



Syddansk Universitet

COMADRE

a global database of animal demography

Salguero-Gómez, Roberto; Jones, Owen; Archer, C. Ruth; Bein, Christoph; de Buhr, Hendrik; Farack, Claudia; Gottschalk, France; Hartmann, Alexander; Henning, Anne; Hoppe, Gabriel; Römer, Gesa; Ruoff, Tara; Sommer, Veronika; Wille, Julia; Voigt, Jakob; Zeh, Stefan; Vieregg, Dirk; Buckley, Yvonne M; Che-Castaldo, Judy; Hodgson, David; Scheuerlein, Alexander; Caswell, Hal; Vaupel, James W.

Published in:
BioRxiv

DOI:
[10.1101/027821](https://doi.org/10.1101/027821)

Publication date:
2016

Document version
Early version, also known as pre-print

Document license
CC BY-NC-ND

Citation for published version (APA):
Salguero-Gómez, R., Jones, O., Archer, C. R., Bein, C., de Buhr, H., Farack, C., ... Vaupel, J. W. (2016). COMADRE: a global database of animal demography. BioRxiv. DOI: 10.1101/027821

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 09. Sep. 2018

COMADRE: a global database of animal demography

Roberto Salguero-Gómez^{1,2*,§}, Owen R. Jones^{3,4,§}, C. Ruth Archer^{1,5}, Christoph Bein¹, Hendrik de Buhr¹, Claudia Farack¹, Fränze Gottschalk¹, Alexander Hartmann¹, Anne Henning¹, Gabriel Hoppe¹, Gesa Römer¹, Tara Ruoff⁶, Veronika Sommer¹, Julia Wille¹, Jakob Voigt¹, Stefan Zeh¹, Dirk Vieregg¹, Yvonne M. Buckley^{2,7}, Judy Che-Castaldo⁸, David Hodgson⁹, Alexander Scheuerlein¹, Hal Caswell^{10,11}, & James W. Vaupel^{1,3,12}

¹Laboratory of Evolutionary Biodemography Laboratory, Max Planck Institute for Demographic Research. Konrad-Zuse-Straße 1, Rostock, DE-18057, Germany.

²ARC Centre of Excellence for Environmental Decisions, School of Biological Sciences, The University of Queensland. Goddard building #8, St. Lucia, Queensland 4072, Australia.

³Max-Planck Odense Center on the Biodemography of Aging, University of Southern Denmark, Odense, Denmark.

⁴Department of Biology, University of Southern Denmark, Odense, Denmark.

⁵MaxNetAging School, Max Planck Institute for Demographic Research. Konrad-Zuse-Straße 1, DE-18057 Rostock, Germany.

⁶Department of Plant Science and Landscape Architecture and Department of Entomology, University of Maryland. College Park, MD 20742 USA.

⁷School of Natural Sciences, Zoology & Trinity Centre for Biodiversity Research, Trinity College Dublin, Dublin 2, Ireland.

⁸National Socio-Environmental Synthesis Center. 1 Park Place, Annapolis, MD 21401, USA.

⁹Centre for Ecology and Conservation. College of Life and Environmental Sciences. University of Exeter, Cornwall Campus TR10 9FE, UK.

¹⁰Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1090 GE Amsterdam, The Netherlands.

¹¹Biology Department MS-34, Woods Hole Oceanographic Institution. Woods Hole, MA 02543-1050, USA.

¹²Population Research Institute. Duke University. Durham, NC 27708-0309, USA.

* Corresponding author: salguero@demogr.mpg.de; compadre-contact@demogr.mpg.de

§ Equal contribution

Summary

1. The open-access scientific philosophy has been widely adopted and proven to promote considerable progress in the fields of ecology and evolution. Open-access global databases now exist on animal migration, the distribution of species, and conservation status, to mention a few. However, a gap exists for databases on population dynamics spanning the rich diversity of the animal kingdom. This information is fundamental to our understanding of the conditions that have shaped variation in animal life histories and their relationships with the environment. Furthermore, an animal population's schedules of reproduction and mortality determine its invasive potential, and its risk of local extinction, which are at the core of conservation biology.
2. Matrix population models (MPMs) are among the most widely used demographic tools by animal ecologists. MPMs project population dynamics in terms of reproduction, mortality, and development over the entire life cycle. The results of MPMs have direct biological interpretations, facilitating comparisons among animal species as different as *Caenorhabditis elegans*, *Loxodonta africana* and *Homo sapiens*.
3. Thousand of animal demographic records exist in the form of MPMs, but they are dispersed throughout the literature, rendering comparative analyses difficult. Here, we introduce the COMADRE Animal Matrix Database version 1.0.0, an open-source online repository containing data on 402 species worldwide, from 272 studies, with a total of 1,575 population projection matrices. COMADRE also contains ancillary information (e.g. ecoregion, taxonomy, biogeography, etc.) that facilitates interpretation of the numerous demographic metrics that can be derived from its MPMs.

4. *Synthesis*: We introduce the COMADRE Animal Matrix Database, a resource for animal demography. The open access nature of this database, together with its ancillary information will facilitate comparative analysis, as will the growing availability of databases focusing on other traits, and tools to query and combine them. Through future frequent updates of COMADRE, and its integration with other online resources, we encourage animal ecologists to tackle global ecological and evolutionary questions with unprecedented statistical power.

Keywords: big data; comparative approach; matrix population model; animal population ecology; population growth rate; open access.

Introduction

An understanding of the drivers and consequences of variation in reproduction and mortality throughout the life cycle is at the heart of population biology, evolution, ecology, and allied fields (Metcalf & Pavard 2006; Salguero-Gómez & de Kroon 2010; Salguero-Gómez *et al.* 2015). Although demography is essential to understand and predict population dynamics, no single repository integrates these data. This is mainly because most biological data sources are scattered and biological data types are heterogeneous (Hoffmann *et al.* 2014). Moreover, demographic data pose challenges for standardization due to the different formats and terminology (Lebreton 2012; Conde *et al.* unpublished). This makes it challenging to create a single demographic data repository across multiple species. However, there are important efforts towards compiling these data such as the Global Population Dynamics Database (GPDD, Inchausti & Halley 2001) and the Living Planet Index (LPI, Collen *et al.* 2009) holding population time-series data, BIDDABA (Lebreton *et al.* 2012), and the Primate Life History Database (PLHD, Strier *et al.* 2010) containing demographic information for birds and primates respectively. Although these examples have advanced the field of population biology, they are limited in either demographic detail (GPDD, LPI) or taxonomic scope (BIDDABA, PLHD, WBI).

A mechanistic understanding of how and why populations invade, grow, decline, or go locally extinct, requires data and methods that provide insights into age/size/ontogeny-based structure, such as Matrix Population Models (*MPMs* hereafter; Caswell 2001). *MPMs* have become the staple method describing the structured demography of animal populations. The widespread use of *MPMs* stems from their well-understood mathematical foundations and tractability (Caswell 2001),

coupled with the clear biological interpretations of the analytical outputs (de Kroon *et al.* 1986; Silvertown, Franco & Menges 1996; de Kroon, van Groenendael & Ehrén 2000). Briefly, an MPM divides the life cycle into discrete stages and projects the population through time in terms of probabilities of survival and of transitions among stages, and of the contributions to sexual or clonal reproduction at each stage. The stages of the life cycle can be chosen based on a compromise between the biology of the species and the availability of data, and the projection interval can vary from days (e.g. Buston & García 2007) to years (e.g. Edmunds *et al.* 2015), depending on the species and question.

As is the case with plants (Salguero-Gómez *et al.* 2015), a large number of MPMs have been published on species in the animal kingdom since the models were introduced in the 1940s (Bernadelli 1941; Leslie 1945) (Figure 1). Underlining the general utility of MPMs, these models have been used to address diverse topics including conservation biology (e.g., Crouse, Crowder & Caswell 1987; Jenouvrier *et al.* 2012), evolutionary biology (e.g., Kawecki 1995), ecotoxicology (e.g., Charles *et al.* 2009), invasion biology (e.g., Neubert & Parker 2004), and resource management (e.g., Salomon *et al.* 2013). MPMs have been employed to study species as taxonomically distinct as *Caenorhabditis elegans*, *Loxodonta africana* and *Homo sapiens*, and in geographically diverse regions with studies in every major biome (Figure 2.A & 2.B).

Despite the growing availability of published MPMs and the fact that such models are inherently comparable, there have been few attempts to use MPMs in comparative analyses. Notable exceptions are the work by Sæther and Backer (2000) on birds, and Heppell, Caswell and Crowder (2000) on mammals, Vélez-Espino, Fox and McLaughlin (2006) on fish, and van de Kerk *et al.* (2013) on order

Carnivora. These works illustrate the power of comparative approaches for robust generalizations by relating demographic estimates made from MPMs to interactions with the environment that form the basis for the evolution of life histories. One reason for the lack of comparative studies has historically been the paucity of readily available demographic data, compared to genetic data (e.g. Benson *et al.* 2013). This changed earlier this year, when Salguero-Gómez and colleagues (2015) released a database on plant demography, COMPADRE. Since its foundation in 1990, COMPADRE has prompted over 35 comparative plant demography studies ranging from senescence (Silvertown, Franco & Perez-Ishiwara 2001), to short-term population dynamics (Stott, Townley & Hodgson 2011), to the link between functional traits and demography (Adler *et al.* 2014). Here, we announce the release of COMPADRE's sister database, COMADRE, which contains MPMs and associated metadata from the animal kingdom.

