bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ipmr: Flexibly implement Integral Projection Models in R

Sam C. Levin *[1,2,3], Dylan Z. Childs [4], Aldo Compagnoni [1,2,3], Sanne Evers [1,2,5], Tiffany M. Knight [1,2,5‡], Roberto Salguero-Gómez [3‡]

[1]Institute of Biology, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108 Halle (Saale), Germany

[2]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

[3]Department of Zoology, 11a Mansfield Rd, University of Oxford, Oxford, OX1 3SZ, UK

[4]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, S10 2TN, UK

[5]Department of Community Ecology, Helmholtz Centre for Environmental Research-UFZ, Theodor-Lieser-Straße 4, 06120, Halle (Saale), Germany

[‡]Joint senior authors

[*]Corresponding Author:

Sam C. Levin

Puschstrasse 4, 04103 Leipzig, Germany

email: levisc8@gmail.com

Running title - ipmr: Integral projection models in R

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Abstract

1. Integral projection models (IPMs) are an important tool for studying the dynamics of populations structured by one or more continuous traits (*e.g.* size, height, color). Researchers use IPMs to investigate questions ranging from linking drivers to plant population dynamics, planning conservation and management strategies, and quantifying selective pressures in natural populations. The popularity of stage-structured population models has been supported by *R* scripts and packages (*e.g.* `IPMpack`, `popbio`, `popdemo`, `lefko3`) aimed at ecologists, which have introduced a broad repertoire of functionality and outputs. However, pressing ecological, evolutionary, and conservation biology topics require developing more complex IPMs, and considerably more expertise to implement them. Here, we introduce `ipmr`, a flexible *R* package for building, analyzing, and interpreting IPMs.

2. The `ipmr` framework relies on the mathematical notation of the models to express them in code format. Additionally, this package decouples the model parameterization step from the model implementation step. The latter point substantially increases `ipmr`’s flexibility to model complex life cycles and demographic processes.

3. `ipmr` can handle a wide variety of models, including density dependence, discretely and continuously varying stochastic environments, and multiple continuous and/or discrete traits. `ipmr` can accommodate models with individuals cross-classified by age and size. Furthermore, the package provides methods for demographic analyses (*e.g.* asymptotic and stochastic growth rates) and visualization (*e.g.* kernel plotting).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

4. `ipmr` is a flexible *R* package for integral projection models. The package substantially reduces the amount of time required to implement general IPMs. We also provide extensive documentation with six vignettes and help files, accessible from an R session and online.

Keywords: elasticity, integral projection model, life history, population dynamics, population growth rate, sensitivity, structured populations

Introduction

Integral projection models (IPMs) are an important and widely used tool for ecologists studying structured population dynamics in discrete time. Since the paper introducing IPMs over two decades ago (Easterling et al., 2000), at least 255 peer-reviewed publications on at least 250 plant species and 60 animal species have used IPMs (Table S1, Fig. S1). These models have addressed questions ranging from invasive species population dynamics (*e.g.* Crandall & Knight, 2017), effect of climate drivers on population persistence (*e.g.* Compagnoni et al., 2020), evolutionary stable strategies (*e.g.* Childs et al., 2004), and rare/endangered species conservation (*e.g.* Ferrer-Cervantes et al., 2012).

The IPM was introduced as alternative to matrix population models, which model populations structured by discrete traits (Caswell, 2001). Some of the advantages of using an IPM include (i) the ability to model populations structured by continuously distributed traits, (ii) the ability to flexibly incorporate discrete and continuous traits in the same model (*e.g.* seeds in a seedbank and a height structured plant population (Crandall & Knight, 2017), or number of females, males, and age-1 recruits for fish species (Erickson et

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

al., 2017)), (iii) efficient parameterization of demographic processes with familiar regression methods (Coulson, 2012), and (iv) the numerical discretization of continuous kernels (see below) means that the tools available for matrix population models are usually also applicable for IPMs. Furthermore, researchers have developed methods to incorporate spatial dynamics (Jongejans et al., 2011), environmental stochasticity (Rees & Ellner, 2009), and density/frequency dependence into IPMs (Adler et al., 2010, Ellner et al., 2016). These developments were accompanied by the creation of software tools and guides to assist with IPM parameterization, implementation, and analysis. These tools range from *R* scripts with detailed annotations (Coulson, 2012, Merow et al., 2014, Ellner et al., 2016) to *R* packages (Metcalf et al., 2013, Shefferson et al., 2020).

