DOI: 10.1111/2041-210X.14426

#### **RESEARCH ARTICLE**

# **Retrodiction of forest demography: Backward simulation with** reverse matrix models

Nicolas Picard<sup>1</sup> | Jack Andrew Hansen<sup>2</sup> | Nancy Harris<sup>3</sup> | Jingjing Liang<sup>2</sup>

<sup>1</sup>GIP Ecofor, Paris, France

<sup>2</sup>Forest Advance Computing and Artificial Intelligence Lab (FACAI), Department of Forestry and Natural Resources, Purdue University, West Lafayette, Indiana, USA

<sup>3</sup>World Resources Institute, Washington, District of Columbia, USA

#### Correspondence

Nicolas Picard Email: nicolas.picard.1991@polytechnique. org

#### Funding information

Department of Forestry and Natural Resources, Purdue University; World Resources Institute, Grant/Award Number: 40004269

Handling Editor: Maria Pennino

#### Abstract

- 1. Reconstructing past ecological population dynamics and demographic events is crucial for understanding the dynamics of ecological processes, evaluating the impact of environmental changes and making informed conservation decisions. In forest ecosystems, retrodiction (i.e. the backward projection of ecological populations) plays a pivotal role in understanding historical forest carbon levels and the factors that have influenced their variation over time, because forest demography is a major determinant of the amount of carbon stored in forest ecosystems. The persistent lack of quantitative methods has been a significant obstacle in retrodicting forest demography, especially in applications of a broad geographical scale. While there is a wealth of models for predicting future forest conditions, models that can project these conditions backward in time are scarce.
- 2. This study presents reverse matrix model (RMM), an innovative retrodiction modelling approach grounded in the principles of transition matrix models. RMM is designed to deduce past demographic characteristics of ecological populations using current data, making it one of the first models capable of projecting the fine-scale dynamics of forest demography into the past.
- 3. We assessed the retrodictive performance of RMM by fitting it to a dataset of a disturbed tropical rainforest in French Guiana in 2001-2023, then comparing the retrodictions to observations back to 1983 when the disturbance occurred. We further empirically evaluated the viability of retrodiction over a defined duration by inverting the density-dependent matrix model by Lin et al. (1996), which predicts the dynamics of northern hardwoods in the United States.
- 4. The case studies demonstrate significant potential for RMM application in various domains of forestry and conservation, including ecosystem management and conservation planning, global change impact assessment and biodiversity monitoring.

#### **KEYWORDS**

backcasting, backward projection, forest dynamics, matrix population model, retrodiction, transition matrix

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made. © 2024 The Author(s). Methods in Ecology and Evolution published by John Wiley & Sons Ltd on behalf of British Ecological Society.

#### 1 | INTRODUCTION

The backward projection of ecological populations, sometimes also referred to as retrodiction or backcasting, is the process of reconstructing past population dynamics and demographic events. Retrodiction is crucial for understanding the dynamics of ecological processes, evaluating the impact of environmental changes and making informed conservation decisions. Early applications of retrodiction were primarily in the field of population genetics. Kingman (1982) introduced the coalescent process, which utilizes the genealogical relationships among individuals to infer past population dynamics. Hudson (1990) and others (Gill et al., 2016) extended these ideas, applying the coalescent theory to infer past population sizes, migration rates and other demographic parameters.

In forest demography, which quantifies the structure and dynamics of tree populations within forest ecosystems, retrodiction can be used to infer historical forest conditions, such as the distribution of forest tree population by species and size. To this end, forest growth models, typically used to predict future forest conditions (Pretzsch, 2009), can be adapted to project forest conditions backward in time based on current forest conditions and historical data such as land use, fire history, or climate records (Bonnicksen & Stone, 1982). In climate change research, forest growth models can be used to understand how forests have responded to past climate variability and change (Barth et al., 2015). In forest conservation and restoration, backward projections can provide baseline or reference conditions to help restore ecosystems to their historical conditions, or understand how far current conditions have deviated from the natural state (Manning et al., 2006). These applications demonstrate the versatility of forest growth models not only in forecasting future scenarios but also in providing insights into the past, which is essential for informed forest management and conservation planning.

