005) and Shyu et al. (2013) used periodic seasonal models to accommodate life cycles in which some stages are only present for part of the annual cycle. The approach (Caswell, 2001, section 13.1) is powerful and general.

3.2 | Unrealistic stage-specific survival

Issues in parameterising stage-specific survival, whilst easy to diagnose, can result in an array of unnatural life histories. Transition and

survival probabilities are bounded between 0 (i.e. the event never happens) and 1 (i.e. always occurs). As such, the stage-specific survival of an MPM, the summed nonreproductive elements in a given column of the MPM \mathbf{A} must not exceed 1. When it does, individuals in that stage have an unrealistic chance of surviving >100%, resulting in an incorrect representation of the organism's life history. Stage-specific survival values >1 typically arise due to rounding errors, typos, inclusion of unstated a/sexual reproductive events. As such, it is generally advised to omit these MPMs in comparative analysis (Jones et al., 2014). Unstated a/sexual reproductive events occur when a given element a_{ij} in the MPM \mathbf{A} contains both survival-dependent processes, such as growth/shrinkage, but also fertility, and these have not been reported separately. Ideally, authors would carefully identify whether various vital rates are being confounded with survival-dependent demographic processes in each MPM element. For the comparative demographer using COMPADRE and COMADRE, we recommend either avoiding MPM models where stage-specific survival >1 or altering the model so that the stage-specific survival is fixed to a maximum of 1 (e.g. Buckley et al., 2010).

In many published MPMs, some life stages have an estimated survival probability of 1 or an incomplete life cycle, likely the result of small sample size or rare event along the life history of the species. Perfect survival (i.e. mortality=0) is unlikely to be accurate, and may need to be estimated or imputed (Johnson et al., 2018). A reproducible approach to infer realistic survival and transition values was recently proposed by Tremblay et al. (2021), using a Bayesian approach to estimate parameter values using priors in addition to the observed data to obtain posterior MPMs. An advantage of this approach is that the confidence intervals of the parameters that represent probabilities (i.e. stasis, transition, survival) are obtained from a beta distribution. This advantage of using a Bayesian inferred multinomial Dirichlet distribution for estimating the mean values is

that the researchers can infer variance and skew of the posterior distributions to further inform MPM construction and demographic inference (e.g. Tremblay et al., 2009a, 2009b). And finally, since sample size can be a key driver of unrealistic stage-specific survival, sample size and uncertainty (e.g. confidence interval, standard deviation) must be reported to (1) relay the precision of the estimated survival value to your audience and (2) for accurate inclusion of survival values in meta-analyses and comparative methods.

3.3 | Incorrectly parameterizing fertility

Fertility often presents a challenge to constructing accurate MPMs. This challenge is partly due to the ambiguity of the term 'fertility'. The issue arises when the per-capita contributions of reproductive adults to new recruits (e.g. eggs, neonates, seeds, etc.) do not represent the links over the full projection interval of the study. Remember that the entry a_{ij} in an MPM is the (expected) number of stage i individuals at $t + 1$ per stage j individual at time t . If stage i is some kind of 'newborn' individual, then a_{ij} must include all the processes between time t and time $t + 1$ (Caswell, 2001, p. 61). Reproductive output, in turn, is a composite demographic process of the number of offspring produced in a reproductive event and the relevant survival that will penalise how many new offspring will *actually* make it to the next observation. Failure to accommodate this vital rate decomposition can result in the introduction of a one-timestep lag into the organism's life cycle, as newly created offspring spend a projection interval 'in limbo' before their onward transitions. The best-known example is in the classic model of teasel *Dipsacus sylvestris* by Werner and Caswell (1977), in which flowering plants at time t were described as producing seeds at time $t + 1$, which only germinated to seedlings at time $t + 2$. The issue was discussed and corrected in Caswell (2001). Furthermore, this issue has been reported, for instance, in reproductive structures such as seeds that do not actually undergo a permanent seed bank. An MPM with this issue will typically (Kendall et al., 2019), but not always (Nguyen et al., 2019), underestimate the asymptotic population growth rate, λ . Naturally the challenge will then be in estimating the relative importance of the seed bank and the lifespan on nongerminated seeds. The effect of incorrectly parameterizing fertility on λ is greatest in cases of extreme growth, such as invasive species, or extreme decline, such as critically endangered species (Rueda-Cediel et al., 2018). Furthermore, this issue can also cause overestimation of the transient envelope (see Ezard et al., 2010). Thus, we recommend reporting the fertility vital rate formulas with the associated MPMs and clearly identifying the values of these underlying vital rates (as in Box 1).

3.4 | Indirectly calculating vital rates

Estimating vital rates often involves combinations of direct and indirect measurements. Direct measurement empirically derives vital rates from individual-based data where identified individuals are censused

multiple times, as in cohort life table studies, mark-recapture methods and many quadrat studies of marked plants. However, vital rates can be hard to observe in species with high offspring production, complex phenology and/or small population sizes (Beissinger & Westphal, 1998). Consequently, recruitment estimates are often supplemented into MPMs from controlled conditions; examples include the laboratory (Jouvet et al., 2018), greenhouse (Gontijo & Carvalho, 2020), zoo (Clubb et al., 2009) and botanic garden (Jiménez-Valdés et al., 2010). Since some MPM methods require a full life cycle to obtain key metrics (e.g. transient metrics: Stott et al., 2011), external study sites or literature sources are often used to parameterize components of the MPM to 'close the loop' in incomplete life cycles (Omeyer et al., 2021). However, captive populations may not represent wild population dynamics (Clubb & Mason, 2003), particularly in regards to survival (Che-Castaldo et al., 2021) or reproduction (Clubb et al., 2009).

Another method to indirectly estimate vital rates involves using ex-situ methods to obtain upper and lower bounds on recruitment (or other vital rates) and explore the parameter space within those bounds. The approach was introduced by Caswell et al. (1998) in a study of the effects of bycatch mortality on the harbour porpoise. Age-specific survival and fertility schedules were selected from other species with similar life cycles, re-scaled to match the longevity of the harbour porpoise, and used to produce uncertainty distributions for population growth and the effects of the measured bycatch. Reporting the distribution and associated parameters provide a measure of uncertainty from which to inform the construction of an MPM (Tenhumberg et al., 2008). Furthermore, the use of hierarchical models to estimate missing values and borrowing strength from other populations or species may improve parameter estimation (James et al., 2021; Tremblay & McCarthy, 2014).

And lastly, integrated population models represent a valuable framework for indirectly estimating the demographic rates and population dynamics (size and structure) by combining data sources, particularly combining longitudinal individual data with population census data (Plard et al., 2019; Schaub & Kéry, 2021). Integrated population models allow for the construction of population models (including MPMs) by (1) combining data sources, (2) defining a life history a priori (this is often some form of stage-structured population model) and (3) quantifying the maximum likelihood of demographic rates encoded in the life history given the data sources. Integrated population models are particularly useful when uncertainty around data acquisition is known (e.g. in capture-mark-recapture studies) (Riecke et al., 2019).

3.5 | Population vector

An estimate of the structure of the population, classified by age or stage, is a useful piece of information when available, but it will only sometimes be available. Current population structure provides a logical starting point for projections of short-term and long-term population viability (Werner & Peacock, 2019). Furthermore, using the population vector (i.e. abundance and stage distribution) for projections helps to account for the effects of transient dynamics, which measure the

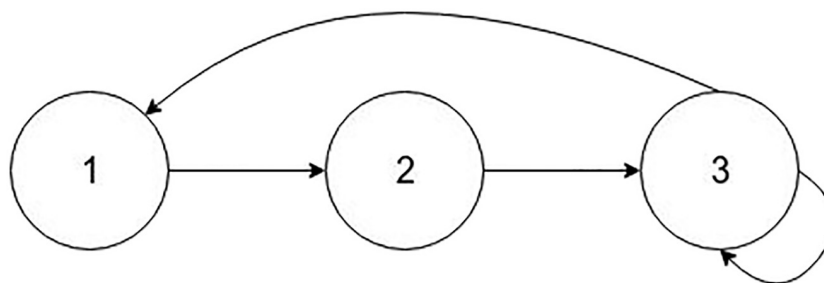
BOX 1 Example presentation of a hypothetical three-stage plant matrix population model (MPM) using a clear and explicit presentation of data applicable to most MPM construction techniques.

(A) Matrix type

A simple deterministic density-independent matrix.*

*This free text field allows for the brief description of matrix type. If the matrix is structured by one variable the matrix is simple. If not, the matrix is considered general (e.g. age x stage). Deterministic refers to if the demographic rates that build the MPM are held constant (deterministic) or drawn from a distribution (stochastic). Density-independent versus density-dependent indicates if the demographic rates are or are not influenced by population density.