The main objectives of the COMADRE team are (i) to find, digitize, and systematically error-check published animal MPMs and supplement them with additional information (Table 1), (ii) to offer such information on an open access basis, and (iii) to develop *R* scripts to facilitate comparative analyses. The data described here are available at www.comadre-db.org. In this paper, we briefly describe COMADRE, and highlight the major differences and similarities between COMADRE and the sister-database on plants COMPADRE. In addition, we briefly report some geographic, taxonomic and modelling biases inherent in the database. Finally, we detail our vision for how COMADRE will expand and develop in the future, linking to other already existing open-access databases to address timely questions in animal ecology and evolution.

COMADRE 1.0.0: A historical perspective

The accumulated number of publications reporting MPMs for animals has increased dramatically since MPMs were introduced in the 1940s (Figure 1).

Important contributions to this history come from the introduction of new *types* of MPMs, and new *methods* for analysing them.

Matrix population models were largely ignored for twenty years after the work of Leslie (1945). This is partly because Leslie had also helped introduce life table calculations of population growth rate into ecology, and those methods were more computationally feasible in the days before computers (Caswell 2001)¹.

The rediscovery of MPMs in the 1960s can be credited to three papers (Keyfitz 1964; Lefkovitch 1965; Rogers 1966). All of these papers focused on animals (yes, humans *are* animals). Keyfitz (1964) presented MPMs as tools for projecting population growth; his book (Keyfitz 1968) influenced a generation of animal ecologists. The first presentations of MPMs had assumed that age was the only *i*-state variable. Lefkovitch (1965), based on studies of laboratory populations of stored product insect pests, explicitly proposed stage-classified models based on the stages of the insect life cycle. Rogers (1966) introduced spatial, or multiregional, models for human populations, classifying individuals by age and spatial location, and modelling mortality, fertility, and migration between locations.

Other types of MPMs were introduced in the following years. The first seasonal, periodic MPM appeared in 1964 (Darwin & Williams 1964) in a study of seasonal harvesting as a control strategy for rabbits. The first density-dependent models appeared in 1969; Pennycuick *et al.* (1969) analyzed a population of great

¹ A glance at Hutchinson's (1978) population ecology text, based on a course he taught for many years at Yale, will show how influential life table methods were, and how intimately connected the approaches of animal demographers were to those of human demographers.

tits (*Parus major*), based on field data. Rabinovich (1969) comparing several density-dependent models, including a MPM to analyze laboratory populations of a parasitoid wasp. The first stochastic model for an animal population was the analysis by Cohen, Christensen and Goodyear (1983) of recruitment fluctuations in striped bass (*Morone saxatilis*). Invasion models, using matrix integrodifference equations, were first applied to bird populations by Caswell, Lensink and Neubert (2003).

Analytical methods have developed in parallel with their applications to animal populations. Many of these are listed in Figure 1. Some of these developments have provided new ways of constructing models (photo-identification methods, mark-recapture methods, vec-permutation matrix methods). Others have provided ways to extract additional information from the resulting MPM (sensitivity and elasticity analyses; LTRE decomposition analyses; stability and bifurcation analyses for nonlinear models; Markov chain methods for analysis of longevity, heterogeneity, and individual stochasticity; reactivity and amplification analyses). The introduction of new methods is not slowing down; if anything it is accelerating.

The COMADRE Animal Matrix Database was founded at the Max Planck Institute for Demographic Research (MPIDR) by Salguero-Gómez in 2011 (Figure 1), soon after joined by Jones, as well as a core committee, a science committee, and a team of digitizers (Supporting Information Appendix S1). The motivation for the creation of a database containing MPMs for animals was based on the success of its sister database, the COMPADRE Plant Matrix Database (Salguero-Gómez *et al.* 2015). Four years after its foundation, the COMADRE digitalization team has digitized, standardized, error-checked and supplemented information contained in over 400 species. As with the commitment for COMPADRE, more data will be

released periodically (Figure 1) through the COMADRE online portal (www.comadre-db.org).

What is in the COMADRE portal?

The COMADRE portal (www.comadre-db.org) facilitates open access to the *R* data object that contains the database itself, as well as the COMADRE user's guide. The latter contains details on the organization of the data object, the meaning and possible values for the variables within, and information on error-checks and quality controls. Additionally, Frequently Asked Questions (FAQs) can be found in the online portal (<http://www.comadre-db.org/Comadre/Help>).

The basic data item in COMADRE is the population projection matrix. A basic (*i.e.*, linear and time-invariant) MPM can be written

$$\mathbf{n}(t+1) = \mathbf{A} \mathbf{n}(t) \quad \text{eqn 1}$$

where \mathbf{n} is a vector giving the abundance of a set of age/size/ontogenetic classes and \mathbf{A} is a population projection matrix. The structure of the projection matrix \mathbf{A} depends on the choice of life cycle stages and the projection interval.

In COMADRE, the projection matrix is decomposed as

$$\mathbf{A} = \mathbf{U} + \mathbf{F} + \mathbf{C} \quad \text{eqn 2}$$

where \mathbf{U} is the matrix describing transitions and survival of extant individuals, and \mathbf{F} and \mathbf{C} are the matrices describing production of new individuals by sexual and clonal reproduction, respectively. Some studies do not measure reproduction, reporting only the transition matrix \mathbf{U} . In these cases, this is reflected in the variable *MatrixFec* (see Table 1 and COMADRE User's Guide for details). The column sums of \mathbf{U} give the survival probabilities of the stages, and thus should not exceed 1. Some studies

report **U** matrices whose column sums do exceed 1; these are noted in the database (variable *SurvivalIssue* in Table 1) and must be treated with caution.

The simple model (1) can be extended in several ways. *Seasonal* MPMs divide the year into seasons (not necessarily of the same length) and report a projection matrix \mathbf{A}_i for season i ; the database entries for such seasonal models report all of the seasonal matrices. *Stochastic, density-dependent, and environment-dependent* MPMs are increasingly common in animal studies. In such cases, the MPM can be written

$$\mathbf{n}(t+1) = \mathbf{A}[t, \mathbf{n}(t), \mathbf{E}(t)] \mathbf{n}(t) \quad \text{eqn 3}$$

where $\mathbf{E}(t)$ is some measure of environmental conditions. Such a model is associated not with a single projection matrix, but with a function that, given a time and/or environment and/or population vector, returns a projection matrix. Because such functions require a different data structure, such MPMs are not included in COMADRE 1.0.0, but we will include them in future versions.

Associated with the projection matrices is a rich set of descriptive information and metadata; thus the *R* object *COMADRE_v.1.0.0.Rdata* contains three main branches: *metadata*, *matrixClass* and *mat*. The metadata can be accessed in *R* with the command `comadre$metadata`, and it includes information about taxonomy, additional details of the study including its source, geo-location, and some details about the specific MPMs (Variables 1 through 54 in Table 1). Information about the classes used to construct the specific MPM can be accessed with the *R* command `comadre$matrixClass`. Lastly, the population projection matrices can be retrieved with the command `comadre$mat`. Data pertaining to particular matrices can thus be obtained using *R*'s data indexing facilities i.e. `comadre$metadata[n,`

`] and comadre$matrixClass[[n]]` will return the metadata and class information pertaining to the n th matrix (`comadre$mat[[n]]`).

In some cases, the original data source provided information that allowed us to split the full life-cycle matrix ($matA$) into survival-dependent processes ($matU$), sexual reproduction ($matF$), and clonal reproduction, ($matC$) as described in equation (2); see variable *MatrixSplit* in Table 1. These matrices can be obtained with ease: `comadre$mat[[n]]$matA`, `comadre$mat[[n]]$matU` etc. Splitting the matrices in this way allows for faster, semi-automatic calculation of demographic output on hundreds of records in a few seconds. The unique matrix-specific indices allow the user to relate the information contained in the three branches of COMADRE with just a few lines of code (see examples in our GitHub repository, linked through Supplementary Online Material S3).

COMADRE and COMPADRE: similarities

The core data in both COMADRE and COMPADRE are the population projection matrices that make up MPMs. A comparison of Table 1 in this manuscript and Table 1 in the introduction to COMPADRE (Salguero-Gómez *et al.* 2015) reveals a number of similarities. Moreover, the data quality controls are the same for COMADRE and for COMPADRE. These were detailed in an earlier publication (Salguero-Gómez *et al.* 2015). Due to its importance, however, we emphasize the variable *SurvivalIssue* (Table 1). The stage-specific survival of any column sums of $matU$ must be a value between 0 and 1. Values greater than 1 render most analyses of survival and longevity impossible. When probabilities exceeded the error margin for rounding error and were considerably greater than 1, authors were contacted for clarification. In some cases (<13% of MPMs with this issue), these personal communications

have resulted in amendments from the originally published matrices, or in the re-assignment of proportions of each matrix element in *matA* to the submatrices *matU*, *matF* and *matC* (Table 1). MPMs with this concern are periodically checked and, when necessary, additional clarification is requested from the authors and stored in the variable “Observation” (Table 1). Currently, only 1.2% of the MPMs (19 out of the 1,575) in version 1.0.0 have at least one life stage with survival >1.