Despite the array of resources available to researchers, implementing an IPM is still not a straightforward exercise. For example, an IPM that simulates a population for 100 time steps requires the user to either write or adapt from published guides multiple functions (*e.g.* to summarize demographic functions into the proper format), implement the numerical approximations of the model's integrals, ensure that individuals are not accidentally sent beyond the integration bounds ("unintentional eviction", *sensu* Williams et al., 2012), and track how the population state changes over the course of a simulation. Stochastic IPMs present further implementation challenges. In addition to the aforementioned elements, users must generate the sequence of environments that the population experiences. There are multiple ways of simulating environmental stochasticity, each with their own strengths and weaknesses (Metcalf et al. 2015).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

`ipmr` manages these key details while providing the user flexibility in their models. `ipmr` uses the `rlang` package for metaprogramming (Henry & Wickham, 2019), which enables `ipmr` to provide a miniature domain specific language for implementing IPMs. `ipmr` aims to mimic the mathematical syntax that describes IPMs as closely as possible (Fig. 1, Tables 1 and 2). This *R* package can handle models with individuals classified by a mixture of any number of continuously and discretely distributed traits. Furthermore, `ipmr` introduces specific classes and methods to deal with both discrete and continuous stochastic environments, density-independent and -dependent models, as well as age-structured populations. `ipmr` decouples the parameterization (*i.e.* regression model fitting) and implementation steps (*i.e.* converting the regression parameters into a full IPM), and does not attempt to help users with the parameterization task. This provides greater flexibility in modeling trait-demography relationships, and enables users to specify IPMs of any functional form that they desire.

Terminology and IPM construction

An IPM describes how the abundance and distribution of trait values (also called *state variables/states*, denoted z and z') for a population changes in discrete time. The distribution of trait values in a population at time t is given by the function $n(z,t)$. A simple IPM for the trait distribution z' at time $t+1$ is then

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

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

$K(z',z)$, known as the *projection kernel*, describes all possible transitions of existing individuals and recruitment of new individuals from t to $t+1$, generating a new trait distribution $n(z',t+1)$. L, U are the lower and upper bounds for values that the trait z can have, which defines the *domain* over which the integration is performed. The integral $\int_L^U n\,(z,t)dz$ gives the total population size at time t.

To make the model more biologically interpretable, the projection kernel $K(z',z)$ is usually decomposed into *sub-kernels* (Eq 2). For example, a projection kernel to describe a lifecycle where individuals can survive, transition to different state values, and reproduce via sexual and asexual pathways, can be decomposed as follows

$$K(z',z) = P(z',z) + F(z',z) + C(z',z), \quad (2)$$

where $P(z',z)$ is a sub-kernel describing transitions due to survival and trait changes of existing individuals, $F(z',z)$ is a sub-kernel describing per-capita sexual contributions of existing individuals to recruitment, and $C(z',z)$ is a sub-kernel describing per-capita asexual contributions of existing individuals to recruitment. The sub-kernels are typically comprised of functions derived from regression models that relate an individual's trait value z at time t to a new trait value z' at $t+1$. For example, the P kernel for Soay sheep (*Ovis aries*) on St. Kilda (Eq 3) may contain two regression models: (i) a logistic regression of survival on log body mass (Eq 4), and (ii) a linear regression of log body mass at $t+1$ on log body mass at t (Eq 5-6). In this example, f_g is a normal probability density function with μ_g given by the linear predictor of the mean, and with σ_g computed from the standard deviation of the residuals from the linear regression model.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