(B) Life cycle diagram



(C) Census description

Census attribute	Example
Census duration	May 2021 to May 2022
Location (latitude-longitude)	Sheffield, UK (53°24'41.5"N 1°30'02.3"W)
Projection interval	1 year
Mode of reproduction	Birth-pulse (taking place in June and July)
Census type	Pre-reproductive census

(D) Stage names and classification

Stage number	Stage name	Classifying criteria
1	Seedling	A recently germinated individual that has fewer than four leaves and has not developed the radial rosette structure. The leaves are less than three centimetres in length
2	Rosette	An individual with a pronounced radial morphology in leaf structure. The leaves are between three and six centimetres in length
3	Adult plant	An individual with a pronounced radial morphology and has the ability to produce flowers. The leaves are greater than six centimetres in length

(E) Vital rate definitions

Vital rate	Definition	Data source
S_{ij}	Probability of survival from class j to class i	Field site
N_s	Number of seeds per fruit	Field site
N_{fx}	Number of fruits produced by an individual in size category x	Field site
P_{rx}	Probability of reproduction of an adult in category x	Field site
P_s	Probability of seed germination and seedling survival across the projection interval	Greenhouse

BOX 1 (Continues)

(F) Vital rate values

Vital rate	Estimate*	Standard error**	Sample size (individuals)***
S_{21}	0.400	0.100	80
S_{32}	0.850	0.050	160
S_{33}	0.900	0.020	160
N_s	1000	150	80
N_{f3}	2.000	300	80
P_{r3}	0.300	0.040	200
P_s	0.005	0.001	500

*If these estimates are dependent on population size (i.e. density-dependence) or in response to an environment variable (i.e. environmental stochasticity), the estimate should be communicated as a function (e.g. $S_{33} \sim 0.9 + \beta_{\text{precipitation}} \times 0.01$ where $\beta_{\text{precipitation}} \sim N(5, 1)$). Furthermore, if this is a point estimate, researchers should indicate if the values represent mean or median values.

**This measure of uncertainty may also be the estimate's standard deviation, variance or a confidence/credible interval at the discretion of the researcher.

***Indicate if the unit of measurement/replication is at the level of individual organisms or at the level of groups (e.g. cohorts, colonies, families).

(G) MPM formula (May 2021 to May 2022)

0	0	$N_s \times N_{f3} \times P_{r3} \times P_s$
S_{21}	0	0
0	S_{32}	S_{33}

(H) Population vector (May 2021)

Stage	Number	95% Confidence interval*
1	1350	1150–1550
2	550	530–570
3	300	290–310

*This measure of uncertainty may also be the estimate's standard deviation, variance or a confidence/credible interval at the discretion of the researcher.

effect of nonstationary population structure on near-term population growth rates (Capdevila et al., 2020). Reported population stage or age vectors reflect two key components: the actual population structure at the census in time t and the methodological choices. This second component is critical to accurately represent the studied population.

Across the development and curation of COMPADRE and COMADRE, we have noticed two sources of error that affect population vector estimation. The first error is a detection bias, where researchers identify certain st/ages with a higher rate of detection over more cryptic stages (e.g. adult plants versus seed banks). The second error is the misappropriation of methods used to quantify demographic rates as a basis for estimating stage abundance. This second error stems from a misunderstanding of the difference between estimating rates and the estimation of numbers. To measure demographic rates, researchers sometimes increase the sampling effort of certain st/age classes over other st/age classes during a census. This differential sampling effort across st/ages is particularly common when st/age classes

have survival probabilities close to its limits (i.e. 0 and 1). For instance, in tree demography, there are typically only a few very large individuals per area examined. Thus, oftentimes researchers supplement the sample size of this category by sampling outside of the predefined area (Jones & Hubbell, 2006). However, many types of demographic rate estimation do not provide any information on numbers and structure. Cohort life tables, that follow a cohort of individuals as they age, are blind to the structure of the population in which the cohort develops. Indeed, there may be no such population (e.g. the entire history of laboratory cohort-based demography going back to Pearl in the 1920s (Pearl et al., 1927)). Mark-recapture estimation of rates from longitudinal data draws all its inference from the marked individuals and makes no inferences about the number and structure of the unmarked. The literature on mark-recapture methods for estimating rates recognises that estimating population numbers is thus much more difficult than estimating rates (Lebreton et al., 1992) and requires different mark-recapture models (e.g. the inclusion of trap happiness/shyness in vital

rate estimation). Therefore, if projection from an actual structure is desired, that initial condition may be more appropriately measured in a separate census, rather than extracted from the measurements of rates that inform the MPM.

3.6 | Omitting cryptic life stages

The identification and estimation of vital rates in cryptic stages poses a challenge in population ecology. Cryptic stages represent points along an organism's life cycle that are somewhat hidden from or overlooked by population ecologists when building population models (Doak et al., 2002). A life stage could be cryptic because it is logistically challenging to observe or observable but indistinguishable from a similar seeming class (Nguyen et al., 2019). In plants, cryptic stages can emerge from seed banks for plants, such as orchids, where the seeds are too small to be identified in the field (Paniw et al., 2017) or some herbaceous perennials (e.g. *Astragalus scaphoides*) where prolonged periods of vegetative dormancy can allow individuals to stay underground for one or more growing seasons (Gremer & Sala, 2013). Additionally, animals can exhibit cryptic stages by undergoing diapause or delays in development due to adverse environment conditions (*Aedes albopictus*: Jia et al., 2016). Pelagic seabirds (albatrosses, petrels, penguins) often spend pre-reproductive stages, sometimes of many years durations, at sea, completely cryptic until returning to the breeding colony as adults. Sophisticated multistate mark-recapture methods can provide estimates of parameters for these parts of the life cycle (e.g. Jenouvrier et al. (2018), using the multievent algorithm of Choquet et al. (2009)). Omitting a cryptic life stage can reduce the biological realism of an MPM and alter the number of stages in the MPM, which can further impact demographic outputs (Salguero-Gómez & Plotkin, 2010; Tenhumberg et al., 2009). In some cases, cryptic life stages will only be identified via a multidisciplinary approach including field and laboratory methods, coupled with Bayesian frameworks to integrate data and prior knowledge (e.g. Paniw et al., 2017).

3.7 | One-sex versus two-sex models

Much of demography focuses on females, under the assumptions that fertility is determined by females without limitation by males (see Caswell, 2001, p. 568). Such models may include males (e.g. Hunter et al., 2010), but if reproduction is determined by female rates (i.e. are female-dominant), males represent a set of stages that do not contribute to population growth. Most existing animal MPMs are female-based and female-dominant (Salguero-Gómez et al., 2016) because, given sampling biases towards mammals and birds (Conde et al., 2019), it is oftentimes not feasible, or necessary for the research question, to track male reproductive interactions (Archer et al., 2022). These studies typically assume a 1:1 sex ratio, sex-congruent vital rates and that reproduction is not male-limited (Compagnoni et al., 2017; Miller & Compagnoni, 2022). While one-sex models are common in animal MPMs (currently 77% in COMADRE v. 4.21.8), care must be taken

to not make assumptions about sex-ratio dependent dynamics within these systems (Archer et al., 2022). Indeed, these assumptions may not be met when any of the following are true: there is a bias in sex ratio (Archer et al., 2022), there is reproductive skew (Sky et al., 2022), or a high sensitivity of population dynamics in response to mating choice (Veran & Beissinger, 2009). Furthermore, sex-dependent detectability can further confound estimates of sex-ratio and their associated impacts on vital rates if not taken into account. Two-sex models that do not assume dominance by one sex are nonlinear and require specification of a mating function that describes fertility as a function of male and female abundance (Caswell, 2001). Defining such mating functions is generally difficult or impossible, except in the particularly easy case of strict monogamy (Jenouvrier et al., 2010).

Reporting sex ratios can greatly expand the scope of a study (Shyu & Caswell, 2016a, 2016b); for example evaluating the impact of sex ratio and the Allee effect (Boukal & Berec, 2002). Unfortunately, this reporting is rarely done in work archived in COMPADRE and COMADRE. Moreover, if there are differences in vital rate values between sexes, such as survival, growth, and/or reproductive output, a one-sex MPM may neglect important processes (Archer et al., 2022; Caswell, 2001, p. 568). In plants, reporting two-sex dynamics is even more rare (0.2% in COMPADRE v. 6.22.5). However, this low percentage likely reflects the rarity of dioecy or other mating systems with two or more sexes in plants (Käfer et al., 2017) and the commonness of polygamous mating systems which makes male-limited reproduction rare (see Compagnoni et al., 2017; Miller & Compagnoni, 2022).

3.8 | Irreducibility and ergodicity

The property of irreducibility has implications for the eigenvalue spectrum of a matrix, and hence biologically relevant outputs (e.g. population growth rates, stable stage structures). These implications are well known in the literature on MPMs (Caswell, 2001). An irreducible matrix is one in which the life cycle graph is completely connected; that is, there exists a direct or indirect path from any stage to any other stage. It is sometimes asserted that reducible MPMs are somehow invalid; they are not. There are (at least) four situations in which reducible matrices naturally occur.

1. Life cycles with post-reproductive stages. The post-reproductive stages can make no contribution to the potentially reproductive stages (e.g. MPMs for humans, orcas).
2. Two-sex models with dominance by one sex (usually females, but could be male). In a female-dominant model, all reproduction is credited to females. Males are produced by females, but make no contribution to the female part of the life cycle (e.g. Hunter et al. (2010) for polar bears).
3. Spatial models in which dispersal is one-directional, as in river systems or oceanic currents.
4. Age \times stage-classified models (Caswell, 2009; Caswell & Salguero-Gómez, 2013). In these models, reproduction produces (by definition) individuals in age class 1, but the model includes all

combinations of age and stage, including impossible combinations of age class 1 and stages that do not exist at age 1.

Reducibility may or may not be easy to spot from the life cycle graph, but it can be tested numerically. The matrix \mathbf{A} is irreducible if and only if the matrix $(\mathbf{I} + \mathbf{A})^{s-1}$ is positive (Caswell, 2001). Irreducibility, together with primitivity, is a sufficient condition for ergodicity, guaranteeing that the population will converge to the same stable structure regardless of the initial condition. A reducible matrix may not have this property; clearly, for example, a population started with only post-reproductive individuals will not converge to the same structure as one started with some pre-reproductive individuals. With regard to ergodicity, it is also known that an MPM is ergodic if and only if all entries of its dominant left eigenvector (\mathbf{v}) are positive (Stott et al., 2010). In short, despite appropriate model structure and correct parameterisation, demographic data may lead to reducible and/or nonergodic matrices.