COMADRE and COMPADRE: differences

In spite of the similarities, animals pose some important differences that cannot be fully accommodated by the database framework of COMPADRE. The following there are key differences between the two databases:

- The variables *GrowthType*, *DicotMonocot* and *AngioGymno*, which are specific to plants, naturally do not exist in COMADRE.
- The variable *TPLVersion*, which identifies the taxonomic validity of plant names from The Plant List (<http://www.theplantlist.org>), is here substituted with CoLCheckOK, its analog for the animal kingdom via The Catalogue of Life (CoL) (<http://www.catalogueoflife.org>). This variable simply indicates whether or not (TRUE/FALSE) the taxonomy given in COMADRE has been validated at CoL. In just three cases the species were not present in CoL. In all other cases the taxonomic placement has been validated as correct.
- We have added the variable *MatrixFec*, which indicates whether the reproductive component (matrices **F** and/or **C**) of the matrix model is missing or not. Users are cautioned to carefully examine models in which the reproductive components are missing before using them for any demographic analyses that require the full life cycle (e.g. population growth rates λ and its

elasticities/sensitivities, damping ratio ρ , etc.). However, other metrics are still valid in these models (e.g. life expectancy from *matU*; table 1). *MatrixFec* will also be added to the next version of COMPADRE.

- Unlike in COMPADRE, where we reconstructed a phylogeny for plant species, a phylogeny for most animal species in COMADRE has been recently published (Hedges *et al.* 2015). Furthermore, species-level resolved trees also exist for some taxonomic groups such as mammals (Bininda-Emonds *et al.* 2007), birds (Jetz *et al.* 2012), or reptiles (Pyron, Burbrink & Wiens 2013).

Scope and coverage of COMADRE

The current version in the COMADRE portal contains an unprecedented sample size for information on animal population dynamics: 1,575 MPMs from 272 studies corresponding to 402 species according to the authors, or 349 accepted taxonomically according to the Catalogue of Life. This represents a substantial improvement in sample size and ancillary information (Table 1) considered to date, including important comparative works examining various aspects of life history traits and population dynamics of mammals (50 species in Heppel, Caswell and Crowder 2000), birds (49 species in Sæther & Backer 2000), fish (88 species in Vélez-Espino, Fox & McLaughlin 2006), order Carnivora (285 species in van de Kerk *et al.* 2013) and various animal taxa (17 species in Jones *et al.* 2014). It must also be noted that the information analysed in those studies was not made publically available, although M. van de Kerk kindly provided COMADRE with her MPMs.

COMADRE offers a broad geographic coverage of animal population dynamics (Figure 2.A). Information in COMADRE 1.0.0 includes MPMs from all continents except Antarctica – although MPMs for Antarctic species do exist and will

be released in future version of COMADRE (e.g., Emperor penguin, Jenouvrier *et al.* 2012; Antarctic petrel, Descamps *et al.* 2015). Importantly, geographic gaps do exist in our knowledge of animal demography in certain regions, including Oceania (8.13% of MPMs), and Asia (2.3%; Figure 2.B). Together, the USA (31.7%), Canada (8.8%), Australia (5.2%), and Kenya (4.8%) comprise over half the MPMs in COMADRE 1.0.0, and a clear bias exists towards terrestrial studies at low elevations (Figure 2.C). Unfortunately, few studies report MPMs from biodiversity hotspots such as Honduras, Guatemala, the Democratic Republic of Congo, Paraguay, India, and Indonesia. Furthermore, even some developed countries, such as Saudi Arabia, Italy, Greece, Ireland, Brazil and France, are under-represented.

Individual and seasonal population projection matrices (Table 1 #42) together total over 50% (783) of the projection matrices in COMADRE 1.0.0, representing unique combinations of studies \times species \times populations \times treatments \times periods (Figure 3.A). The remaining 755 projection matrices are element-by-element arithmetic means of other matrices, or constructed based on data from multiple sources (“pooled”). Given the intra-annual (Figure 4.A), inter-annual (Figure 4.C) and spatial replication (Figure 4.B) in many studies, the high proportion of mean and pooled matrices suggests a tendency in animal demographic studies to publish only summary MPMs. We encourage authors to do so as part of the supplementary materials for their papers. Authors willing to share additional matrices to be archived in COMADRE can do so by submitting the materials at comadre-contact@demogr.mpg.de.

Most studies in COMADRE are of natural populations in the wild (87%, Figure 3.B), and under unmanipulated conditions (80%; Figure 3.C). Most of the demographic studies in COMADRE are based on females only (68%; Figure 3.D);

this is common practice in animal demographic studies (particularly in mammals), since quantifying reproduction is usually easier in females than in males. We have noted, in the variable “Observations” (Table 1 #52), when the primary sex ratio was stated by the author as differing from 1:1 (female:male). For the vast majority of matrices (97%; Figure 3.E), we have successfully split the full matrix **A** into its sub-components of survival (**U**), sexual reproduction (**F**) and clonal reproduction (**C**), and only 4% of the MPMs do not incorporate reproductive information (Figure 3.F).

The data in COMADRE 1.0.0 represent a wide range of animal groups (Figure 3.G). However, there are some strong taxonomic biases. Mammals represent 44.4% of the MPMs in the current version of COMADRE, followed by birds (17.9%), bony fish (10.3%), and reptiles (6.7%). We include very few MPMs for amphibians, despite global concerns for their conservation status (Beebee & Griffiths 2005; Wake & Vredenburg 2008) or for insects (2.71%), despite their high species richness, estimated to comprise the majority of the animal kingdom (Hedges *et al.* 2015). The latter is particularly surprising since the early developments of MPMs focused on insects thanks to their clearly structured population dynamics (Lefkovitch 1965; Rabinovich 1969). Aside from bony fish (Actinopterygii), we also lack significant amounts of demographic information on marine organisms in COMADRE, including corals (5.5%), bivalves (1.33%), sponges (0.8%), sea urchins (0.12%), and cartilaginous fish (0.2%). No information in COMADRE 1.0.0 exists for the infraclasses Marsupialia (kangaroos, wallabies, koala, possums, opossums, wombats, etc.) or order Struthioniformes (kiwis, emu, ostriches, etc.), neither in the additional 800 species records that are currently being digitized and error-checked for future version releases.

The replication of studies through time and space is highly variable in COMADRE. Yet, the average duration of studies in COMADRE 1.0.0 (15.69 years \pm 13.80 S.D.; Figure 4.C) is greater than in plant MPM studies (Salguero-Gómez et al. 2015). Long duration is essential for many demographic studies in the animal kingdom, as some animals, such as the clam *Arctica islandica*, giant tortoises (*Geochelone nigra*, *G. gigantea*), rockfish (*Sebastes* sp.), and the bowhead whale (*Balaena mysticetus*), can reach over 150 years of age (de Magalhaes & Costa 2009). Notable demographic studies using MPMs parameterized with long time-series include *Vipera aspis* (17 years, Altwegg et al. 2005), *Ursus americanus* (22 years, Mitchell et al. 2009), *Delphinus delphis* (35 years, Mannocci et al. 2012), *Recurvirostra avosetta* (40 years, Hill 1988), *Elatobium abietinum* (41 years, Estay et al. 2012), *Marmota flaviventris* (44 years, Ozgul et al. 2009), *Haliaeetus albicilla* (62 years, Krüger, Grünkorn & Struwe-Juhl 2010), *Diomedea exulans* (51 years, Barbraud et al. 2013), and *Aythya affinis* (72 years, Koons et al. 2006).

In contrast to the duration, the average number of populations considered in each study is low, averaging 3.31 \pm 5.26 (S.D.). The low spatial replication currently limits much-needed understanding of the geographic variability of demographic rates within species. Recently initiated efforts to increase spatial replication for certain species in the plant kingdom (e.g. on *Plantago lanceolata*, PlantPopNet, www.plantago.plantpopnet.com) are a useful model that could be replicated in future work on animals. It is perhaps not surprising that the animal studies with highest spatial replication in COMADRE 1.0.0 focus on humans, with the foundational archive of human MPMs compiled by Keyfitz and Flieger (1968), which covers populations from 156 countries. We note that analyses of spatial and other kinds of variability in animal population studies are becoming more sophisticated due to the

use of model selection methods to explicitly include environmental variables (e.g., Thomson, Cooch & Conroy 2009), and the concept of “spatial replication” used in plant studies may acquire a different meaning to that used in most animal studies.

Unlocking global analyses

“I’m not interested in your data; I’m interested in merging your data with other data.