Forest demography is a major determinant of the amount of carbon stored in forest ecosystems (McDowell et al., 2020). Due to inconsistent data, methods and assumptions employed in existing forest carbon monitoring systems (Harris et al., 2021), large uncertainties in the magnitude and direction of the response of the terrestrial carbon cycle to climate change still remain unquantified (Pugh et al., 2020). Consequently, retrodiction plays a pivotal role in our comprehension of historical forest carbon levels and the factors that have influenced their variation over time. Understanding these dynamics is essential for reconstructing historical carbon storage benchmarks (Gasser et al., 2020; Magerl et al., 2019), which can inform predictive models and help in planning conservation and management strategies to mitigate climate change impacts.

Transition matrix models are widely used in many disciplines (Leslie, 1945; Lewis, 1942; Salguero-Gómez et al., 2015), especially in forestry where they have been employed to study the dynamics of almost all the forest types around the world (Liang & Picard, 2013). In the typology of forest growth models, matrix models stand between stand-level models and individual tree models (Porté & Bartelink, 2002). Stand-level models are useful for monospecific even-aged stands, but become inappropriate for mixed or unevenaged stands because they do not consider within-stand variability

in individual size or species. In comparison, individual-based models predict the temporal changes of each individual tree on a stand, but due to their computing intensive nature, they have limited applications in large-scale studies. Matrix models provide a good compromise between computing requirements over large scale and capacity to deal with individual variability.

Matrix models have been applied to study almost all the aspects of forestry (Ohse et al., 2023). In forest ecology, matrix models have been used to study the demography of natural succession (e.g. Alvarez-Buylla, 1994) and the impacts of climate changes (Liang et al., 2011), and natural disturbances (Price & Bowman, 1994). In forest management, matrix models have been applied to evaluate forest carbon sequestration and biomass potential (Ingram, 2012; Liang, 2012), economic outcomes (Ingram & Buongiorno, 1996) and ecological impacts (Ma et al., 2016, 2020) of various management regimes, including the impact on the habitat of endangered wildlife species (Zhou et al., 2008).

The persistent lack of quantitative methods has been a significant obstacle in retrodicting forest demography, especially in applications of a broad geographical scale. While there is a wealth of models for predicting future forest conditions, models that can project these conditions backward in time are scarce. The mathematical methods that underpin matrix modelling pave the way for retrodicting matrix models. Matrix models are related to several mathematical fields, such as discrete-time finite-space Markov chains (Bruner & Moser, 1973; Caswell, 2001, Chap. 5) and partial differential equations (Caswell, 2001, Chap. 8; Takada & Hara, 1994). The retrodiction of Markov chains (Surace & Scandi, 2023) and the time-reversal of partial differential equations thus provide hints for retrodicting matrix models. However, matrix models do not entirely fall within the scope of these mathematical fields, so that none of them completely defines how to retrodict matrix models.

This study presents reverse matrix model (RMM), a retrodiction modelling approach grounded in the principles of transition matrix models. RMM is designed to deduce past demographic characteristics of ecological populations using current data, marking it one of the first models capable of projecting the fine-scale dynamics of forest demography into the past. Two related guestions were addressed: (i) How can a RMM be fitted to data? (ii) Given a forward matrix model fitted to data, how can it be inverted using RMM or another inversion method? For question (i), we hypothesize that RMM has a retrodictive capacity similar to the predictive capacity of matrix models, as they are grounded in the same theoretical framework. For question (ii), we hypothesize that, because some information is lost in the forward direction, inversion is possible only within a bounded time frame. We assessed the retrodictive performance of RMM by fitting it to a dataset of a disturbed tropical rainforest in French Guiana in 2001-2023, then comparing the retrodictions to observations back to 1983 when the disturbance occurred. In addition to RMM, we also devised explicit methods to invert a matrix model from a forward to a backward recurrence relationship. We empirically evaluated the duration of time during which the backward dynamics followed the same trajectory as the forward dynamics by inverting the densitydependent matrix model by Lin et al. (1996), which predicts the dynamics of northern hardwoods in the United States.