4 | FULL REPRESENTATION OF AN MPM IN PUBLICATIONS

In this section, we justify the need for clear presentation of MPMs in the scientific literature, suggest where to archive MPMs open-access, and discuss how these two actions benefit the original authors, publishing journal, readers, comparative demographers, meta-analysts and the discipline at large. A table containing corresponding data to include when publishing MPMs along with a rationale for inclusion and examples of good practice can be found in Table S1.

4.1 | Partitioning demographic processes

It is important to define what each matrix element in an MPM represents. Various demographic processes can overlap into the same matrix element in an MPM, particularly in species with a fast

and/or plastic lifecycle relative to the MPM projection interval. For example, the value in an MPM that represents the link between large individuals at time t and smaller individuals at time $t + 1$ might correspond to sexual reproduction, clonal reproduction, fission, retrogression, or a composite of multiple processes. The mathematical derivations of key life history traits (e.g. generation time, life expectancy, rate of senescence, degree of iteroparity) require that these processes be clearly separated (Jones et al., 2022). This is critical for the family of analyses based on Markov chains; the matrix \mathbf{U} defines the transient state transitions in an absorbing Markov chain (Caswell, 2011, 2013). By reporting the underlying demographic rate structure in a life cycle diagram and its consequent full matrix population model \mathbf{A} , one can separate matrices into survival-dependent processes (e.g. progression/growth, retrogression/shrinkage, fission, fusion, stasis) in the submatrix \mathbf{U} , sexual reproduction in the submatrix \mathbf{F} and clonal reproduction in the submatrix \mathbf{C} (Figure 4). Importantly, both \mathbf{F} and \mathbf{C} submatrix elements must incorporate survival according to census type (i.e. pre-/post-reproductive census).

Reporting the matrix \mathbf{A} as well as the submatrices \mathbf{U} , \mathbf{F} and \mathbf{C} lends two key benefits: (1) explicitly indicates how the values in \mathbf{A} are generated from underlying vital rates; and (2) the submatrices can be used to calculate a vast plethora of demographic measures that cannot be calculated from \mathbf{A} alone, such as longevity (mean and variance), occupancy times (means and variances), lifetime reproductive output (means and variances), net reproductive rate, generation time and entropy (Keyfitz entropy (Keyfitz, 1968) and Demetrius' entropy (Demetrius, 1992)) just to name a few.

4.2 | Attribution of secondary data sources

Secondary data sources are critical for reproducibility. These data sources provide information and support for methodologies used in MPM construction. In some cases, MPMs simply use secondary data to complete the life cycle, whereas others are constructed purely from secondary sources (see Table S1). Secondary sources include

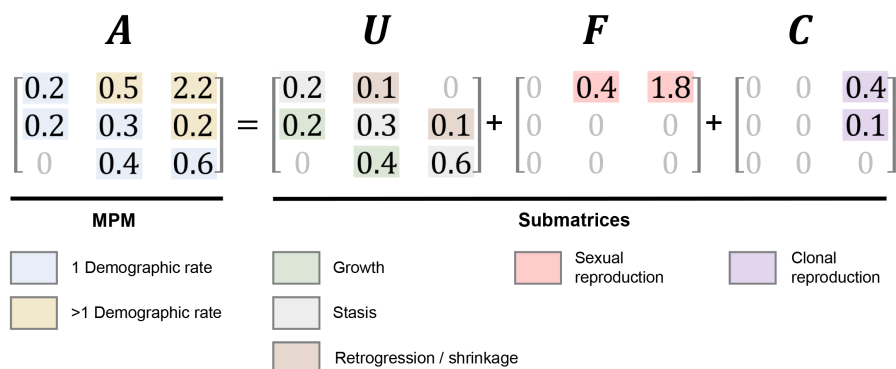


FIGURE 4 Decomposing an MPM into its submatrices allows for the isolation for otherwise masked vital rates. Matrix \mathbf{A} represents the MPM. Since individual transitions can be represented by multiple demographic rates (e.g. retrogression, sexual reproduction and clonal reproduction), decomposing \mathbf{A} into its \mathbf{U} , \mathbf{F} and \mathbf{C} submatrices allows for targeted demographic inference about what demographic transitions are driving the dynamics of the population.

previous studies, data from databases, simulations, indirect observations and theoretical estimates. The use of secondary data sources may mean that the final MPM does not accurately represent vital rate trade-offs, and so should be recognised in the methods section of the publication or its Supporting Information. Sufficient communication of secondary sources includes the source of the data, whether a point estimate, confidence interval, or distribution was integrated with the primary data source as well as the rationale for their inclusion. For example, Omeyer et al. (2021) presents a table of data sources used in construction of the MPM. If these secondary sources are not recognised in tandem with the MPM, the inferred demographic processes, however realistic they may be, may not pass peer review nor uptake by the scientific community. In turn, clear communication of these secondary sources is highly recommended.

4.3 | Archival of information in COMPADRE and COMADRE

We propose that the COMPADRE and COMADRE matrix databases provide the most appropriate way of archiving and accessing MPMs. While we recognise that there are other ecological database repositories (e.g. dryad: <https://datadryad.org/>; figshare: <https://figshare.com>; zenodo: <https://zenodo.org>), the open access COMPADRE and COMADRE databases (<https://compadre-db.org/Contribute>) provide a dedicated data archival platform, specifically for MPMs, allowing direct contributions from researchers as well as digitization of published MPMs by our data validation teams. The web-based data entry portal provides a structured data curation process (i.e. from screening, to standardisation, to validation) that can accommodate MPMs of different dimensions and for diverse life histories. On entry, MPMs are complemented with relevant biogeographic variables and details on census methodology in COMPADRE and COMADRE. Details of the original publication, including DOI and citation functionality (see <https://compadre-db.org/Education/article/obtaining-references>), are stored alongside each MPM to ensure that their contribution towards any future publication is recognised. All data are archived long-term through Oxford Open Access and Bodleian Library support.

Other recent enhancements to COMPADRE and COMADRE will further aid the research community. Previously, the databases were only accessible via download of an R-object file which contained all matrices in that version of the database. The database is now accessible via a queryable website (<https://compadre-db.org/QueryDatabase>) that allows users to find and download individual matrices. We also strive to empower researchers and educators with teaching materials (<https://compadre-db.org/Education>) and the production of new R-packages (Jones et al., 2022) for ease and scalability of MPM-related research. Along with these materials, all details of the database structure and workflow are open-access (<https://jonesor.github.io/CompadreGuides/user-guide.html>). These improvements to the databases and their interface structure have been directly targeted to equip demographers with more tools to

conduct research with and train students on MPMs along with increasing database transparency to ensure best research practices.

5 | A STANDARD PROTOCOL FOR REPORTING MPMs

Here, we introduce a proposed checklist for how to report an MPM in publications (Box 1). We recommend using the checklist when designing data collection as well as when writing up the MPM for publication. We recommend using this template as Supporting Information for published MPMs as it allows for the clear communication of model construction in addition to ease in integrating published MPMs into the COMPADRE and COMADRE databases.

6 | THE THEORY DOES NOT STAND STILL: NONLINEARITY, ENVIRONMENT-DEPENDENCE AND MULTISTATE MODELS

MPMs have become a predominant approach in the toolbox of population ecologists partly due to their simplicity of construction and analysis. But the theory underlying matrix-based demography does not stand still, and in the last 20 years it has enlarged dramatically. These new methods produce models whose structure does not fit into the frameworks for reporting that seemed so comprehensive in the past. These recent advances in MPM theory and methods, enable researchers to link population dynamics and demography to environmental conditions and multiple individual traits (e.g. sex and age (Childs et al., 2016); age and kinship (Caswell, 2019b, 2020)) rather than a single trait. These advances also offer benefits for the study of population responses to extreme climate (Jenouvrier et al., 2022), as well as more nuanced investigations of comparative and evolutionary demography (Childs et al., 2016). In turn, in this section, we overview some exciting areas of structured demography that can open novel research questions for the modern demographer and list some of the challenges they pose for communication and reporting.

6.1 | Nonlinear dynamics

Nonlinear MPMs are those in which entries of the projection matrix depend on the population state (numbers and structure) and may be frequency- or density-dependent. Frequency-dependent nonlinearities depend only on the relative abundance of stages; they occur in two-sex models in which mating depends on the relative abundance of males and females, and in population genetic models where dynamics depend on the relative abundance of genotypes (de Vries & Caswell, 2019). Density-dependent models depend on the abundance and structure of the population; recent examples include Pardini et al. (2009) and Shyu et al. (2013) for analyses of control strategies for garlic mustard *Alliaria petiolata* and de Vries et al. (2020) for laboratory studies of pesticide resistance in *Tribolium*.

The analysis of nonlinear MPMs focuses on demographic outcomes different from those of linear models; equilibria, attractors, bifurcations, oscillations and stability (see Caswell, 2001, Chapters 16 and 17, and Cushing et al., 2003 for the most detailed analysis yet). However, what makes these models problematic for the current status of COMPADRE and COMADRE is that the unit of the model is not a matrix, but rather a matrix *function*, in which the entries of the projection matrix are functions of the state of the population. Sensitivity analyses are available to study pretty much any demographic outcome in response to any parameter (Caswell, 2019a), but reporting the functions that define the MPM is not at all standardised.

6.2 | Environment-dependence

A similar problem arises in environment-dependent MPMs. In such models, some or all of the demographic rates are functions of some aspects of the environment; for example, polar bears as functions of statistics of Arctic sea ice (Regehr et al., 2010), sifaka as functions of rainfall (Lawler et al., 2009), the emperor penguin as a function of seasonal sea ice patterns in the Antarctic (Jenouvrier et al., 2012) and the North Atlantic right whale as functions of time and of trends in time (Fujiwara & Caswell, 2001). As with nonlinear MPMs, the model is not a matrix, but a function that maps from the environmental variable(s) to the entries in the matrix. Protocols for reporting such functions are not yet available but are important to develop.