$$P(z', z) = s(z) * g(z', z), \quad (3)$$

$$Logit(s(z)) = \alpha_s + \beta_s * z, \quad (4)$$

$$g(z', z) = f_g(z', \mu_g, \sigma_g), \quad (5)$$

$$\mu_g = \alpha_g + \beta_g * z. \quad (6)$$

Analytical solutions to the integral in Eq 1 are usually not possible (Ellner & Rees, 2006). However, numerical approximations of these integrals can be constructed using a numerical integration rule. A commonly used rule is the midpoint rule (more complicated and precise methods are possible and will be implemented, though are not yet, see Ellner et al., 2016, Chapter 6). The midpoint rule divides the domain $[L, U]$ into m artifical size bins centered at z_i with width $h = (U - L)/m$. The midpoints $z_i = L + (i - 0.5) * h$ for $i = 1, 2, ..., m$. The midpoint rule approximation for Eq 1 then becomes:

$$n(z_j, t + 1) = h \sum_{i=1}^{m} K\,(z_j, z_i) n(z_i, t) \quad (7)$$

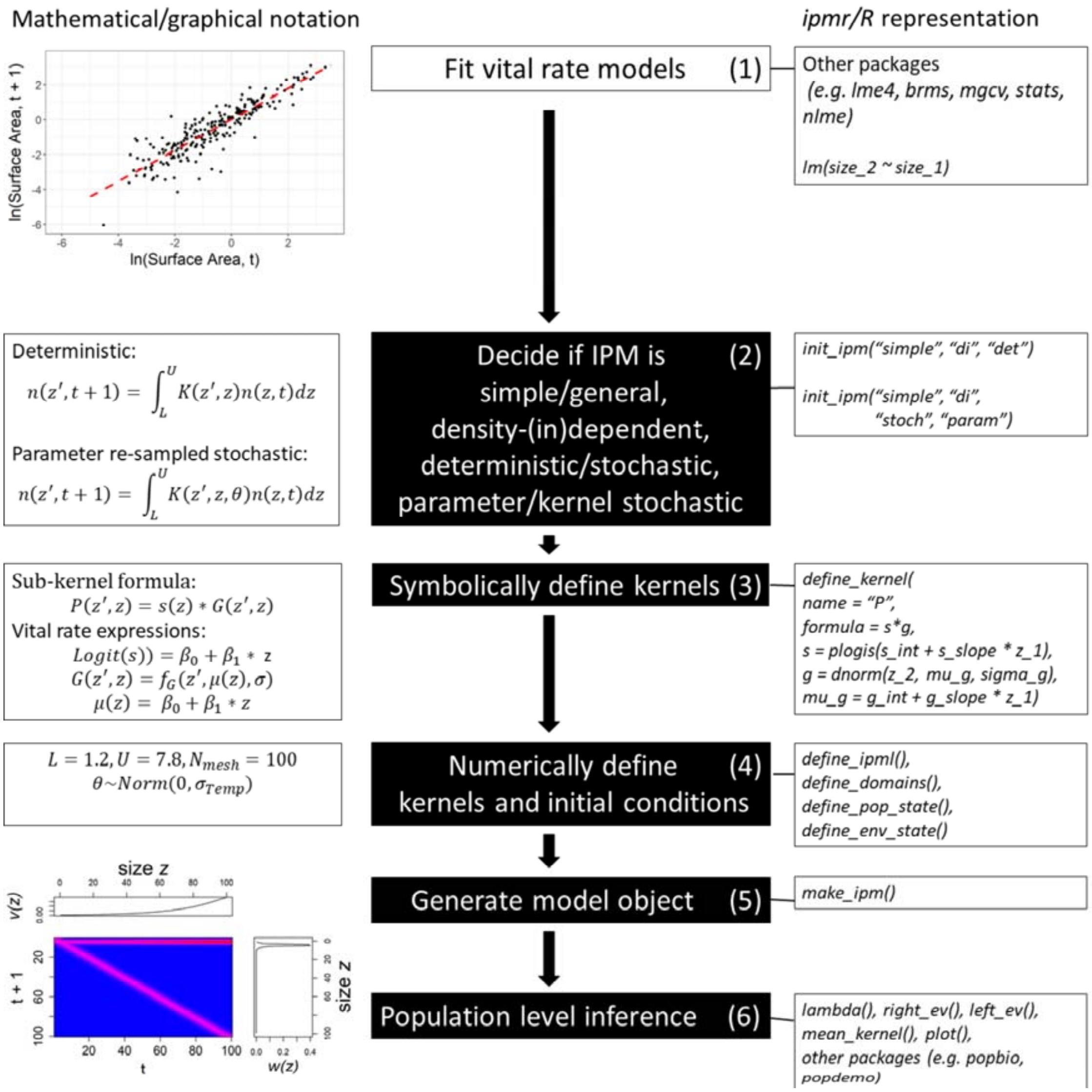
In practice, the numerical approximation of the integral converts the continuous projection kernel into a (large) discretized matrix. A matrix multiplication of the discretized projection kernel and the discretized trait distribution then generates a new trait distribution, a process referred to as *model iteration* (*sensu* Easterling et al., 2000).