#### 2 | MATERIALS AND METHODS

#### 2.1 | Matrix model

In a matrix model of forest dynamics, a population is represented by a vector **N** giving the abundance of individuals in *K* diameter classes at a given time *t*. Time is discrete. The dynamics is determined by the matrix recurrence relationship  $\mathbf{O} = \mathbf{GN} + \mathbf{R}$ , where  $\mathbf{O}$  is the vector of tree abundances at the subsequent time t + 1,  $\mathbf{G}$  is a  $K \times K$  growth transition matrix, and  $\mathbf{R}$  is a vector of length *K* whose first element gives the number of recruited individuals between times t and t + 1and all other elements are null. By iterating the forward recurrence relationship that computes  $\mathbf{O}$  from  $\mathbf{N}$ , future dynamics can be predicted. Between times t and t + 1, an individual may either move up from its diameter class to the next one, stay alive in its diameter class, or die. Hence, the growth transition matrix is written:

$$\mathbf{G} = \begin{bmatrix} q_{1} & & \mathbf{0} \\ p_{1} & q_{2} & & \\ & \ddots & \ddots & \\ \mathbf{0} & & p_{K-1} & q_{K} \end{bmatrix}$$
(1)

where  $q_i$  is the stasis rate in class *i* and  $p_i$  is the upgrowth rate from class *i* to *i* + 1. The mortality rate in class *i* is obtained as  $m_i = 1 - p_i - q_i$ .

A particular case is when the first element of **R** is a linear function of **N**, that is written as <sup>t</sup>**fN** where superscript t denotes the transpose and  $\mathbf{f} = (f_1, \dots, f_K)$  is a vector of *K* fecundities. Most of the time, newly recruited trees cannot be related to their parent trees, so that the average fecundity  $\mathbf{f} = ({}^{t}\mathbf{1}_{K}\mathbf{R}) / ({}^{t}\mathbf{1}_{K}\mathbf{N})$  is the only estimable quantity. Then, the matrix relationship becomes:  $\mathbf{O} = \mathbf{U}\mathbf{N}$ , where  $\mathbf{U}$  is a Usher (1966, 1969) transition matrix:

$$\boldsymbol{U} = \begin{bmatrix} q_1 + f & f & \dots & f \\ p_1 & q_2 & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & p_{K-1} & q_K \end{bmatrix}$$

A matrix model predicts the transitions and per capita asexual or sexual contributions between the states of a life cycle (Caswell, 2001). In forest dynamics, the states of the life cycle are diameter classes, and transitions are stasis, upgrowth and mortality (Figure 1). The trajectory of an individual along the life cycle can be described by a Markov chain. Given a dataset giving the state of every tree at time t and t + 1, the matrix model is calibrated by counting, for each diameter class *i*, the number  $M_i$  of trees that die, the number  $Q_i$  of trees that stay alive in diameter class *i*, and the number  $P_i$  of trees that move up from diameter class *i* to *i* + 1 (Figure 2). Transition rates are computed with respect to the number of living trees in each diameter class (i.e. excluding the number of dead trees). Thus, the forward transitions rates are: for class *i* = 1, ..., K - 1,  $m_i = M_i / N_i$ ,  $q_i = Q_i / N_i$  and  $p_i = P_i / N_i$ , where  $N_i = M_i + Q_i + P_i$  is the number of living trees in diameter class *i* at time *t*; for the last class,  $m_K = M_K / N_K$  and  $q_K = Q_K / N_K$ , where  $N_K = M_K + Q_K$  is the number of living trees in diameter class *K* at time *t*.

Matrix models with fixed transition rates may not be sufficiently flexible to account for population dynamics (Caswell, 2001, Chap. 16; Roberts & Hruska, 1986). Transition matrix may actually vary between populations depending on their density. Density-dependent models where the transition rates are functions of **N** provide more realistic predictions when density-dependent processes such as competition affect vital rates (Favrichon, 1998; Lin et al., 1996). Density-dependent models are often fitted using a two-step process (Solomon et al., 1986). First, the transition rates for different populations at varying densities are estimated as described above. Second, these transition rates are regressed against population density.

#### 2.2 | Reverse matrix model

A backward matrix model can be defined by reversing the direction of the transitions between the states of the life cycle (i.e. by reversing the direction of the arrows in Figure 1). Because the life cycle description of matrix models complies with a Markov chain, the retrodiction of Markov chains offers a prima facie way to develop RMMs. Let *i* and *j* be two states of the life cycle that are connected by an arrow, and let  $X_t$  be the state of an individual at time *t*. By definition, the transition rate from *i* to *j* in the matrix model is the probability  $Pr(X_{t+1} = j | X_t = i)$ . Retrodicting the matrix model consists in computing the backward transition rate from *j* to *i*, that is the probability  $Pr(X_t = i | X_{t+1} = j)$ . The retrodiction of



FIGURE 1 Life cycle for a matrix growth model of forest dynamics. Circles represent the diameter classes. Arrows represent the possible transitions between two consecutive time steps. Between time t and t + 1, a tree in diameter class i dies with probability  $m_i$ , stays alive in class i with probability  $q_i$  or moves up from class i to i + 1 with probability  $p_i$ .