6.3 | Multistate models

An exciting emerging area of demographic research is the construction and analysis of multistate MPMs, in which individuals are classified by more than one state variable. This includes age and stage (Caswell & Salguero-Gómez, 2013), stage and spatial location (Hunter & Caswell, 2005), stage and genotype (de Vries & Caswell, 2019), stage and infection status (Klepac & Caswell, 2011), age and unmeasured heterogeneity (Hartemink et al., 2017), and stage-specific incidence of disease (Caswell & Van Daalen, 2021). A detailed presentation of the methods is given in (Caswell et al., 2018) and the extension to more than two state axes (so-called hyperstate matrices) is given in (Roth & Caswell, 2016). The incorporation of additional states enables researchers to tease apart various sources of individual heterogeneity, the variance of life history outcomes for individuals from the same population model, and to ask deeper comparative and evolutionary questions. For example, maternal age has a strong impact on vital rates in monogonont rotifers (Bock et al., 2019). Applying vec-permutation methods (Caswell, 2012) to build multistate MPMs has allowed researchers to quantify the population-level impacts of the observed maternal age effect and to investigate the evolutionary processes that can lead to this type of senescence in rotifers (Hernández et al., 2020). Multidimensional MPMs and Markov chain approaches have been particularly important in the

study of 'luck' in life histories, which explores why some individuals live long and prosper, while others do not (Snyder & Ellner, 2018). In studies of 'luck', variance among individuals for a life history outcome is partitioned into contributions from between-group and within-group variation (e.g. Snyder & Ellner, 2018; van Daalen & Caswell, 2017). Examples of sources of individual heterogeneity include maternal age (van Daalen et al., 2022), birth-year environment (Snyder & Ellner, 2022), and genetic variation (Steiner et al., 2021). The within-group variation is called individual stochasticity or 'luck' and arises from the fact that vital rates are probabilistic processes.

These models pose a challenge for reporting because the MPM consists not of a single matrix, but of four sets of matrices. Consider an age \times stage-classified model. It is composed of a set of matrices giving transitions among stages for each age class, a set of matrices \mathbf{D} giving age transitions for each stage, a set of matrices \mathbf{F} giving stage-specific fertility for each age, and a set of matrices \mathbf{H} that allocate newborn offspring to the appropriate ages. These are assembled into block structured transition and fertility matrices from which all the usual demographic outcomes can be calculated and related to both age and stage (e.g. see Caswell and Salguero-Gómez (2013) for an analysis of selection gradients for both age and size).

7 | DISCUSSION

Demographic research has come a long way since the introduction of age-based (Leslie, 1945) and stage-based matrix models (Lefkovich, 1965). Advances in this field have been fuelled partly by clear communication of methods and associated code. We aim to continue this expansion with MPM communication.

As the depth and breadth of the literature continues to expand, we are starting to build a comprehensive picture of demography across the spectrum of life (Adler et al., 2014; Healy et al., 2019; Salguero-Gómez et al., 2017). Through the work of the COMPADRE and COMADRE databases, we have come to appreciate the utility and opportunities of a standardised way of compiling MPMs. Indeed, a significant portion of the time (>50%) we spend curating these databases is actually not on digitising, error-checking, and complementing data, but on contacting authors for clarification and request of missing data and metadata. Through this arduous process, we have identified valuable—yet typically missing—information in MPMs. Whilst the missing data highlighted here as being particularly important primarily reflects the interests and perspectives of comparative demographers, including the data outlined in the standardised method would benefit demography as a whole.

This paper intends to act as a useful reference for authors, editors, reviewers, managers/conservationists and comparative demographers. Furthermore, we hope this manuscript will promote a constructive discussion on the purpose, construction and presentation of stage-based demographic information. **Box 1** contains a comprehensive example of the key information we believe should be incorporated into the publication of any MPM. Should the methods suggested here be adopted, there will be clear benefits for the

growth of the COMPADRE and COMADRE demographic databases; however, we believe these benefits extend beyond COMPADRE and COMADRE users towards the whole field of population ecology and fields that use MPMs for their own inference (e.g. conservation biology and biodiversity monitoring). A greater level of detail and transparency when describing how and why an MPM is produced will result in greater accuracy, accessibility, reproducibility and citability—this has clear benefits to the field as a whole and to individual researchers. In addition, greater consistency and transparency facilitates peer review, and indeed, these guidelines may offer a tool that can be cited by associate editors and peer-reviewers who may frequently advocate some (or all) of the steps suggested herein. Furthermore, adoption of the steps suggested here may increase confidence in the results presented and facilitate learning/uptake of MPMs by early career researchers.

Finally, we close with a caution. We have used the term ‘accurate’ at points throughout this paper, applied to MPMs, but we must acknowledge that there is no such thing as an accurate model, be it an MPM or any other type. A model is a series of choices, choices of aspects that are included and aspects that are neglected. Model selection techniques such as AIC (Anderson & Burnham, 2002) make these choices explicit and measure their support in terms of likelihood. But even without using the explicit statistical method, the message is clear. Choices of *i*-state variables, of projection intervals, of types of time variation, of functional dependence on a chosen set of environmental factors and so forth, all of these are inaccurate. The point is not to seek for accuracy: it is to be clear about communicating the choices you made in constructing the model, the analyses you chose to apply and the interpretation of the results. An ‘accurate’ model of an ecological system, experimental or observational, would be as complicated as the real system. That does not end well (Borges, 1999).

AUTHOR CONTRIBUTIONS

This paper was conceptualised at a workshop hosted by Jenni McDonald, Danielle L. Buss and Dave Hodgson. Subsequently, Samuel J. L. Gascoigne, Simon Rolph, Christina M. Hernández and Roberto Salguero-Gómez generated the first draft of the manuscript with ideas contributed from all authors. Samuel J. L. Gascoigne conducted the survey. Samuel J. L. Gascoigne, Daisy Sankey, Nagalakshmi Nidadavolu and Adrian S. Stell Pičman conducted the screen of papers from COMPADRE and COMADRE. All authors contributed ideas in drafts of the manuscript.

ACKNOWLEDGEMENTS

We thank the hundreds of population ecologists who have contributed open-access matrix population models ready for fully reproducible research, and those who have, throughout the last 15 years, answered our emails asking for additional data and metadata. We also thank Chloé R. Nater and one anonymous reviewer for their helpful comments. Conversations leading to this manuscript were initiated during a workshop held at the University of Exeter Cornwall campus with support of NERC grant (NE/N006798/1) to Jenni

McDonald and Dave Hodgson. We acknowledge the support of the Evolutionary Demography Laboratory at the Max Planck Institute for Demographic Research (MPIDR) in the development of the COMPADRE Plant Matrix Database and COMADRE Animal Matrix Database, and the maintenance of COMPADRE & COMADRE through the distributed network of digitising nodes including MPIDR, University of Oxford, University of Exeter, Southampton University, Trinity College Dublin, Lincoln Park Zoo, University of Southern Denmark and iDiv. Roberto Salguero-Gómez was supported by a NERC Independent Research Fellowship (NE/M018458/1). Christina M. Hernández was supported by a US-NSF grant (DEB-1933497). Pol Capdevila was supported by a Maria Zambrano Next Generation EU Fellowship. Judy P. Che-Castaldo, Owen R. Jones, Roberto Salguero-Gómez and Chelsea C. Thomas were supported by an NSF Advances in Bioinformatics Development Award (#DBI-1661342). Danielle L. Buss was funded by NERC Discovery Science grant NE/L007770/1. Tiffany M. Knight was supported by the Helmholtz Association and the Alexander von Humboldt foundation. Hal Caswell was supported by the European Research Council, Advanced Grant 788195. Lastly, this paper is in memoriam of our dear friend and colleague James W. Vaupel, who sadly passed before the submission of this manuscript. His multiple contributions to demography will no doubt outlive multiple Bristlecone pine generation times.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14164>.

DATA AVAILABILITY STATEMENT

The code used in this paper can be found at Zenodo (Gascoigne, 2023, <https://zenodo.org/record/7967075>) and on GitHub at https://github.com/SamuelGascoigne/Standard_MPM_Protocol.