Equations 1 and 2 are an example of a *simple IPM*. A critical aspect of `ipmr`'s functionality is the distinction between *simple IPMs* and *general IPMs*. A simple IPM incorporates a single continuous state variable. Equations 1 and 2 represent a simple IPM because there is only one continuous state, z, and no additional discrete states. A general IPM models one or

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

more continuous state variables, and/or discrete states. General IPMs are useful for modelling species with more complex life cycles. Many species' life cycles contain multiple life stages that are not readily described by a single state variable. Similarly, individuals with similar trait values may behave differently depending on environmental context. For example, Bruno et al. (2011) modeled aspergillosis impacts on sea fan coral (*Gorgonia ventalina*) population dynamics by creating a model where colonies were cross classified by tissue area (continuously distributed) and infection status (a discrete state with two levels - infected and uninfected). Coulson, Tuljapurkar & Childs (2010) constructed a model for Soay sheep where the population was structured by body weight (continuously distributed) and age (discrete state). Mixtures of multiple continuous and discrete states are also possible. Indeed, the vital rates of many species with complex life cycles are often best described with multivariate state distributions (Caswell & Salguero-Gómez, 2013). A complete definition of the simple/general distinction is given in Ellner et al. (2016, Chapter 6).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.


Mathematical/graphical notation
ipmr/R representation
ln(Surface Area, t + 1)
ln(Surface Area, t)
Fit vital rate models (1)
Other packages
(e.g. lme4, brms, mgcv, stats, nlme)
Deterministic:
$n(z',t+1) = \int_L^U K(z',z)n(z,t)dz$
Parameter re-sampled stochastic:
$n(z',t+1) = \int_L^U K(z',z,\theta)n(z,t)dz$
Decide if IPM is simple/general, density-(in)dependent, deterministic/stochastic, parameter/kernel stochastic (2)
Sub-kernel formula:
$P(z',z) = s(z) * G(z',z)$
Vital rate expressions:
$Logit(s)) = \beta_0 + \beta_1 * z$
$G(z',z) = f_G(z',\mu(z),\sigma)$
$\mu(z) = \beta_0 + \beta_1 * z$
Symbolically define kernels (3)
$L = 1.2, U = 7.8, N_{mesh} = 100$
$\theta \sim Norm(0, \sigma_{Temp})$
Numerically define kernels and initial conditions (4)
size z
v(z)
t + 1
t
w(z)
Generate model object (5)
Population level inference (6)
other packages (e.g. popbio, popdemo)

```
lm(size_2 ~ size_1)

init_ipm("simple", "di", "det")

init_ipm("simple", "di",
         "stoch", "param")

define_kernel(
name = "P",
formula = s*g,
s = plogis(s_int + s_slope * z_1),
g = dnorm(z_2, mu_g, sigma_g),
mu_g = g_int + g_slope * z_1)

define_ipml(),
define_domains(),
define_pop_state(),
define_env_state()

make_ipm()

lambda(), right_ev(), left_ev(),
mean_kernel(), plot(),
```

Figure 1: There are generally 6 steps in defining an IPM with `ipmr`. (1) Vital rate models are fit to demographic data collected from field sites. This step requires the use of other packages, as `ipmr` does not contain facilities for regression modeling. The figure on the left shows the fitted relationship between size at t and $t+1$ for Carpobrotus spp. in Case Study 1. (2) The next step is deciding what type of IPM is needed. This is determined by both the research question and the data used to parameterize the regression models. This process is initiated with `init_ipm()`. In step (3), kernels are defined using `ipmr`'s syntax to represent kernels and

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

vital rate functions. (4) Having defined symbolic representations of the model, the numerical definition is given. Here, the integration rule, domain bounds, and initial population conditions are defined. For some models, initial environmental conditions can also be defined. (5) `make_ipm()` numerically implements the `proto_ipm` object, (6) which can then be analyzed further. The figure at the bottom left shows a $K(z', z)$ kernel created by `make_ipm()` and `make_iter_kernel()`. The line plots above and to the right display the left and right eigenvectors, extracted with `left_ev()` and `right_ev()`, respectively.