Your data will never be as exciting as what I can merge it with”

Tim Berners-Lee

The scientific promise of the COMADRE Animal Matrix Database does not reside exclusively in its hundreds of MPMs, but also in the many outputs that can be derived from them, and the possibility to put them in a broader spatial, ecological and evolutionary context using other open access databases and user collected or compiled data. Users of COMADRE can find several *R* scripts available to manipulate and interact with matrices, and derive basic demographic outputs (Supporting Information Appendix S3, and our growing GitHub repository). Users are welcome to explore these or other more developed open-source libraries (Stubben & Milligan 2007; Stott, Hodgson & Townley 2012; Metcalf *et al.* 2013), and to carry out their own calculations based on methods for the analyses of MPMs (e.g., Caswell 2001; Morris & Doak 2002).

The schedules of growth, survival, and reproduction and the associated population performance metrics available through COMADRE will enable further comparative analyses of life history variation and population performance relative to the environment. For example, information in COMADRE can be integrated with existing repositories for other data such as genetic sequences (GenBank; Benson *et al.* 2013), distribution and occurrences (GBIF; Flemons *et al.* 2007), and

conservation status and threats (BirdLife, <http://www.birdlife.org/datazone/>; IUCN Red List, <http://www.iucnredlist.org>). Data on species-level life history traits are also available for specific taxonomic groups including vertebrates (AnAge; de Magalhaes & Costa 2009), mammals (Ernest 2003; PanTHERIA, Jones et al. 2009), amphibians (Trochet *et al.* 2014), fish (FishBase), and reptiles (SCALETOOL, www.scale-project.net). Lastly, an upcoming resource, DATLife (Scheuerlein *et al.* unpublished), containing age-specific mortality and ancillary data on animal species will be of particular interest to supplement COMADRE, as methods already exist to convert life tables into MPMs and *vice versa* (Caswell 2001). In addition to this rich and rapidly growing body of data, a diverse set of tools are emerging that will facilitate these large-scale comparative analyses including the R packages *taxize* (Chamberlain & Szöcs 2013), *letsR* (Vilela & Villalobos, 2015) which facilitate taxonomic matching and macroecological analyses respectively.

The compilation of demographic data in COMADRE will also enable the identification of gaps in our knowledge of animal population dynamics and, as is now happening for plants, will catalyze new studies at broad spatial scales. The open-access publication of both COMPADRE and COMADRE databases will facilitate further comparative demographic analyses across plant and animal kingdoms (see Jones *et al.* 2014) enabling tests of life history and population dynamics theory across a wide range of species with contrasting life histories. We suggest that researchers revisit the canonical tenets of animal life history to confront established theories with data compilations that are vastly richer than was available 30 years ago.

While the COMADRE team strives to make the database as accurate as possible, we make no claims, promises, or guarantees about the accuracy,

completeness, or adequacy of the database, and expressly disclaim liability for errors and omissions in the contents of the database. No warranty of any kind is given for any particular use of COMADRE. Users who detect apparent errors are encouraged to contact us at comadre-contact@demogr.mpg.de.

Although this first release contains 402 species, we have already identified over 800 additional animal species with MPMs, and our on-going efforts will release them as they become fully-digitized, error-checked and supplemented in the coming years (Salguero-Gómez *et al.* unpublished). Finally, researchers using the data archived here are encouraged to cite also the original sources in their works (Supporting Information Appendix S4).

We are extremely grateful to the many ecologists, zoologists, and evolutionary biologists who have made the projection matrices from their MPMs available for publication in this open access database. Some of the data stored here must rank among the most valuable (and most expensive to collect) biological information in existence. The providers of data have shared the vision of the COMADRE leaders, that important data should be made available to all interested parties for free. In return, we simply encourage demographers, conservation biologists, ecologists and evolutionary biologists worldwide to mine this database and find out as much as they can about global, phylogenetic and ecological patterns in animal life history and population dynamics.

Acknowledgments

The original author list whose works constitute the motivation for the COMADRE Animal Matrix Database version 1.0.0 can be found in the Supporting Information Appendix S4. The authors' willingness to share unpublished data and clarify questions on published materials is an inspiration for the new era of big, open, data. COMADRE is currently supported by the Evolutionary Demography Laboratory at the Max Planck Institute for Demographic Research (MPIDR). For a list of funding support since its inception see Supporting Information Appendix S5. Logistical and computational support was provided by the IT team of the MPIDR.

Data accessibility statement: The data associated with this manuscript can be accessed at www.comadre-db.org

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**, 740-745.
- Altwegg, R., Dummermuth, S., Anholt, B. R. & Flatt, T. (2005) Winter weather affects asp viper *Vipera aspis* population dynamics through susceptible juveniles. *Oikos*, **110**, 55-66.
- Barbraud, C., Tuck, G. N., Thomson, R., Delord, K. & Weimerskirch, H. (2013) Fisheries bycatch as an inadvertent human-induced evolutionary mechanism. *PLoS One*, **8**, e60353.
- Beebee, T. J. & Griffiths, R. A. (2005) The amphibian decline crisis: a watershed for conservation biology? *Biological Conservation*, **125**, 271-285.
- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. (2013) GenBank. *Nucleic Acids Research*, **41**, D36-D42.
- Bernardelli, H. (1941) Population waves. *Journal of the Burma Research Society*, **31**, 1-18.
- Bininda-Emonds, O. R., Cardillo, M., Jones, K. E., MacPhee, R. D., Beck, R. M., Grenyer, R., Price, S. A., Vos, R. A., Gittleman, J. L. & Purvis, A. (2007) The delayed rise of present-day mammals. *Nature*, **446**, 507-512.
- Brault, S., & Caswell, H. (1993) Pod-specific demography of killer whales (*Orcinus orca*). *Ecology*, **74**(5), 1444-1454.
- Buston, P. M. & García, M. B. (2007) An extraordinary life span estimate for the clown anemonefish *Amphiprion percula*. *Journal of Fish Biology*, **70**, 1710-1719.
- Caswell, H. (1989) *Matrix Population Models*. 1st edition. John Wiley & Sons, Ltd. Massachusetts, USA.
- 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), 1226-1238.

- Caswell, H. (2001). *Matrix Population Models*. 2nd edition. John Wiley & Sons, Ltd. Massachusetts, USA.
- Caswell, H. (2008) Perturbation analysis of nonlinear matrix population models. *Demographic Research*, **18**(3), 59-116.
- Caswell, H. (2009) Stage, age and individual stochasticity in demography. *Oikos*, **118**(12), 1763-1782.
- Caswell, H., Lensink, R., & Neubert, M. G. (2003) Demography and dispersal: life table response experiments for invasion speed. *Ecology*, **84**, 1968-1978.
- Caswell, H., & Neubert, M. G. (2005) Reactivity and transient dynamics of discrete-time ecological systems. *Journal of Difference Equations and Applications*, **11**(4-5), 295-310.
- Chamberlain, S. A. & Szöcs, E. (2013) taxize: taxonomic search and retrieval in R. *F1000Research*, **2**. doi: 10.12688/f1000research.2-191.v2
- Charles, S., Billoir, E., Lopes, C. & Chaumot A. (2009) Matrix population models as relevant modelling tools in ecotoxicology. *Ecotoxicology Modeling*, **2**, 261-298.
- Cohen, J. E., Christensen, S. W. & Goodyear, C. P. (1983) An age-structured fish population model with random survival of eggs: Calculation of asymptotic growth rates and application to Potomac River striped bass. *Canadian Journal of Fisheries and Aquatic Science*, **40**, 2170-2183.
- Collen B., Loh J., Whitmee S., McRae L., Amin R. & Baillie J.E.M. (2009) Monitoring change in vertebrate abundance: the Living Planet Index. *Conservation Biology*, **23**, 317-327.
- Crouse, D. T., Crowder, L. B. & Caswell, H. (1987) A stage-based population model for loggerhead sea turtles and implications for conservation ecology. *Ecology*, **68**, 1412-1423.
- Darwin, J. H. & Williams, R. M. (1964) The effect of time of hunting on the size of a rabbit population. *New Zealand Journal of Science*, **7**, 341-352.
- de Kroon, H., Plaisier, A., van Groenendael, J. & Caswell, H. (1986) Elasticity: the relative contribution of demographic parameters to population growth rate. *Ecology*, **67**, 1427-1431.