ORCID

Samuel J. L. Gascoigne  <https://orcid.org/0000-0002-2984-1810>

Simon Rolph  <https://orcid.org/0000-0001-6755-9456>

Christina M. Hernández  <https://orcid.org/0000-0002-7188-8217>

Connor Bernard  <https://orcid.org/0000-0003-2175-5712>

Erola Fenollosa  <https://orcid.org/0000-0002-6189-2124>

Oliver G. Spacey  <https://orcid.org/0000-0002-0280-8201>

Maja Kajin  <https://orcid.org/0000-0001-9996-5897>

Anna C. Vinton  <https://orcid.org/0000-0002-8279-1736>

C. Ruth Archer  <https://orcid.org/0000-0003-2501-425X>

Jean H. Burns  <https://orcid.org/0000-0001-6140-9094>

Danielle L. Buss  <https://orcid.org/0000-0001-5362-240X>

Hal Caswell  <https://orcid.org/0000-0003-4394-6894>

Judy P. Che-Castaldo  <https://orcid.org/0000-0002-9118-9202>

Dylan Z. Childs  <https://orcid.org/0000-0002-0675-4933>

Pol Capdevila  <https://orcid.org/0000-0002-2842-4302>

Aldo Compagnoni  <https://orcid.org/0000-0001-8302-7492>
 Elizabeth Crone  <https://orcid.org/0000-0002-5287-221X>
 Thomas H. G. Ezard  <https://orcid.org/0000-0001-8305-6605>
 Dave Hodgson  <https://orcid.org/0000-0003-4220-2076>
 Tiffany M. Knight  <https://orcid.org/0000-0003-0318-1567>
 Owen R. Jones  <https://orcid.org/0000-0001-5720-4686>
 Eelke Jongejans  <https://orcid.org/0000-0003-1148-7419>
 Jenni McDonald  <https://orcid.org/0000-0002-5050-750X>
 Brigitte Tenhumberg  <https://orcid.org/0000-0003-3801-3192>
 Andrew J. Tyre  <https://orcid.org/0000-0001-9736-641X>
 Satu Ramula  <https://orcid.org/0000-0001-7795-0352>
 Iain Stott  <https://orcid.org/0000-0002-1399-3596>
 Raymond L. Tremblay  <https://orcid.org/0000-0002-8588-4372>
 James W. Vaupel  <https://orcid.org/0000-0003-0783-3905>
 Roberto Salguero-Gómez  <https://orcid.org/0000-0002-6085-4433>

REFERENCES

- Adler, P. B., Salguero-Gómez, R., Compagnoni, A., Hsu, J. S., Ray-Mukherjee, J., Mbeau-Ache, C., & Franco, M. (2014). Functional traits explain variation in plant life history strategies. *Proceedings of the National Academy of Sciences of the United States of America*, 111(27), 10019. <https://doi.org/10.1073/pnas.1410430111>
- Anderson, D. R., & Burnham, K. P. (2002). Avoiding pitfalls when using information-theoretic methods. *The Journal of Wildlife Management*, 66(3), 912–918.
- Archer, C. R., Paniw, M., Vega-Trejo, R., & Sepil, I. (2022). A sex skew in life-history research: The problem of missing males. *Proceedings of the Royal Society B: Biological Sciences*, 289. <https://doi.org/10.1098/rspb.2022.1117>
- Baudisch, A., Salguero-Gómez, R., Jones, O. R., Wrycza, T., Mbeau-Ache, C., Franco, M., & Colchero, F. (2013). The pace and shape of senescence in angiosperms. *Journal of Ecology*, 101(3), 596–606. <https://doi.org/10.1111/1365-2745.12084>
- Beissinger, S. R., & Westphal, M. I. (1998). On the use of demographic models of population viability in endangered species management. *The Journal of Wildlife Management*, 62(3), 821. <https://doi.org/10.2307/3802534>
- Bernard, C., Santos, G. S., Deere, J. A., Rodriguez-Caro, R., Capdevila, P., Kusch, E., Gascoigne, S. J. L., Jackson, J., & Salguero-Gómez, R. (2023). MOSAIC: A unified trait database to complement structured population models. *Scientific Data*, 10(335). <https://doi.org/10.1038/s41597-023-02070-w>
- Bock, M. J., Jarvis, G. C., Corey, E. L., Stone, E. E., & Gribble, K. E. (2019). Maternal age alters offspring lifespan, fitness, and lifespan extension under caloric restriction. *Scientific Reports*, 9(1), 1–16. <https://doi.org/10.1038/s41598-019-40011-z>
- Borges, J. L. (1999). On exactitude in science. In J. L. Borges (Ed.), *Collected fictions*, translated by Andrew Hurley. Penguin Random House. Originally published as *Del Rigor en la Ciencia*, in Borges, J. L. (1954), *Historia Universal de la Infamia*.
- Boukal, D. S., & Berec, L. (2002). Single-species models of the Allee effect: Extinction boundaries, sex ratios and mate encounters. *Journal of the Torrey Botanical Society*, 218, 375–394. <https://doi.org/10.1006/yjtbi.3084>
- Bruijning, M., Jongejans, E., & Turcotte, M. M. (2019). Demographic responses underlying eco-evolutionary dynamics as revealed with inverse modelling. *Journal of Animal Ecology*, 88(5), 768–779. <https://doi.org/10.1111/1365-2656.12966>
- Buckley, Y. M., Ramula, S., Blomberg, S. P., Burns, J. H., Crone, E. E., Ehrlén, J., Knight, T. M., Pichancourt, J. B., Quedsted, H., & Wardle, G. M. (2010). Causes and consequences of variation in plant population growth rate: A synthesis of matrix population models in a phylogenetic context. *Ecology Letters*, 13(9), 1182–1197. <https://doi.org/10.1111/j.1461-0248.2010.01506.x>
- Capdevila, P., Stott, I., Beger, M., & Salguero-Gómez, R. (2020). Towards a comparative framework of demographic resilience. *Trends in Ecology & Evolution*, 35(9), 776–786. <https://doi.org/10.1016/j.tree.2020.05.001>
- Caswell, H. (2001). *Matrix population models: Construction, analysis, and interpretation* (2nd ed.). Sinauer.
- Caswell, H. (2009). Stage, age and individual stochasticity in demography. *Oikos*, 118(12), 1763–1782. <https://doi.org/10.1111/j.1600-0706.2009.17620.x>
- Caswell, H. (2011). Beyond R0: Demographic models for variability of lifetime reproductive output. *PLoS ONE*, 6(6), e20809. <https://doi.org/10.1371/journal.pone.0020809>
- Caswell, H. (2012). Matrix models and sensitivity analysis of populations classified by age and stage: A vec-permutation matrix approach. *Theoretical Ecology*, 5(3), 403–417. <https://doi.org/10.1007/s12080-011-0132-2>
- Caswell, H. (2013). Sensitivity analysis of discrete Markov chains via matrix calculus. *Linear Algebra and its Applications*, 438(4), 1727–1745. <https://doi.org/10.1016/j.laa.2011.07.046>
- Caswell, H. (2019a). *Sensitivity analysis: Matrix methods in demography and ecology*. Springer Nature.
- Caswell, H. (2019b). The formal demography of kinship: A matrix formulation. *Demographic Research*, 41(24), 679–712. <https://doi.org/10.4054/DemRes.2019.41.24>
- Caswell, H. (2020). The formal demography of kinship II: Multistate models, parity, and sibship. *Demographic Research*, 42(38), 109–114. <https://doi.org/10.4054/DemRes.2020.42.38>
- Caswell, H., Brault, S., Read, A. J., & Smith, T. D. (1998). Harbor porpoise and fisheries: An uncertainty analysis of incidental mortality. *Ecological Applications*, 8(4), 1226. <https://doi.org/10.2307/2640975>
- Caswell, H., de Vries, C., Hartemink, N., Roth, G., & van Daalen, S. F. (2018). Age × stage-classified demographic analysis: A comprehensive approach. *Ecological Monographs*, 88(4), 560–584. <https://doi.org/10.1002/ecm.1306>
- Caswell, H., & Salguero-Gómez, R. (2013). Age, stage and senescence in plants. *Journal of Ecology*, 101(3), 585–595. <https://doi.org/10.1111/1365-2745.12088>
- Caswell, H., & Van Daalen, S. (2021). Healthy longevity from incidence-based models: More kinds of health than stars in the sky. *Demographic Research*, 45(December), 397–452. <https://doi.org/10.4054/DEMRES.2021.45.13>
- Che-Castaldo, J., Havercamp, K., Watanuki, K., Matsuzawa, T., Hirata, S., & Ross, S. R. (2021). Comparative survival analyses among captive chimpanzees (*Pan troglodytes*) in America and Japan. *PeerJ*, 9, 1–21. <https://doi.org/10.7717/peerj.11913>
- Che-Castaldo, J., Jones, O. R., Kendall, B. E., Burns, J. H., Childs, D. Z., Ezard, T. H. G., Hernandez-Yanez, H., Hodgson, D. J., Jongejans, E., Knight, T., Merow, C., Ramula, S., Stott, I., Vindenes, Y., Vindenes, Y., Yokomizo, H., & Salguero-Gómez, R. (2020). Comments to 'Persistent problems in the construction of matrix population models'. *Ecological Modelling*, 416, 1–3. <https://doi.org/10.1016/j.ecolmodel.2019.108913>
- Childs, D. Z., Sheldon, B. C., & Rees, M. (2016). The evolution of labile traits in sex- and age-structured populations. *Journal of Animal Ecology*, 85(2), 329–342. <https://doi.org/10.1111/1365-2656.12483>
- Choquet, R., Rouan, L., & Pradel, R. (2009). Program E-SURGE: A software for fitting multievent models. In D. L. Thomson, E. G. Cooch, & M. J. Conroy (Eds.), *Modeling demographic processes in marked populations* (pp. 845–865). Springer. <https://doi.org/10.1007/978-0-387-78151-8>