Case study 1 - A simple IPM

One use for IPMs is to evaluate potential performance and management of invasive species in their non-native range (*e.g.* Erickson et al., 2017). Calculating sensitivities and elasticities of λ to kernel perturbations can help identify conservation management strategies (de Kroon et al., 1986, Caswell, 2001, Baxter et al., 2006, Ellner et al., 2016). Bogdan et al. (2020) constructed a simple IPM for a *Carpobrotus* species growing north of Tel Aviv, Israel. The model includes four regressions, and an estimated recruit size distribution. Table 1 provides the mathematical formulae, the corresponding R model formula, and the `ipmr` notation for each one. The case study materials also offer an alternative implementation that uses the generic `predict()` function to generate the same output. The final part of the case study provides examples of functions that compute kernel sensitivity and elasticity, the per-generation growth rate, and generation time for the model, as well as how to visualize these results.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Case study 2 - A general age × size IPM

We use an age- and size-structured IPM from Ellner et al. (2016) to illustrate how to create general IPMs with `ipmr`. This case study demonstrates is the suffix syntax for vital rate and kernel expressions, which is a key feature of `ipmr` (highlighted in bold in the 'ipmr' column in Table 2). The suffixes appended to each variable name in the `ipmr` formulation correspond to the sub- and/or super-scripts used in the mathematical formulation. `ipmr` internally expands the model expressions and substitutes the range of ages and/or grouping variables in for the suffixes. This allows users to specify their model in a way that closely mirrors its mathematical notation, and saves users from the potentially error-prone process of re-typing model definitions many times or using `for` loops over the range of discrete states. The case study then demonstrates how to couple age-specific survival and fertility with the model outputs.

Discussion of additional applications

We have shown above how `ipmr` handles a variety of model implementations that go beyond the capabilities of existing scripts and packages. The underlying implementation based on metaprogramming should be able to readily incorporate future developments in parameterization methods. Regression modeling is a field that is constantly introducing new methods. As long as these new methods have functional forms for their expected value (or a method for `predict()`), `ipmr` should be able to implement IPMs using them.

Finally, one particularly useful aspect of the package is the `proto_ipm` data structure. The `proto_ipm` is the common data structure used to represent every model class in `ipmr` and

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

provides a concise, standardized format for representing IPMs. Furthermore, the `proto_ipm` object is created without any raw data, only functional forms and parameters. We are in the process of creating the PADRINO IPM database using `ipmr` and `proto_ipms` as an “engine” to re-build published IPMs using only functional forms and parameter estimates. This database could act as an IPM equivalent of the popular COMPADRE and COMADRE matrix population model databases (Salguero-Gómez et al., 2015, Salguero-Gómez et al., 2016). Recent work has highlighted the power of syntheses that harness many structured population models (Adler et al., 2013, Salguero-Gómez et al., 2016, Compagnoni et al., 2020). Despite the wide variety of models that are currently published in the IPM literature, `ipmr`’s functional approach is able to reproduce nearly all of them without requiring any raw data at all.

Citation list

1. Adler, P.B., Ellner, S.P. & Levine, J.M. (2010). Coexistance of perennial plants: an embarassment of niches. Ecology Letters 13: 1019-1029. https://doi.org/10.1111/j.1461-0248.2010.01496.x

2. 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 111(2): 740-745. https://doi.org/10.1073/pnas.1315179111

3. Baxter, P.W.J., McCarthy, M.A., Possingham, H.P., Menkhorst, P.W. & McLean, N. (2006). Accounting for management costs in sensitivity analyses of matrix population models.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Conservation Biology 20(3): 893-905. https://doi.org/10.1111/j.1523-1739.2006.00378.x