- de Kroon, H., van Groenendael, J. & Ehrlén, J. (2000) Elasticities: a review of methods and model limitations. *Ecology*, **81**, 607-618.
- de Magalhaes, J. P. & Costa, J. (2009) A database of vertebrate longevity records and their relation to other life-history traits. *Journal of Evolutionary Biology*, **22**, 1770-1774.
- Decamps, S., Tarroux, A., Varpe, Ø., Yoccoz, N. G., Tveraa, T. & Larentsen, S. H. (2015) Demographic effects of extreme weather events: snow storms, breeding success, and population growth rate in a long-lived Antarctic seabird. *Ecology and Evolution*, **5**, 314-325.
- Edmunds, P. J. (2015) A quarter-century demographic analysis of the Caribbean coral, *Orbicella annularis*, and projections of population size over the next century. *Limnology and Oceanography*, **60**, 840-855.
- Ernest, S. M. (2003) Life history characteristics of placental non-volant mammals: Ecological Archives E084-093. *Ecology*, **84**, 3402-3402.
- Estay, S. A., Lima, M., Labra, F. A. & Harrington, R. (2012) Increased outbreak frequency associated with changes in the dynamic behaviour of populations of two aphid species. *Oikos*, **121**, 614-622.
- Flemons, P., Guralnick, R., Krieger, J., Ranipeta, A. & Neufeld, D. (2007) A web-based GIS tool for exploring the world's biodiversity: The Global Biodiversity Information Facility Mapping and Analysis Portal Application (GBIF-MAPA). *Ecological Informatics*, **2**, 49-60.
- Hedges, S. B., Marin, J., Suleski, M., Paymer, M. & Kumar, S. (2015) Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution* X, XXX-XXX. doi: 10.1093/molbev/msv037
- Heppell, S., Caswell, H. & Crowder, L. B. (2000) Life histories and elasticity patterns: perturbation analysis for species with minimal demographic data. *Ecology*, **81**(3), 654-665.
- Heppell, S., Pfister, C. & de Kroon, H. (2000) Elasticity analysis in population biology: methods and applications. *Ecology*, **3**, 605-606.
- Hill, D. (1988) Population dynamics of the avocet (*Recurvirostra avosetta*) breeding in Britain. *Journal of Animal Ecology*, **2**, 669-683.

- Hoffmann A, Penner J, Vohland K, Cramer W, Doubleday R, Henle K, Kõljalg U, Kühn I, Kunin W, Negro J, Penev L, Rodríguez C, Saarenmaa H, Schmeller D, Stoev P, Sutherland W, Ó Tuama É, Wetzel F, Häuser C (2014) The need for an integrated biodiversity policy support process – Building the European contribution to a global Biodiversity Observation Network (EU BON). *Nature Conservation*, **6**: 49-65.
- Hutchinson, G.E. (1978) *An Introduction to Population Ecology*. Yale University Press, New Haven, USA.
- Inchausti, P. & Halley, J. (2001) Investigating long-term ecological variability using the global population dynamics database. *Science*, **293**, 655-657.
- Jenouvrier, S., Holland, M., Strøve, 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**, 2756-2770.
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. (2012) The global diversity of birds in space and time. *Nature*, **491**, 444-448.
- Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C., Connolly, C., Cutts, M. J., Foster, J. K., Grenyer, R., Habib, M., Plaster, C. A., Price, S. A., Rigby E. A., Rist, R., Teacher, A., Bininda-Emonds, O. R. P., Gittleman, J. L., Mace, G. M., Purvis, A. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, **90**, 2648-2648.
- Jones, O. R., Scheuerlein, A., Salguero-Gómez, R., Camarda, C. G., Schaible, R., Casper, B. B., Dahlgren, J. P., Ehrlén, J., García, M. P., 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**, 169-173.
- Kawecki, T. J. (1995) Demography of source-sink populations and the evolution of ecological niches. *Evolutionary Ecology*, **9**, 38-44.
- Keyfitz, N. (1964) The population projection as a matrix operator. *Demography*, **1**, 56-73.

- Keyfitz, N. (1968) *Introduction to the Mathematics of Population*. Addison-Wesley Pub. Co. California, USA.
- Keyfitz, N. & Flieger, W. (1968) *World Population: an Analysis of Vital Data*. University of Chicago Press. Chicago, USA.
- Koons, D. N., Rotella, J. J., Willey, D. W., Taper, M., Clark, R. G., Slattery, S., Brook, R. W., Corcoran, R. M. & Lovvorn, J. R. (2006) Lesser scaup population dynamics: what can be learned from available data? *Avian Conservation and Ecology*, **1**, art. 6.
- Krüger, O., Grünkorn, T. & Struwe-Juhl, B. (2010) The return of the white-tailed eagle (*Haliaeetus albicilla*) to northern Germany: Modelling the past to predict the future. *Biological Conservation*, **143**, 710-721.
- Lebreton, J. D., Devillard, S., Popy, S., Desprez, M., Besnard, A. & Gaillard, J. M. (2012) Towards a vertebrate demographic data bank. *Journal of Ornithology*, **152**, 617-624.
- Lefkovitch, L. P. (1965) The study of population growth in organisms grouped by stages. *Biometrika*, **35**, 183-212.
- Leslie, P. H. (1945) The use of matrices in certain population mathematics. *Biometrika*, **33**, 183-212.
- Levin, S. A., & Goodyear, C. P. (1980) Analysis of an age-structured fishery model. *Journal of Mathematical Biology*, **9**(3), 245-274.
- Mannocci, L., Dabin, W., Augeraud-Véron, E., Dupuy, J. F., Barbraud, C. & Ridoux, V. (2012) Assessing the impact of by-catch on dolphin populations: the case of the common dolphin in the eastern North Atlantic. *PLoS One*, **7**, e32615.
- Metcalf, C. J. E. & Pavard, S. (2007) Why evolutionary biologists should be demographers. *Trends in Ecology & Evolution*, **22**, 205-212.
- 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**, 195-200.

- Mitchell, M. S., Pacifici, L. B., Grand, J. B. & Powell, R. A. (2009) Contributions of vital rates to growth of a protected population of American black bears. *Ursus*, **20**, 77-84.
- Morris, W. F. & Doak, D. F. (2002) *Quantitative Conservation Biology*. Sinauer, Sunderland, Massachusetts, USA.
- Neubert, M. G. & Parker, I. M. (2004) Projecting rates of spread for invasive species. *Risk Analysis*, **24**, 817-831.
- Nichols, J. D., Sauer, J. R., Pollock, K. H., & Hestbeck, J. B. (1992) Estimating transition probabilities for stage-based population projection matrices using capture-recapture data. *Ecology*, **73**(1), 306-312.
- Ozgul, A., Oli, M. K., Armitage, K. B., Blumstein, D. T. & van Vuren, D. H. (2009) Influence of local demography on asymptotic and transient dynamics of a yellow-bellied marmot metapopulation. *The American Naturalist*, **173**, 517-530.
- Pennycuik, L. (1969) A computer model of the Oxford great tit population. *Journal of Theoretical Biology*, **22**, 381-400.
- Pyron, R. A., Burbrink, F. T. & Wiens, J. J. (2013) A phylogeny and revised classification of *Squamata*, including 4161 species of lizards and snakes. *BMC evolutionary biology*, **13**, 93.
- Rabinovich, J. E. (1969) The applicability of some population growth models to a single-species laboratory population. *Annals of the Entomological Society of America*, **62**, 437-442.
- Rogers, A. (1966) Matrix methods of population analysis. *Journal of the American Institute of Planners*, **32**, 40-44.
- Salguero-Gómez, R. & de Kroon, H. (2010) Matrix projection models meet variation in the real world. *Journal of Ecology*, **98**, 250-254.
- 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., Wille, J., Zeh, S., Davison, R., Vieregg, D., Baudisch, A., Altwegg, R., Colchero, F., Dong, M., de Kroon, H., Lebreton, J.-D., Metcalf, C. J. E., Neel, M. M., Parker, I. M., Takada, T., Valverde, T., Vélez-Espino, L.

- A., Wardle, G. M., Franco, M., Vaupel, J. W. (2015) The COMPADRE Plant Matrix Database: an open online repository for plant demography. *Journal of Ecology*, **103**, 202-218.
- Sæther, B. E. & Bakke, Ø. (2000) Avian life history variation and contribution of demographic traits to the population growth rate. *Ecology*, **81**, 642-653.
- Salomon, Y., McCarthy, M. A., Taylor, P. & Wintle, B. A. (2013) Incorporating uncertainty of management costs in sensitivity analyses of matrix population models. *Conservation Biology*, **27**, 134-144.
- Silvertown, J., Franco, M. & Menges, E. (1996) Interpretation of elasticity matrices as an aid to the management of plant populations for conservation. *Conservation Biology*, **10**, 591-597.
- Silvertown, J., Franco, M. & Perez-Ishiwara, R. (2001) Evolution of senescence in iteroparous perennial plants. *Evolutionary Ecology Research*, **3**, 393-412.
- Stott, I., Hodgson, D. J. & Townley, S. (2012) popdemo: an R package for population demography using projection matrix analysis. *Methods in Ecology and Evolution*, **3**, 797-802.
- Stott, I., Townley, S. & Hodgson, D. J. (2011) A framework for studying transient dynamics of population projection matrix models. *Ecology Letters*, **14**, 959-970.
- Strier, K. B., Altmann, J., Brockman, D. K., Bronikowski, A. M., Cords, M., Fedigan, L. M., Lapp, H., Liu, X., Morris, W. F., Pusey, A. E., Stoinski, T. S. & Alberts, S. C. (2010) The Primate Life History Database: a unique shared ecological data resource. *Methods in Ecology and Evolution*, **1**, 199-211.
- Stubben, C. & Milligan, B. (2007) Estimating and analyzing demographic models using the popbio package in R. *Journal of Statistical Software*, **22**, 1-23.
- Thomson, D. L., Cooch, E. G. & Conroy, M. J. (2009) *Modeling Demographic Processes in Marked Populations*. Springer-Verlag, New York, USA.
- Trochet, A., Moulherat, S., Calvez, O., Stevens, V. M., Clobert, J. & Schmitter, D. S. (2014) A database of life-history traits of European amphibians. *Biodiversity Data Journal*, **2**, e4123.