- Clubb, R., Rowcliffe, M., Lee, P., Mar, K. U., Moss, C., & Mason, G. J. (2009). Fecundity and population viability in female zoo elephants: Problems and possible solutions. *Animal Welfare*, 18(3), 237–247.
- Clubb, R., & Mason, G. (2003). Captivity effects on wide-range carnivores. *Nature*, 425(6957), 473–474.
- Compagnoni, A., Steigman, K., & Miller, T. E. X. (2017). Can't live with them, can't live without them? Balancing mating and competition in two-sex populations. *Proceedings of the Royal Society B: Biological Sciences*, 284(1865), 2–11. <https://doi.org/10.1098/rspb.2017.1999>
- Conde, D. A., Staerk, J., Colchero, F., da Silva, R., Schöley, J., Baden, H. M., Jouvett, L., Fa, J. E., Syed, H., Jongejans, E., Meiri, S., Gaillard, J. M., Chamberlain, S., Wilcken, J., Jones, O. R., Dahlgren, J. P., Steiner, U. K., Bland, L. M., Gomez-Mestre, I., ... Vaupel, J. W. (2019). Data gaps and opportunities for comparative and conservation biology. *Proceedings of the National Academy of Sciences of the United States of America*, 116(19), 9658–9664. <https://doi.org/10.1073/pnas.1816367116>
- Cooch, E. G., Gauthier, G., & Rockwell, R. F. (2003). Apparent differences in stochastic growth rates based on timing of census: A cautionary note. *Ecological Modelling*, 159(2–3), 133–143. [https://doi.org/10.1016/S0304-3800\(02\)00292-2](https://doi.org/10.1016/S0304-3800(02)00292-2)
- Coulson, T., Catchpole, E. A., Albon, S. D., Morgan, B. J. T., Pemberton, J. M., Clutton-Brock, T. H., Crawley, M. J., & Grenfell, B. T. (2001). Age, sex, density, winter weather, and population crashes in Soay sheep. *Science*, 292(5521), 1528–1531. <https://doi.org/10.1126/science.292.5521.1528>
- Cushing, J. M., Constantino, R. F., Dennis, B., Desharnais, R., & Henson, S. M. (2003). *Chaos in ecology: Experimental nonlinear dynamics*. Elsevier.
- Davis, K. J. (2022). Managed culls mean extinction for a marine mammal population when combined with extreme climate impacts. *Ecological Modelling*, 473(November 2021), 110122. <https://doi.org/10.1016/j.ecolmodel.2022.110122>
- De Magalhães, J. P., & Costa, J. (2009). A database of vertebrate longevity records and their relation to other life-history traits. *Journal of Evolutionary Biology*, 22(8), 1770–1774. <https://doi.org/10.1111/j.1420-9101.2009.01783.x>
- de Vries, C., & Caswell, H. (2019). Selection in two-sex stage-structured populations: Genetics, demography, and polymorphism. *Theoretical Population Biology*, 130, 160–169. <https://doi.org/10.1016/j.tpb.2019.07.012>
- de Vries, C., Desharnais, R. A., & Caswell, H. (2020). A matrix model for density-dependent selection in stage-classified populations, with application to pesticide resistance in *Tribolium*. *Ecological Modelling*, 416(November 2018), 108875. <https://doi.org/10.1016/j.ecolmodel.2019.108875>
- Demetrius, L. (1992). Growth rate, population entropy, and evolutionary dynamics. *Theoretical Population Biology*, 41(2), 208–236. [https://doi.org/10.1016/0040-5809\(92\)90044-T](https://doi.org/10.1016/0040-5809(92)90044-T)
- Doak, D. F., Morris, W. F., Pfister, C., Kendall, B. E., & Bruna, E. M. (2005). Correctly estimating how environmental stochasticity influences fitness and population growth. *The American Naturalist*, 166(1), E14–E21. <https://doi.org/10.1086/430642>
- Doak, D. F., Thomson, D., & Jules, E. S. (2002). Population viability analysis for plants: Understanding the demographic consequences of seed banks for population health. In S. R. Beissinger & D. R. McCullough (Eds.), *Population viability analysis* (pp. 312–337). University of Chicago Press.
- Ebert, T. A. (1999). Populations methods in demography. *Methods & Demography*, January 1999, 2.
- Ellner, S. P., Childs, D. Z., & Rees, M. (2016). *Data-driven modelling of structured populations*.
- Emery, S. M., & Gross, K. L. (2005). Effects of timing of prescribed fire on the demography of an invasive plant, spotted knapweed *Centaurea maculosa*. *Journal of Applied Ecology*, 42(1), 60–69. <https://doi.org/10.1111/j.1365-2664.2004.00990.x>
- Ezard, T. H. G., Bullock, J. M., Dalglish, H. J., Millon, A., Pelletier, F., Ozgul, A., & Koons, D. N. (2010). Matrix models for a changeable world: The importance of transient dynamics in population management. *Journal of Applied Ecology*, 47(3), 515–523. <https://doi.org/10.1111/j.1365-2664.2010.01801.x>
- Ferreira, M. S., Kajin, M., Cerqueira, R., & Vieira, M. V. (2016). Marsupial population dynamics in a tropical rainforest: Intraspecific competition and nonlinear effect of rainfall. *Journal of Mammalogy*, 97(1), 121–127. <https://doi.org/10.1093/jmammal/gyv161>
- Franco, M., & Silvertown, J. (1996). Life history variation in plants: An exploration of the fast-slow continuum hypothesis. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 351(1345), 1341–1348. <https://doi.org/10.1098/rstb.1996.0117>
- Fujiwara, M., & Caswell, H. (2001). Demography of the endangered North Atlantic right whale. *Nature*, 414(6863), 537–541. <https://doi.org/10.1038/35107054>
- Gascoigne, S. (2023). SamuelGascoigne/Standard_MPM_Protocol: Initial release (1.0.0). Zenodo, <https://doi.org/10.5281/zenodo.7967075>
- Gerstner, K., Moreno-Mateos, D., Gurevitch, J., Beckmann, M., Kambach, S., Jones, H. P., & Seppelt, R. (2017). Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. *Methods in Ecology and Evolution*, 8(6), 777–784. <https://doi.org/10.1111/2041-210X.12758>
- Gontijo, L. M., & Carvalho, R. M. R. (2020). Using life stage-structured matrix models to determine natural enemy:pest release ratios for augmentative biological control. *Journal of Applied Entomology*, 144(5), 364–372. <https://doi.org/10.1111/jen.12734>
- Gremer, J. R., & Sala, A. (2013). It is risky out there: The costs of emergence and the benefits of prolonged dormancy. *Oecologia*, 172(4), 937–947. <https://doi.org/10.1007/s00442-012-2557-8>
- Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, 555(7695), 175–182. <https://doi.org/10.1038/nature25753>
- Hartemink, N., Missov, T. I., & Caswell, H. (2017). Stochasticity, heterogeneity, and variance in longevity in human populations. *Theoretical Population Biology*, 114, 107–116. <https://doi.org/10.1016/j.tpb.2017.01.001>
- Healy, K., Ezard, T. H. G., Jones, O. R., Salguero-Gómez, R., & Buckley, Y. M. (2019). Animal life history is shaped by the pace of life and the distribution of age-specific mortality and reproduction. *Nature Ecology and Evolution*, 3(8), 1217–1224. <https://doi.org/10.1038/s41559-019-0938-7>
- Hernández, C. M., van Daalen, S. F., Caswell, H., Neubert, M. G., & Gribble, K. E. (2020). A Demographic and Evolutionary Analysis of Maternal Effect Senescence. *Proceedings of the National Academy of Sciences of the United States of America*, 117(28), 16431–16437. <https://doi.org/10.1073/pnas.1919988117>
- Hunter, C. M., & Caswell, H. (2005). Selective harvest of sooty shearwater chicks: Effects on population dynamics and sustainability. *Journal of Animal Ecology*, 74(4), 589–600. <https://doi.org/10.1111/j.1365-2656.2005.00929.x>
- Hunter, C. M., Caswell, H., Runge, M. C., Regehr, E. V., Amstrup, S. C., & Stirling, I. (2010). Climate change threatens polar bear populations: A stochastic demographic analysis. *Ecology*, 91(10), 2883–2897. <https://doi.org/10.1890/09-1641.1>
- James, T. D., Salguero-Gómez, R., Jones, O. R., Childs, D. Z., & Beckerman, A. P. (2021). Bridging gaps in demographic analysis with phylogenetic imputation. *Conservation Biology*, 35(4), 1210–1221. <https://doi.org/10.1111/cobi.13658>
- Jelbert, K., Buss, D., McDonald, J., Townley, S., Franco, M., Stott, I., Jones, O., Salguero-Gómez, R., Buckley, Y., Knight, T., Silk, M., Sargent, F., Rolph, S., Wilson, P., & Hodgson, D. (2019). Demographic amplification is a predictor of invasiveness among plants. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-13556-w>
- Jenouvrier, S., Aubry, L. M., Barbraud, C., Weimerskirch, H., & Caswell, H. (2018). Interacting effects of unobserved heterogeneity