4. Bogdan, A., Levin, S.C., Salguero-Gómez, R., Knight, T.M. (2020). Demographic analysis of Israeli Carpobrotus populations: management strategies and future directions. bioRxiv. https://doi.org/10.1101/2020.12.08.415174

5. Bruno, J.F., Ellner, S.P., Vu, I., Kim, K., & Harvell, C.D. (2011). Impacts of aspergillosis on sea fan coral demography: modeling a moving target. Ecological Monographs 81(1): 123-139. https://doi.org/19.1890/09-1178.1

6. Caswell, H. (2001) Matrix population models: construction, analysis, and interpretation, 2nd edn. Sunderland, MA: Sinauer Associates Inc

7. 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

8. Childs, D.Z., Rees, M., Rose, K.E., Grubb, P.J., & Ellner, S.P. (2004). Evolution of size-dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model. Proceedings of the Royal Society B 271(1547): 425-434. https://doi.org/10.1098/rpsb.2003.2597

9. Compagnoni, A., Levin, S.C., Childs, D.Z., Harpole, S., Paniw, M., Roemer, G., Burns, J.H., Che-Castaldo, J., Rueger, N., Kunstler, G., Bennett, J.M., Archer, C.R., Jones, O.R., Salguero-Gomez, R., & Knight, T.M. (2020). Short-lived plants have stronger demographic responses to climate. bioRxiv. https://doi.org/https://doi.org/10.1101/2020.06.18.160135

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

10. Coulson, T.N. (2012). Integral projection models, their construction and use in posing hypotheses in ecology. Oikos 121: 1337-1350. https://doi.org/10.1111/j.1600-0706.2012.00035.x

11. Coulson, T., Tuljapurkar, S., & Childs, D.Z. (2010). Using evolutionary demography to link life history theory, quantitative genetics and population ecology. Journal of Animal Ecology 79: 1226-1240. https://doi.org/10.1111/j.1365-2656.2010.01734.x

12. Crandall, R.M. & Knight, T.M. (2017). Role of multiple invasion mechanisms and their interaction in regulating the population dynamics of an exotic tree. Journal of Applied Ecology 55(2):885-894. https://doi.org/10.1111/1365-2664.13020

13. de Kroon, H., Plaisier, A., van Goenendael, J., & Caswell, H. (1986). Elasticity: the relative contribution of demographic parameters to population growth rate. Ecology 67(5): 1427-1431.

14. Easterling, M.R., Ellner, S.P., & Dixon, P.M. (2000). Size specific sensitivity: applying a new structured population model. Ecology 81(3): 694-708.

15. Ellner, S.P., Childs, D.Z., Rees, M. (2016) Data-driven modelling of structured populations: a practical guide to the integral projection model. Basel, Switzerland: Springer International Publishing AG

16. Ellner, S.P. & Rees, M. (2006). Integral Projection Models for species with complex demography. The American Naturalist 167(3): 410-428.

17. Erickson, R.A., Eager, E.A., Brey, M.B., Hansen, M.J., & Kocovsky, P.M. (2017). An integral projection model with YY-males and application to evaluating grass carp

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

control. Ecological Modelling 361: 14-25. https://doi.org/10.1016/j.ecolmodel.2017.07.030

18. Ferrer-Cervantes, M.E., Mendez-Gonzalez, M.E., Quintana-Ascencio, P-F., Dorantes, A., Dzib, G., & Duran, R. (2012). Population dynamics of the cactus *Mammillaria gaumeri*: an integral projection model approach. Population Ecology 54: 321-334. DOI: https://doi.org/10.1007/s10144-012-0308-7

19. Henry, L., & Wickham, H. (2020). rlang: Functions for Base Types and Core R and 'Tidyverse' Features. R package version 0.4.7. https://CRAN.R-project.org/package=rlang

20. Jongejans, E., Shea, K., Skarpaas, O., Kelly, D., & Ellner, S.P. (2011). Importance of individual and environmental variation for invasive species spread: a spatial integral projection model. Ecology 92(1): 86-97. https://doi.org/10.1890/09-2226.1