- Tuljapurkar, S. (1990) *Population dynamics in variable environments*. Springer, Berlin. Germany.
- Tuljapurkar, S., Steiner, U. K., & Orzack, S. H. (2009) Dynamic heterogeneity in life histories. *Ecology Letters*, **12**(1), 93-106.
- van de Kerk, M., de Kroon, H., Conde, D. A. & Jongejans, E. (2013) Carnivora population dynamics are as slow and as fast as those of other mammals: implications for their conservation. *PLoS One*, **8**, e70354.
- Vilela, B. & Villalobos, F. (2015) letsR: a new R package for data handling and analysis in macroecology. *Methods in Ecology and Evolution*, **X**, XXX-XXX. doi: 10.1111/2041-210X.12401
- Vélez-Espino, L. A., Fox, M. G. & McLaughlin, R. L. (2006) Characterization of elasticity patterns of North American freshwater fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, **63**, 2050-2066.
- Wake, D.B. & Vredenburg, V.T. (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11466-11473.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Constituents of COMADRE.

Appendix S2. COMADRE user's guide.

Appendix S3. COMADRE *R* scripts.

Appendix S4. Extended literature used in COMADRE 1.0.0.

Appendix S5. Funding and extended acknowledgements.

Appendix S6. Author contributions.

Appendix S7. Supporting information references.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

1 **Table 1.** Variables in the COMADRE Animal Matrix Database, organized by six general aspects: taxonomy, source, details of
 2 study, geolocation, and Matrix Population Model (MPM). A more detailed description can be found in the user protocol of
 3 COMADRE at www.comadre-db.org. ⌘ implies information that is study-specific; ⌘ information that is MPM-specific. Variables 1-54
 4 are archived in `comadre$metadata`, variables 55-57 in `comadre$matrixClass`, and variables 58-61 in `comadre$mat` in the
 5 COMADRE *R* data object open access available in the COMADRE online portal.

6

Aspect	Variable	Description
Taxonomy ⌘	1. <i>SpeciesAuthor</i>	Taxonomic species name as used by the author(s) in the publication. When more than one study exist for the same species, these are given sequential numeric suffixes (e.g. <i>Ursus_americanus</i> , <i>Ursus_americanus_2</i> , etc.)
	2. <i>SpeciesAccepted</i>	Currently accepted taxonomic name according to the Catalogue of Life (www.catalogueoflife.org). See the Supplementary Online Material S3 for an <i>R</i> script to check accepted and synonym names from <i>SpeciesAuthor</i> above
	3. <i>CommonName</i>	English common name of <i>SpeciesAccepted</i>
	4. <i>CoLCheckOK</i>	Whether the taxonomy detailed here has been verified at the Catalogue of Life
	5. <i>CoLCheckDate</i>	The date (<i>DDMMYYYY</i>) that the taxonomy was checked at the Catalogue of Life
	6. <i>Infraspecific</i>	Taxonomic infraspecific name of <i>SpeciesAccepted</i> , as used by the author
	7. <i>SpeciesEpithetAccepted</i>	Taxonomic species epithet of study species, as per Catalogue of Life
	8. <i>GenusAccepted</i>	Taxonomic genus of study species, as per Catalogue of Life
	9. <i>GenusAuthor</i>	Taxonomic genus of study species, as in <i>SpeciesAuthor</i>
	10. <i>Family</i>	Taxonomic family of study species
	11. <i>Order</i>	Taxonomic order of study species

Source of information ¶	15. <i>Authors</i>	Last names of full authorship in study
	16. <i>Journal</i>	Abbreviated journal of publication (www.abbreviations.com/jas.php), otherwise stated as “PhD thesis”, “MSc thesis”, “BSc thesis”, “Book”, “Report” or “Internet”
	17. <i>YearPublication</i>	Year of publication of source
	18. <i>DOI/ISBN</i>	Digital object identifier (for manuscripts) or international standard book number (for books), when available; old publications do not have an assigned DOI. An <i>R</i> script is also provided to obtain full citation from manuscripts based on DOI (See Online Supplementary Materials 3)
	19. <i>AdditionalSource</i>	If additional information was obtained from a secondary source, the abbreviated citation is included here (First author’s first last name, abbreviated journal name and publication year; e.g.: “Naujokaitis-Lewis <i>Cons Biol</i> 2009” for <i>Canis latrans</i>)
Details of the study ¶	20. <i>StudyDuration</i>	Years of observation of the population dynamics of the species, calculated as $StudyEnd - StudyStart + 1$ (e.g., 2005 – 2000 + 1 = 6)
	21. <i>StudyStart</i>	Year the study started
	22. <i>StudyEnd</i>	Year the study ended
	23. <i>AnnualPeriodicity</i>	Frequency with which seasonal or annual MPMs were constructed (e.g. 1: once per year; 2: twice per year; 0.2: once every five years)
	24. <i>NumberPopulations</i>	Number of populations examined in the study – These may not match the number of populations with MPMs in COMADRE 1.0.0 if the author has not made available all of the MPMs
	25. <i>MatrixCriteriaSize</i>	Whether and on which biometric aspects of the species was the MPM constructed (e.g. height)
	26. <i>MatrixCriteriaOntogeny</i>	Whether some aspect of developmental stage of the species was used to construct the MPM (e.g. juvenile, reproductive adult)
Location ¶	27. <i>MatrixCriteriaAge</i>	Whether some aspect of developmental stage of the species was used to construct the MPM (e.g. 0, 1, 2 years old)
	28. <i>MatrixPopulation</i>	Name(s) of populations from which the MPM was constructed. When no population name is provided in the source, the name of closest geographic landmark or letters in alphabetical (e.g. “A”, “B”, “C”...) or numerical order (e.g. “1”, “2”, “3”...) are used
	29. <i>LatDeg</i>	Latitudinal degrees of study population
	30. <i>LatMin</i>	Latitudinal minutes of study population

	31. <i>LatSec</i>	Latitudinal seconds of study population
	32. <i>LatNS</i>	Latitudinal cardinal direction: North or South
	33. <i>LonDeg</i>	Longitudinal degrees of study population
	34. <i>LonMin</i>	Longitudinal minutes of study population
	35. <i>LonSec</i>	Longitudinal seconds of study population
	36. <i>LonWE</i>	Longitudinal cardinal direction: West or East
	37. <i>Altitude</i>	Altitude of study population (in meters) obtained from Google Earth
	38. <i>Country</i>	Country or countries where the study population was studied. Only countries currently accepted by the United Nations according to the ISO 3 list were used (http://unstats.un.org/unsd/tradekb/Knowledgebase/Country-Code)
	39. <i>Continent</i>	Continent of the study population
	40. <i>Ecoregion</i>	Description of the terrestrial or aquatic ecoregion, corresponding to Olson et al.'s classification (2001), where the study took place. When the study is undertaken in its majority under controlled, indoor conditions (e.g. laboratory, greenhouse), this is noted as " <i>LAB</i> "
Details of matrix population Model α	41. <i>StudiedSex</i>	Sex(es) considered to construct the MPM (Figure 3.B)
	42. <i>MatrixComposite</i>	MPMs were differentiated between matrices that correspond to a given single population, single treatment and single annual period ("Individual"; Figure 3.A), to a single population, treatment and intra-annual period ("Seasonal"), to a MPM that is the result of element-by-element arithmetic mean ("Mean"), or where the individual-level data were pooled to construct a MPM over various periods, populations and/or treatments ("Pooled"). We must note that by default we calculated the mean MPM when all individual MPMs in the study were made available. The pooled and mean matrices for all the individual, unmanipulated (see <i>MatrixTreatment</i>) MPMs coincide when the sample sizes and stage distributions at time t are the same across all the individual MPMs. Mean MPMs were only calculated by us for unmanipulated individual matrices below
	43. <i>MatrixTreatment</i>	Treatment under which the demographic data used to parameterize the specific MPM was exerted. We specified " <i>Unmanipulated</i> " as those matrices where no human-led experimentation was carried out (Figure 3.F). Users are encouraged to carefully examine variable <i>MatrixObservation</i> (below) for additional pertinent information