- and individual stochasticity in the life history of the southern fulmar. *Journal of Animal Ecology*, 87(1), 212–222. <https://doi.org/10.1111/1365-2656.12752>
- Jenouvrier, S., Caswell, H., Barbraud, C., & Weimerskirch, H. (2010). Mating behavior, population growth, and the operational sex ratio: A periodic two-sex model approach. *The American Naturalist*, 175(6), 739–752. <https://doi.org/10.1086/652436>
- Jenouvrier, S., Holland, M., Stroeve, J., Barbraud, C., Weimerskirch, H., Serreze, M., & Caswell, H. (2012). Effects of climate change on an emperor penguin population: Analysis of coupled demographic and climate models. *Global Change Biology*, 18(9), 2756–2770. <https://doi.org/10.1111/j.1365-2486.2012.02744.x>
- Jenouvrier, S., Long, M. C., Coste, C. F. D., Holland, M., Gamelon, M., Yoccoz, N. G., & Sæther, B. E. (2022). Detecting climate signals in populations across life histories. *Global Change Biology*, 28(7), 2236–2258. <https://doi.org/10.1111/gcb.16041>
- Jia, P., Lu, L., Chen, X., Chen, J., Guo, L., Yu, X., & Liu, Q. (2016). A climate-driven mechanistic population model of *Aedes albopictus* with diapause. *Parasites & Vectors*, 9(1), 1–15. <https://doi.org/10.1186/s13071-016-1448-y>
- Jiménez-Valdés, M., Godínez-Alvarez, H., Caballero, J., & Lira, R. (2010). Population dynamics of *Agave marmorata* Roezl. Under two contrasting management systems in Central Mexico. *Economic Botany*, 64(2), 149–160. <https://doi.org/10.1007/s12231-010-9117-0>
- Johnson, D. J., Needham, J., Xu, C., Massoud, E. C., Davies, S. J., Anderson-Teixeira, K. J., Bunyavejchewin, S., Chambers, J. Q., Chang-Yang, C. H., Chiang, J. M., Chuyong, G. B., Condit, R., Cordell, S., Fletcher, C., Giardina, C. P., Giambelluca, T. W., Gunatilleke, N., Gunatilleke, S., Hsieh, C. F., ... McMahon, S. M. (2018). Climate sensitive size-dependent survival in tropical trees. *Nature Ecology and Evolution*, 2(9), 1436–1442. <https://doi.org/10.1038/s41559-018-0626-z>
- Jones, F. A., & Hubbell, S. P. (2006). Demographic spatial genetic structure of the Neotropical tree, *Jacaranda copaia*. *Molecular Ecology*, 15(11), 3205–3217. <https://doi.org/10.1111/j.1365-294X.2006.03023.x>
- Jones, O. R., Barks, P., Stott, I., James, T. D., Levin, S., Petry, W. K., Capdevila, P., Che-Castaldo, J., Jackson, J., Römer, G., Schuette, C., Thomas, C. C., & Salguero-Gómez, R. (2022). Rcompadre and Rage—Two R packages to facilitate the use of the COMPADRE and COMADRE databases and calculation of life-history traits from matrix population models. *Methods in Ecology and Evolution*, 13(4), 770–781. <https://doi.org/10.1111/2041-210X.13792>
- Jones, O. R., Scheuerlein, A., Salguero-Gómez, R., Camarda, C. G., Schaible, R., Casper, B. B., Dahlgren, J. P., Ehrén, J., García, M. B., Menges, E. S., Quintana-Ascencio, P. F., Caswell, H., Baudisch, A., & Vaupel, J. W. (2014). Diversity of ageing across the tree of life. *Nature*, 505(7482), 169–173. <https://doi.org/10.1038/nature12789>
- Jouvet, L., Rodríguez-Rojas, A., & Steiner, U. K. (2018). Demographic variability and heterogeneity among individuals within and among clonal bacteria strains. *Oikos*, 127(5), 728–737. <https://doi.org/10.1111/oik.04292>
- Käfer, J., Marais, G. A. B., & Pannell, J. R. (2017). On the rarity of dioecy in flowering plants. *Molecular Ecology*, 26(5), 1225–1241. <https://doi.org/10.1111/mec.14020>
- Kendall, B. E., Fujiwara, M., Diaz-Lopez, J., Schneider, S., Voigt, J., & Wiesner, S. (2019). Persistent problems in the construction of matrix population models. *Ecological Modelling*, 406(March), 33–43. <https://doi.org/10.1016/j.ecolmodel.2019.03.011>
- Keyfitz, N. (1968). Changing vital rates and age distributions. *Population Studies*, 22(2), 235–251. <https://doi.org/10.1080/00324728.1968.10405537>
- Klepac, P., & Caswell, H. (2011). The stage-structured epidemic: Linking disease and demography with a multi-state matrix approach model. *Theoretical Ecology*, 4(3), 301–319. <https://doi.org/10.1007/s12080-010-0079-8>
- Lawler, R. R., Caswell, H., Richard, A. F., Ratsirarson, J., Dewar, R. E., & Schwartz, M. (2009). Demography of Verreaux's sifaka in a stochastic rainfall environment. *Oecologia*, 161(3), 491–504. <https://doi.org/10.1007/s00442-009-1382-1>
- Lebreton, J. D., Burnham, K. P., Clobert, J., & Anderson, D. R. (1992). Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. *Ecological Monographs*, 62(1), 67–118. <https://doi.org/10.2307/2937171>
- Lefkovich, L. P. (1965). The study of population growth in organisms grouped by stages. *Biometrics*, 21(1), 1–18.
- Leslie, P. H. (1945). On the use of matrices in certain population. *Biometrika*, 33(3), 183–212.
- Levin, S. C., Evers, S., Potter, T., Guerrero, M. P., Childs, D. Z., Compagnoni, A., Knight, T. M., & Salguero-Gómez, R. (2022). Rpadrino: An R package to access and use PADRINO, an open access database of Integral Projection Models. *Methods in Ecology and Evolution*, 2022(May), 1–7. <https://doi.org/10.1111/2041-210X.13910>
- Li, W. H., Ju, Y. R., Liao, C. M., & Liao, V. H. C. (2014). Assessment of selenium toxicity on the life cycle of *Caenorhabditis elegans*. *Ecotoxicology*, 23(7), 1245–1253. <https://doi.org/10.1007/s10646-014-1267-x>
- McDonald, J. L., Stott, I., Townley, S., & Hodgson, D. J. (2016). Transients drive the demographic dynamics of plant populations in variable environments. *Journal of Ecology*, 104(2), 306–314. <https://doi.org/10.1111/1365-2745.12528>
- Miller, T. E. X., & Compagnoni, A. (2022). Two-sex demography, sexual niche differentiation, and the geographic range limits of Texas bluegrass (*Poa arachnifera*). *The American Naturalist*, 200(1), 17–38. <https://doi.org/10.1086/719668>
- Namkoong, G., & Roberds, J. H. (1974). Extinction probabilities and the changing age structure of redwood forests. *The American Naturalist*, 108(961), 355–368.
- Neubert, M. G., & Caswell, H. (2000). Demography and dispersal: Calculation and sensitivity analysis of invasion speed for structured populations. *Ecology*, 81(6), 1613–1628. [https://doi.org/10.1890/0012-9658\(2000\)081\[1613:DADCAS\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[1613:DADCAS]2.0.CO;2)
- Nguyen, V., Buckley, Y. M., Salguero-Gómez, R., & Wardle, G. M. (2019). Consequences of neglecting cryptic life stages from demographic models. *Ecological Modelling*, 408(July), 108723. <https://doi.org/10.1016/j.ecolmodel.2019.108723>
- Nordstrom, S. W., Dykstra, A. B., & Wagenius, S. (2021). Fires slow population declines of a long-lived prairie plant through multiple vital rates. *Oecologia*, 196(3), 679–691. <https://doi.org/10.1007/s00442-021-04955-2>
- Omeyer, L. C. M., Stokes, K. L., Beton, D., Çiçek, B. A., Davey, S., Fuller, W. J., Godley, B. J., Sherley, R. B., Snape, R. T. E., & Broderick, A. C. (2021). Investigating differences in population recovery rates of two sympatrically nesting sea turtle species. *Animal Conservation*, 24, 832–846. <https://doi.org/10.1111/acv.12689>
- Paniw, M., Maag, N., Cozzi, G., Clutton-Brock, T., & Ozgul, A. (2019). Life history responses of meerkats to seasonal changes in extreme environments. *Science*, 363(6427), 631–635. <https://doi.org/10.1126/science.aau5905>
- Paniw, M., Quintana-Ascencio, P. F., Ojeda, F., & Salguero-Gómez, R. (2017). Accounting for uncertainty in dormant life stages in stochastic demographic models. *Oikos*, 126(6), 900–909. <https://doi.org/10.1111/oik.03696>
- Pardini, E. A., Drake, J. M., Chase, J. M., & Knight, T. M. (2009). Complex population dynamics and control of the invasive biennial *Alliaria petiolata* (Garlic mustard). *Ecological Applications*, 19(2), 387–397. <https://doi.org/10.1890/08-0845.1>
- Pearl, R., Miner, J. R., & Parker, S. L. (1927). Experimental studies on the duration of life. XI. Density of population and life duration in *Drosophila*. *The American Naturalist*, 61, 289–318. <https://doi.org/10.1086/280154>
- Pfister, C. A. (1998). Patterns of variance in stage-structured populations: Evolutionary predictions and ecological implications. *Proceedings of the National Academy of Sciences of the United States of America*, 95(1), 213–218. <https://doi.org/10.1073/pnas.95.1.213>