21. Merow, C., Dahlgren, J.P., Metcalf, C.J.E., Childs, D.Z., Evans, M.E.K., Jongejans, E., Record, S., Rees, M., Salguero-Gomez R., & McMahon, S.M. (2014). Advancing population ecology with integral projection models: a practical guide. Methods in Ecology and Evolution 5: 99-110. https://doi.org/10.1111/2041-210X.12146S

22. Metcalf, C.J.E., Ellner, S.P., Childs, D.Z., Salguero-Gómez, R., Merow, C., McMahon, S.M., Jongejans, E., & Rees, M. (2015). Statistical modelling of annual variation for inference on stochastic population dynamics using Integral Projection Models. Methods in Ecology and Evolution 6(9): 1007-1017. https://doi.org/10.1111/2041-210X.12405

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

23. Metcalf, C. J. E., McMahon, S. M., Salguero-Gómez, R. & Jongejans, E. (2013). IPMpack: an R package for integral projection models. Methods in Ecology and Evolution. 4(2): 195-200. https://doi.org/10.1111/2041-210x.12001

24. Ramula, S., Rees, M. & Buckley, Y. M. (2009). Integral projection models perform better for small demographic data sets than matrix population models: a case study of two perennial herbs. Journal of Applied Ecology 46(5): 1048-1053. https://doi.org/10.1111/j.1365-2664.2009.01706.x

25. Salguero-Gómez, R, Jones, O.R., Archer, C.A., Buckley, Y.M., Che-Castaldo, J., Caswell, C., Hodgson, D., Scheuerlein, A., Conde, D.A., Brinks, E., de Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Roemer, G., Runge, J., Ruoff, T., et al. (2014) The COMPADRE Plant Matrix Database: an online repository for plant population dynamics. Journal of Ecology 103: 202-218. https://doi.org/10.1111/1365-2745.12334

26. Salguero-Gómez, R., Jones, O.R., Archer, C.R., Bein, C., de Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Roemer, G., Ruoff, T., Sommer, V., Wille, J. Voigt, J., Zeh, S., Vieregg, D., Buckley, Y.M., Che-Castaldo, J., Hodgson, D., et al. (2016) COMADRE: a global database of animal demography. Journal of Animal Ecology 85: 371-384. https://doi.org/10.1111/1365-2656.12482

27. Shefferson, R.P., Kurokawa, S., & Ehrlen, J. (2020). LEFKO3: analysing individual history through size-classified matrix population models. Methods in Ecology and Evolution. https://doi.org/10.1111/2041-210X.13526

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

28. Williams, J.L., Miller, T.E.X., & Ellner, S.P. (2012). Avoiding unintentional eviction from integral projection models. Ecology 93(9): 2008-2014. https://doi.org/10.1890/11-2147.1

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 1: Translations between mathematical notation, R's formula notation, and ipmr's notation for the simplified version of Bogdan et al.'s Carpobrotus IPM. The ipmr column contains the expressions used in each kernel's definition. R expressions are not provided for sub-kernels and model iteration procedures because they typically require defining functions separately, and there are many ways to do this step (examples are in the R code for each case study in the appendix). The plogis() function may not be familiar to some users. It is from the 'stats' R package and computes the inverse logit transformation of an expression.

Math Formula	R Formula	ipmr
$\mu^g = \alpha^g + \beta^g * z$	`size_2 ~ size_1, family = gaussian()`	`mu_g = g_int + g_slope * z`
$g(z',z) = f^g(\mu^g,\sigma^g)$	`g = dnorm(z_2, mu_g, sd_g)`	`g = dnorm(z_2, mu_g, sd_g)`
$logit(s(z)) = \alpha^s + \beta^s * z$	`surv ~ size_1, family = binomial()`	`s = plogis(s_int + s_slope * z)`
$log(r^n(z)) = \alpha^{r^n} + \beta^{r^n} * z$	`fec ~ size_1, family = poisson()`	`r_n = exp(r_n_int + r_n_slope * z)`
$logit(r^p(z)) = \alpha^{r^p} + \beta^{r^p} * z$	`repr ~ size_1, family = binomial()`	`r_p = plogis(r_p_int + r_p_slope * z)`
$r^d(z') = f^{r^d}(\mu^{r^d},\sigma^{r^d})$	`dnorm(z_2, mu_f_d, sigma_f_d)`	`r_d = dnorm(z_2, f_d_mu, f_d_sigma)`
$p^r = \frac{\#Recruits(t+1)}{\#flowers(t)}$	`p_r = n_new_recruits / n_flowers`	`p_r = n_new / n_flowers`
$P = s(z) * g(z',z)$		`P = s * g`
$F(z',z) = r^p(z) * r^n(z) * r^d(z') * p^r$		`F = r_p * r_n * r_d * p_r`
$n(z',t+1) = \int_L^U [P(z',z) + F(z',z)]n(z,t)dz$		