44. <i>Captivity</i>	Whether the study species was in its wild setting, or under control conditions (e.g. greenhouse, botanical garden) for most of the demographic data that was collected (Figure 3.E)	
Start and end of study period	44. <i>MatrixStartYear</i>	Beginning year t for MPM A describing the population dynamics between time t and year $t+1$
	45. <i>MatrixStartSeason</i>	Beginning season s for seasonal MPM B describing the population dynamics between season s and season $s+1$
	46. <i>MatrixStartMonth</i>	Beginning month m for seasonal MPM B describing the population dynamics between month m and month $m+1$
	47. <i>MatrixEndYear</i>	End year $t+1$ for MPM A describing the population dynamics between time t and time $t+1$
	48. <i>MatrixEndSeason</i>	End season $s+1$ for seasonal MPM B describing the population dynamics between seasons s and season $s+1$
	49. <i>MatrixEndMonth</i>	End month $m+1$ for seasonal MPM B describing the population dynamics between month m and month $m+1$
50. <i>MatrixSplit</i>	To facilitate the calculation of various demographic properties (e.g. life expectancy η_e , mean age at first reproduction L_a , vital rate sensitivities, etc), the MPM A (<i>matA</i> , below) has been split into survival (<i>matU</i>), sexual (<i>matF</i>), and clonal reproduction (<i>matC</i>) submatrices when sufficient information was provided in the source. In 2.9% of the cases, insufficient information led to us not been able to split A into U , F and C . This matrix is referred to as <i>Indivisible</i> (Figure 3.C)	
51. <i>MatrixFec</i>	In some instances the sexual reproductive component of the life cycle of the organism (see <i>matF</i> below) is not modelled either because it is not of interest to the researcher or because of logistical complications in doing so	
52. <i>Observation</i>	Relevant observation that the user should have in mind when analyzing and interpreting the MPMs. In the present version, >50% of the matrices made available in this version have observations. Observations include, for instance, warnings about the description by the author of an “ <i>Unmanipulated</i> ” population that some researchers may wish to treat as a treatment (e.g. natural fires), among others	
53. <i>MatrixDimension</i>	Dimension of the MPM	
54. <i>SurvivalIssue</i>	Reports maximum stage-specific survival in the submatrix U (below). If this value > 1,	

		users are encouraged to carefully evaluate the matrix
	55. <i>MatrixClassAuthor</i>	Classification of the stages in the life cycle of the study species as described by the author
	56. <i>MatrixClassOrganized</i>	Standardization of <i>MatrixClassAuthor</i> into three stages: <i>prop</i> for propagules, <i>dorm</i> for dormant individuals, and <i>active</i> for individuals active, established individuals. We standardized <i>MatrixClassAuthor</i> in this way to facilitate cross comparisons of various general life cycle stages. Note that further general classifications are possible, for instance, distinguishing reproductive individuals from non-reproductive individuals by evaluating the F and C submatrices
	57. <i>MatrixClassNumber</i>	Sequence of numbered classes from 1 to <i>MatrixDimension</i>
Population Matrix Model α	58. <i>matA</i>	MPM including demographic processes that depend on survival (<i>SubMatrixU</i> below), sexual reproduction (if pertinent and available; <i>SubMatrixF</i> below), and clonal reproduction (if pertinent and available; <i>SubMatrixC</i> below; Figure 3)
	59. <i>matU</i>	Submatrix population model describing only survival-dependent demographic processes (e.g. seedbank, stasis, progression, retrogression, vegetative dormancy, etc). Matrix elements corresponding to sexual and clonal reproduction are filled with zeros
	60. <i>matF</i>	Submatrix population model describing only sexual reproduction. All other matrix elements are filled with zeros
	61. <i>matC</i>	Submatrix population model describing only clonal reproduction. All other matrix elements are filled with zeros

7

Figure legends

Figure 1. Time-line of the cumulative number of studies published up until May 31st 2015 containing matrix population models (MPMs) of animals in peer-reviewed journals, books, reports and theses. The light blue background corresponds to studies released in COMADRE version 1.0.0; Dark blue corresponds to studies currently under inspection, to be incorporated in future versions. Pivotal events in the development of the COMADRE Animal Matrix Database: (a, b) first applications of matrix models in demography (Bernardelli 1941; Leslie 1945), (c) first application of MPMs to human demographic projections (Keyfitz 1964), (d) introduction of theory for stage-classified MPMs (Lefkovitch 1965), (e) first spatial MPM (Rogers 1966), (f) first nonlinear, density dependent MPMs for animal populations (Pennycuik 1969; Rabinovitch 1969), (g) first sensitivity analysis for stage-classified MPMs and calculation of selection gradients for animals (Caswell 1978), (h) first bifurcation analysis of density-dependent MPMs in animals (Levin & Goodyear 1980), (i) first calculation of the stochastic growth rate from an animal MPM (Cohen, Christensen & Goodyear 1983), (j) formalization of elasticity analyses for MPMs (de Kroon *et al.* 1986), (k) first application of elasticity analysis in conservation biology for *Caretta caretta* (Crouse, Crowder & Caswell 1987) and first Life Table Response Experiment analysis (Levin *et al.* 1987), (l) first edition of the *Matrix Population Models: Construction, Analysis, and Interpretation* (Caswell 1989), (m) publication of *Population Dynamics in Variable Environments* (Tuljapurkar 1990), (n) first presentation of multi-state mark-recapture methods for estimating stage-structured MPMs in animals (Nichols *et al.* 1992), (o) development of the first MPM from photo identification data (Brault & Caswell 1993), (p) one of the first studies to detail uncertainty in MPMs (Caswell *et al.* 1998), (q) first special feature on MPMs

(Heppell, Pfister & de Kroon 2000), (r) 2nd edition of *Matrix Population Models* (Caswell 2001), (s) publication of *Quantitative Conservation Biology: Theory and Practice of Population Viability Analysis* (Morris & Doak 2002) summarizing and stimulating applications of MPMs to conservation, (t) first application of matrix integro-difference equations to examine invasion speeds in animal populations (Caswell, Lensink & Neubert 2003), (u) first investigation of nonequilibrium properties, such as reactivity, for MPMs (Caswell & Neubert 2005), (v) first complete perturbation analysis for nonlinear animal MPMs (Caswell 2008), (w) introduction of individual stochasticity analyses for animal MPMs (Caswell 2009; Tuljapurkar *et al.* 2009), (x) foundation of the COMADRE database at the Max Planck Institute for Demographic Research, (y) release of the COMPADRE Plant Population Database 3.0 in www.compadre-db.org (Sept 11th 2014), and (z) online release of the COMADRE Animal Matrix Database version 1.0.0 in www.comadre-db.org.

Figure 2. Geographic representation of animal demographic studies in COMADRE.

A. Worldmap of studies in COMADRE 1.0.0. The points represent studied sites, and have been jittered to highlight temporal replication and close spatial overlap of certain studies. World map shows major habitats as color-coded background (See legend). **B.** Breakdown of studies by continent. **C.** Frequency of MPMs by altitude. Negative values typically indicate marine and freshwater sites.

Figure 3. Classification of Matrix Population Models (MPMs) in COMADRE. **A.** By type of matrix (See *MatrixComposite* in Table 1). **B.** By environmental conditions of studied population (*Captivity*). **C.** By general type of treatment of matrix model under consideration (*MatrixTreatment*). **D.** By whether sex was modelled (*StudySex*). **E.** By

whether the matrix **A** was split into submatrices **U**, **F** and **C** (*MatrixSplit*). **F**. By whether reproduction was modelled (*MatrixFec*). **G**. By taxonomic class representation (*Class*).

Figure 4. General aspects of tempo-spatial replication and stage construction of matrices in COMADRE. **A**. Frequency of matrix models by demographic census periodicity on an annual basis (See *Periodicity* in Table 1). **B**. Number of populations (*NumberPopulations*). **C**. Duration of the study (in years; *StudyDuration*). **D**. Matrix dimensionality (*MatrixDimension*). **E**. Criteria used to construct the matrix model (*MatrixCriteriaSize*, *MatrixCriteriaOntogeny* & *MatrixCriteriaAge*).