- Plard, F., Fay, R., Kéry, M., Cohas, A., & Schaub, M. (2019). Integrated population models: Powerful methods to embed individual processes in population dynamics models. *Ecology*, 100(6), 1–15. <https://doi.org/10.1002/ecy.2715>
- Powers, S. M., & Hampton, S. E. (2019). Open science, reproducibility, and transparency in ecology. *Ecological Applications*, 29(1), 1–8. <https://doi.org/10.1002/eap.1822>
- Regehr, E. V., Hunter, C. M., Caswell, H., Amstrup, S. C., & Stirling, I. (2010). Survival and breeding of polar bears in the southern Beaufort Sea in relation to sea ice. *Journal of Animal Ecology*, 79(1), 117–127. <https://doi.org/10.1111/j.1365-2656.2009.01603.x>
- Reichman, O. J., Jones, M. B., & Schildhauer, M. P. (2011). Challenges and opportunities of open data in ecology. *Science*, 331(6018), 703–705. <https://doi.org/10.1126/science.1197962>
- Riecke, T. V., Williams, P. J., Behnke, T. L., Gibson, D., Leach, A. G., Sedinger, B. S., Street, P. A., & Sedinger, J. S. (2019). Integrated population models: Model assumptions and inference. *Methods in Ecology and Evolution*, 10(7), 1072–1082. <https://doi.org/10.1111/2041-210X.13195>
- Römer, G., Dahlgren, J. P., Salguero-Gómez, R., Stott, I. M., & Jones, O. R. (2021). Plant demographic knowledge is biased towards short-term studies of temperate-region herbaceous perennials. *BioRxiv*, 1–46. <https://doi.org/10.1101/2021.04.25.441327>
- Roth, G., & Caswell, H. (2016). Hyperstate matrix models: Extending demographic state spaces to higher dimensions. *Methods in Ecology and Evolution*, 7(12), 1438–1450. <https://doi.org/10.1111/2041-210X.12622>
- Rueda-Cediel, P., Anderson, K. E., Regan, T. J., & Regan, H. M. (2018). Effects of uncertainty and variability on population declines and IUCN Red List classifications. *Conservation Biology*, 32(4), 916–925. <https://doi.org/10.1111/cobi.13081>
- Sæther, B. E., Coulson, T., Grøtan, V., Engen, S., Altwegg, R., Armitage, K. B., Barbraud, C., Becker, P. H., Blumstein, D. T., Dobson, F. S., Festa-Bianchet, M., Gaillard, J.-M., Jenkins, A., Jones, C., Nicoll, M. A. C., Norris, K., Oli, M. K., Ozgul, A., & Weimerskirch, H. (2013). How life history influences population dynamics in fluctuating environments. *The American Naturalist*, 182(6), 743–759. <https://doi.org/10.1086/673497>
- Salguero-Gómez, R., Jackson, J., & Gascoigne, S. J. L. (2021). Four key challenges in the open-data revolution. *Journal of Animal Ecology*, 90(9), 2000–2004. <https://doi.org/10.1111/1365-2656.13567>
- Salguero-Gómez, R., Jones, O. R., Archer, C. R., Bein, C., de Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Römer, G., Ruoff, T., Sommer, V., Wille, J., Voigt, J., Zeh, S., Viereg, D., Buckley, Y. M., Che-Castaldo, J., ... Vaupel, J. W. (2016). COMADRE: A global data base of animal demography. *Journal of Animal Ecology*, 85, 371–384. <https://doi.org/10.1111/1365-2656.12482>
- Salguero-Gómez, R., Jones, O. R., Archer, C. R., Buckley, Y. M., Che-Castaldo, J., Caswell, H., Hodgson, D., Scheuerlein, A., Conde, D. A., Brinks, E., de Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Römer, G., Runge, J., Ruoff, T., ... Vaupel, J. W. (2015). The COMPADRE Plant Matrix Database: An open online repository for plant demography. *Journal of Ecology*, 103, 202–218. <https://doi.org/10.1111/1365-2745.12334>
- Salguero-Gómez, R., Jones, O. R., Blomberg, S. P., Hodgson, D. J., Zuidema, P. A., & Kroon, H. D. (2017). Erratum: Fast-slow continuum and reproductive strategies structure plant life-history variation worldwide (Proc Natl Acad Sci USA (2015) 113 (230–235)). *Proceedings of the National Academy of Sciences of the United States of America*, 114(45), E9753. <https://doi.org/10.1073/pnas.1717717114>
- Salguero-Gómez, R., & Plotkin, J. B. (2010). Matrix dimensions bias demographic inferences: Implications for comparative plant demography. *The American Naturalist*, 176(6), 710–722. <https://doi.org/10.1086/657044>
- Schaub, M., & Kéry, M. (2021). *Integrated population models: Theory and ecological applications with R and JAGS*. Academic Press.
- Shyu, E., & Caswell, H. (2016a). A demographic model for sex ratio evolution and the effects of sex-biased offspring costs. *Ecology and Evolution*, 6(5), 1470–1492. <https://doi.org/10.1002/ece3.1902>
- Shyu, E., & Caswell, H. (2016b). Frequency-dependent two-sex models: A new approach to sex ratio evolution with multiple maternal conditions. *Ecology and Evolution*, 6(19), 6855–6879. <https://doi.org/10.1002/ece3.2202>
- Shyu, E., Pardini, E. A., Knight, T. M., & Caswell, H. (2013). A seasonal, density-dependent model for the management of an invasive weed. *Ecological Applications*, 23(8), 1893–1905. <https://doi.org/10.1890/12-1712.1>
- Sky, N. H., Jackson, J., Chege, G., Gaymer, J., Kimiti, D., Mutisya, S., Nakito, S., & Shultz, S. (2022). Female reproductive skew exacerbates the extinction risk from poaching in the eastern black rhino. *Proceedings of the Royal Society B: Biological Sciences*, 289(1972), 20220075. <https://doi.org/10.1098/rspb.2022.0075>
- Smith, M., Caswell, H., & Mettler-Cherry, P. (2005). Stochastic flood and precipitation regimes and the population dynamics of a threatened floodplain plant. *Ecological Applications*, 15(3), 1036–1052. <https://doi.org/10.1890/04-0434>
- Snyder, R. E., & Ellner, S. P. (2018). Pluck or luck: Does trait variation or chance drive variation in lifetime reproductive success? *The American Naturalist*, 191(4), E90–E107. <https://doi.org/10.1086/696125>
- Snyder, R. E., & Ellner, S. P. (2022). Snared in an evil time: How age-dependent environmental and demographic variability contribute to variance in lifetime outcomes. *The American Naturalist*, 200(3), E124–E140. <https://doi.org/10.1086/720411>
- Steiner, U. K., Tuljapurkar, S., & Roach, D. A. (2021). Quantifying the effect of genetic, environmental and individual demographic stochastic variability for population dynamics in *Plantago lanceolata*. *Scientific Reports*, 11(1), 1–11. <https://doi.org/10.1038/s41598-021-02468-9>
- Stott, I., Townley, S., Carslake, D., & Hodgson, D. J. (2010). On reducibility and ergodicity of population projection matrix models. *Methods in Ecology and Evolution*, 1(3), 242–252.
- Stott, I., Townley, S., & Hodgson, D. J. (2011). A framework for studying transient dynamics of population projection matrix models. *Ecology Letters*, 14(9), 959–970. <https://doi.org/10.1111/j.1461-0248.2011.01659.x>
- Tenhumberg, B., Louda, S. M., Eckberg, J. O., & Takahashi, M. (2008). Monte Carlo analysis of parameter uncertainty in matrix models for the weed *Cirsium vulgare*. *Journal of Applied Ecology*, 45(2), 438–447. <https://doi.org/10.1111/j.1365-2664.2007.01427.x>
- Tenhumberg, B., Tyre, A. J., & Rebarber, R. (2009). Model complexity affects transient population dynamics following a dispersal event: A case study with pea aphids. *Ecology*, 90(7), 1878–1890. <https://doi.org/10.1890/08-1665.1>
- Tremblay, R. L., & McCarthy, M. A. (2014). Bayesian estimates of transition probabilities in seven small lithophytic orchid populations: Maximizing data availability from many small samples. *PLoS ONE*, 9(7), 1–12. <https://doi.org/10.1371/journal.pone.0102859>
- Tremblay, R. L., Perez, M.-E., Larcombe, M., Brown, A., Quarmby, J., Bickerton, D., French, G., & Bould, A. (2009a). Dormancy in *Caladenia*: A Bayesian approach to evaluating latency. *Australian Journal of Botany*, 57(4), 340–350. <https://doi.org/10.1071/BT08163>
- Tremblay, R. L., Perez, M.-E., Larcombe, M., Brown, A., Quarmby, J., Bickerton, D., French, G., & Bould, A. (2009b). Population dynamics of *Caladenia*: Bayesian estimates of transition and extinction probabilities. *Australian Journal of Botany*, 57(4), 351–360. <https://doi.org/10.1071/BT08167>
- Tremblay, R. L., Tyre, A. J., Pérez, M. E., & Ackerman, J. D. (2021). Population projections from holey matrices: Using prior information to estimate rare transition events. *Ecological Modelling*, 447(August 2020), 109526. <https://doi.org/10.1016/j.ecolmodel.2021.109526>
- Tuljapurkar, S. (1989). An uncertain life: Demography in random environments in population analysis but it ignores variation in population vital rates. *Theoretical Population Biology (English Edition)*, 294, 227–294.

- van Daalen, S. F., & Caswell, H. (2017). Lifetime reproductive output: Individual stochasticity, variance, and sensitivity analysis. *Theoretical Ecology*, 10(3), 355–374. <https://doi.org/10.1007/s12080-017-0335-2>
- van Daalen, S. F., Hernández, C. M., Caswell, H., Neubert, M. G., & Gribble, K. E. (2022). The contributions of maternal age heterogeneity to variance in lifetime reproductive output. *The American Naturalist*, 199(5), 603–616. <https://doi.org/10.1086/718716>
- Veran, S., & Beissinger, S. R. (2009). Demographic origins of skewed operational and adult sex ratios: Perturbation analyses of two-sex models. *Ecology Letters*, 12(2), 129–143. <https://doi.org/10.1111/j.1461-0248.2008.01268.x>
- Werner, P. A., & Caswell, H. (1977). Population growth rates and age versus stage-distribution models for teasel (*Dipsacus sylvestris* Huds.). *Ecology*, 58(5), 1103–1111. <https://doi.org/10.2307/1936930>
- Werner, P. A., & Peacock, S. J. (2019). Savanna canopy trees under fire: Long-term persistence and transient dynamics from a stage-based matrix population model. *Ecosphere*, 10(5). <https://doi.org/10.1002/ecs2.2706>

SUPPORTING INFORMATION

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

Data S1: A survey on matrix communication.

How to cite this article: Gascoigne, S. J. L., Rolph, S., Sankey, D., Nidadavolu, N., Stell Pičman, A. S., Hernández, C. M., Philpott, M. E. R., Salam, A., Bernard, C., Fenollosa, E., Lee, Y. J., McLean, J., Hetti Achchige Perera, S., Spacey, O. G., Kajin, M., Vinton, A. C., Archer, C. R., Burns, J. H., Buss, D. L. ... Salguero-Gómez, R. (2023). A standard protocol to report discrete stage-structured demographic information. *Methods in Ecology and Evolution*, 00, 1–19. <https://doi.org/10.1111/2041-210X.14164>