Table 2: Translations between mathematical notation, R's formula notation, and ipmr's notation for Ellner et al. (2016) Ovis aries IPM. The ipmr column contains the expressions used in each kernel's definition. R expressions are not provided for sub-kernels and model iteration procedures because they typically require defining functions separately, and there are many ways to do this step

(examples are in the R code for each case study in the appendix). ipmr supports a suffix based syntax to avoid repetitively typing out the levels of discrete grouping variables. These are represented as 'a' in the Math column, 'age' in the R formula column, and are highlighted in bold in the ipmr column.

Math Formula	R Formula	ipmr
$Logit(s(z,a)) = \alpha^{s} + \beta_{z}^{s} * z + \beta_{a}^{s} * a$	`surv ~ size_1 + age, family = binomial()`	s_**age** = plogis(s_int + s_z * z_1 + s_a * **age**)
$g(z',z,a) = f^{g}(\mu_{a}^{g}, \sigma_{a}^{g})$	`g = dnorm(size_2, mu_g_age, sigma_g)`	g_**age** = dnorm(z_2, mu_g_**age**, sigma_g)
$\mu^{g}(z,a) = \alpha^{g} + \beta_{z}^{g} * z + \beta_{a}^{g} * a$	`size_2 ~ size_1 + age, family = gaussian()`	mu_g_**age** = g_int + g_z * z + g_a * **age**
$Logit(m^{p}(z,a)) = \alpha^{m^{p}} + \beta_{z}^{m^{p}} * z + \beta_{a}^{m^{p}} * a$	`repr ~ size_1 + age, family = binomial()`	m_p_**age** = plogis(m_p_int + m_p_z * z + m_p_a * **age**)
$Logit(r^{p}(a)) = \alpha^{r^{p}} + \beta_{a}^{r^{p}} * a$	`recr ~ age, family = binomial()`	r_p_**age** = plogis(r_p_int + r_p_a * **age**)
$\mu^{b} = \alpha^{b} + \beta_{z}^{b} * z$	`rc_size_2 ~ size_1, family = gaussian()`	mu_rc_size = rc_size_int + rc_size_z * z
$b(z',z) = f^{b}(\mu^{b}, \sigma^{b})$	`b = dnorm(size_2, mu_rc_size, sigma_rc_size)`	rc_size = dnorm(z_2, mu_rc_size, sigma_rc_size)
$P_{a}(z',z) = s(z,a) * g(z',z,a)$		P_**age** = s_**age** * g_**age** * d_z
$F_{a}(z',z) = s(z,a) * m^{p}(z,a) * r^{p}(a) * b(z',z)/2$		F_**age** = s_**age** * f_p_**age** * r_p_**age** * rc_size / 2

$$n_{0}(z',t+1) = \sum_{a=0}^{M+1} \int_{L}^{U} F_{a}(z',z) n_{a}(z,t) dz$$

$$n_{a}(z',t+1) = \int_{L}^{U} P_{a-1}(z',z) n_{a-1}(z,t) dz$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

$$n_{M+1}(z', t+1) = \int_{L}^{U} [\, P_{M+1}(z', z) n_{M+1}(z, t) + P_{M}(z', z) n_{M}(z, t)] dz$$

Author Contributions: All authors contributed to package design. SCL implemented the package. All authors wrote the first draft of the manuscript.

Funding: R.S-G. was supported by a NERC Independent Research Fellowship (NE/M018458/1). SCL, AC, SE, and TMK were funded by the Alexander von Humboldt Foundation in the framework of the Alexander von Humboldt Professorship of TM Knight endowed by the German Federal Ministry of Education and Research.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.20.440590; this version posted April 21, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.