Figure 1

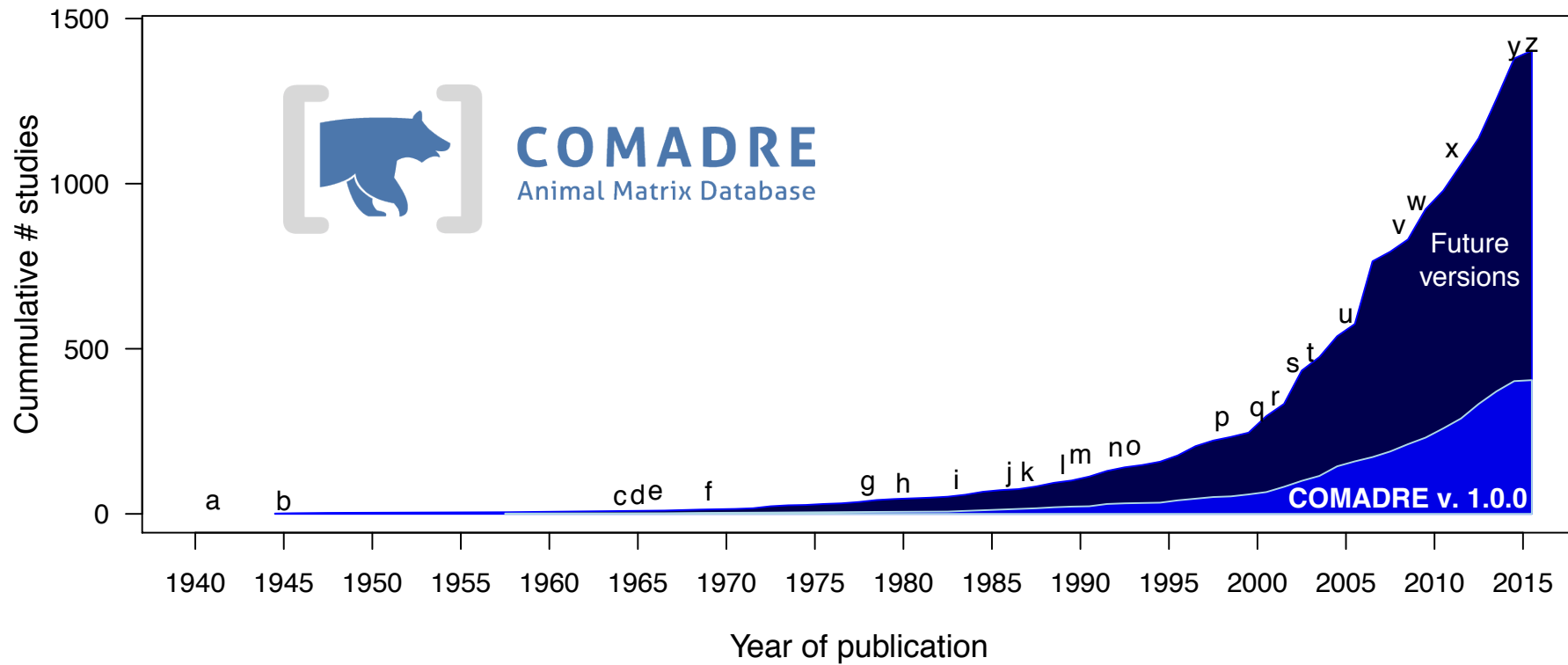


Figure 2

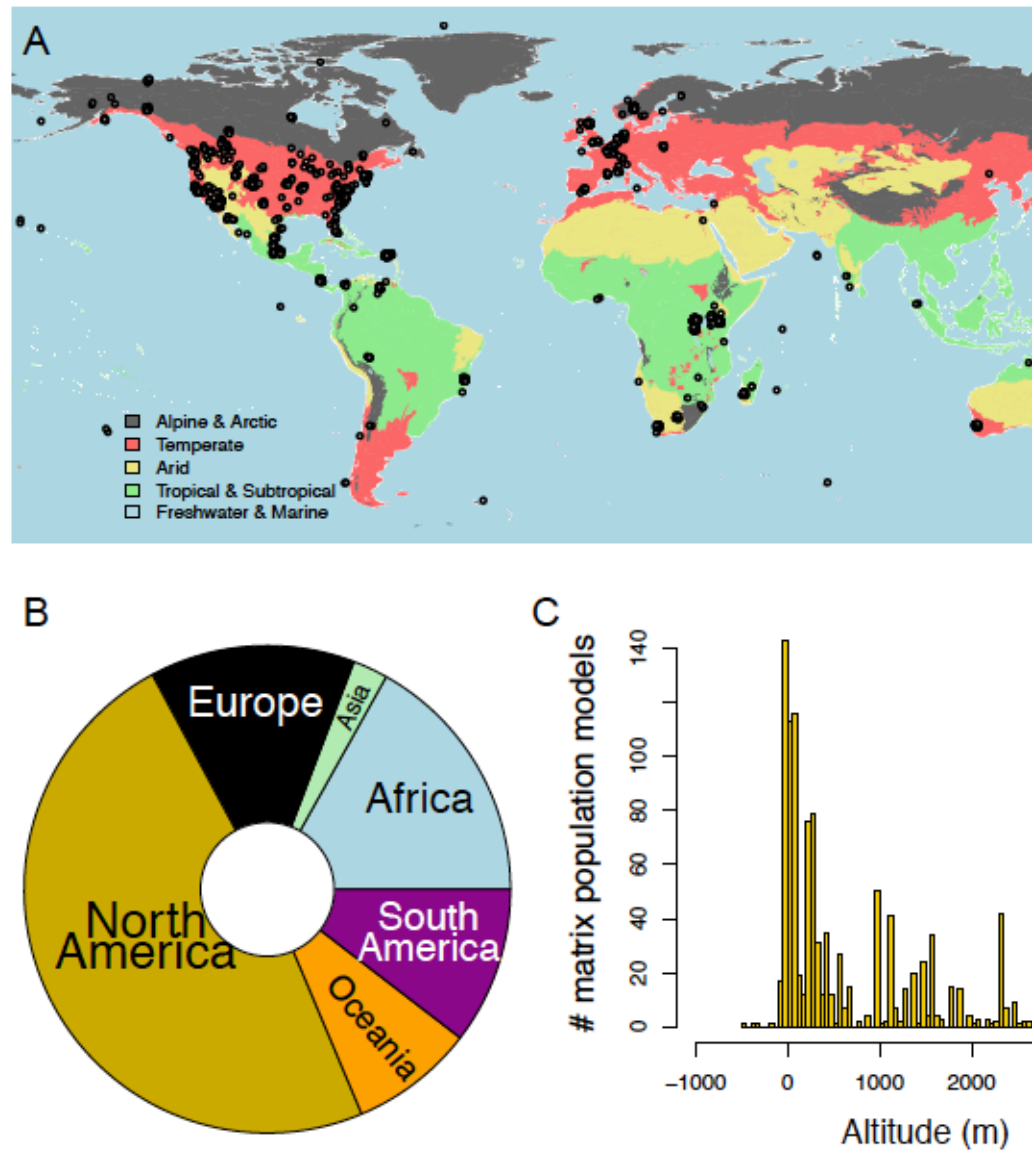


Figure 3

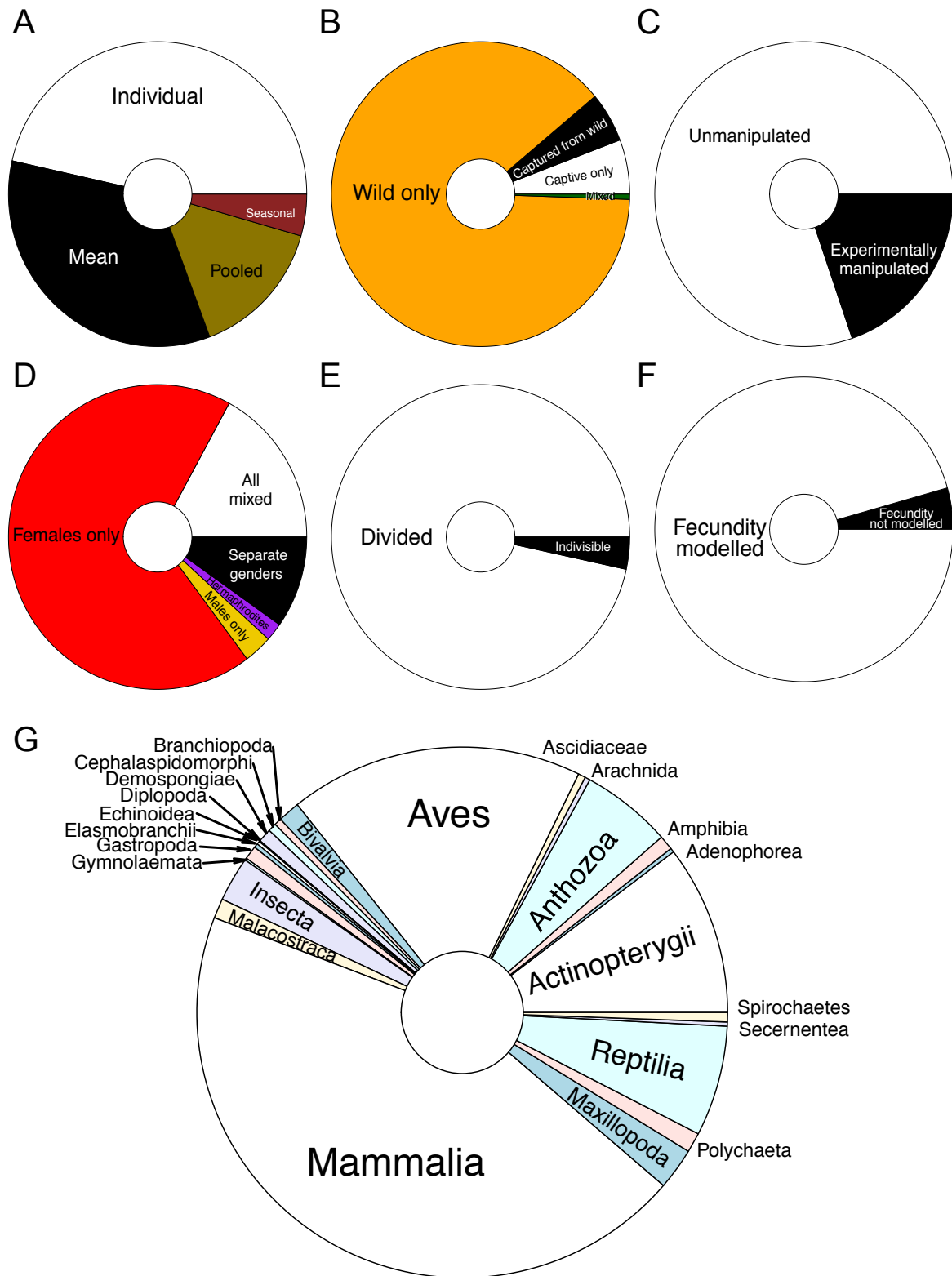


Figure 4