**FIGURE 2** Given a dataset with the observed status (alive or dead) and diameter of *N* trees at times *t* and t + 1, representation of the count of trees in each diameter class at each time step according to their status. Diameter classes are represented by coloured boxes: orange = diameter classes for the forward transitions; blue = diameter classes for the backward transitions. For each diameter class, i = 1, ..., K, the counts of tree are:  $M_i$ , counts of trees that die in diameter class *i* between time *t* and t + 1;  $Q_i$ , count of trees that stay alive in diameter class *i* to i + 1. In the first class, *R* is the number of recruited trees between time *t* and t + 1.

Markov chains is commonly done using Bayes' formula (Surace & Scandi, 2023):

$$\Pr(X_t = i | X_{t+1} = j) = \frac{\Pr(X_{t+1} = j | X_t = i) \Pr(X_t = i)}{\sum_k \Pr(X_{t+1} = j | X_t = k) \Pr(X_t = k)}$$

These calculations are continued in Appendix A. However, this approach has an inherent limitation in defining RMM because the number of past dead individuals in a population is unknown. To compute the number of dead individuals in matrix models, it is enough to know the number of living individuals in each state. However, to compute the backward transitions from death, the cumulative number of dead individuals must be known. In other words, we can define stasis and reverse growth as backward transitions, but we cannot define reverse mortality as a backward transition by solely monitoring living individuals.

Nonetheless, RMM can be defined in a way similar to matrix models by counting the individuals that experience backward transitions and expressing these numbers in relative terms to the number of living individuals in each state. For backward transitions, the number of living trees in diameter class *i* at time t + 1 is:  $O_i = Q_i + P_{i-1}$  for each class i = 2, ..., K (Figure 2). The stasis rate, that is the rate of trees that were already in class i at time t, is:  $r_i = (M_i + Q_i) / O_i$ . The reverse growth rate, that is the rate of trees in diameter class *i* at time t + 1 that were issued from diameter class i - 1, is:  $b_i = P_{i-1} / O_i$ . Contrary to the forward transition rates that sum to one, the sum of backward transition rates is greater than one:  $r_i + b_i = 1 + M_i / O_i$ . The quantity  $r_i + b_i - 1$  now represents the mortality rate in class *i*. The first diameter class is a bit different from the others because it does not have any reverse growth rate but includes the newly recruited trees. These recruited trees have to be removed to compute the stasis rate, which is  $r_1 = (M_1 + Q_1) / (O_1 - R) = 1 + M_1 / Q_1$ , where  $O_1 = Q_1 + R$  is the number of living trees in diameter class 1 at time t + 1.

The backward matrix relationship thus is N = C(O - R), where the reverse growth transition matrix C is

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{b}_2 & \boldsymbol{0} \\ & \boldsymbol{r}_2 & \ddots & \\ & & \ddots & \boldsymbol{b}_K \\ \boldsymbol{0} & & & \boldsymbol{r}_K \end{bmatrix}$$
(2)

As in matrix models, recruitment in RMM can also be expressed proportionally to **O** rather than in an additive form. The backward matrix relationship then is N = VO, where the reverse transition matrix **V** is

$$\mathbf{V} = \begin{bmatrix} r_1 - h & b_2 - h & -h & \dots & -h \\ & r_2 & b_3 & & \mathbf{0} \\ & & \ddots & \ddots & \\ & & & \ddots & b_K \\ \mathbf{0} & & & & r_K \end{bmatrix}$$
(3)

and *h* is the average fecundity rate of the reverse matrix model. Given the counts of individuals in diameter classes, the expression of *h* is  $h = (1 + M_1/Q_1)R/(O_1 + ... + O_K).$ 

Given a dataset with the observed diameters of trees at time tand t + 1, including the diameter of trees that were alive at time t but dead at t + 1 and the diameter of trees that were recruited between t and t + 1, the expressions of  $r_i$ ,  $b_i$  and h allow us to fit RMM to observations. The fitted backward recurrence relationship can then be iterated to retrodict past dynamics.

#### 2.3 | Inversion of the recurrence relationship

When a forward matrix model is already fitted but the goal is to retrodict past dynamics rather than predict future dynamics, the question is to invert the recurrence relation of the forward model. We considered four methods to invert the forward recurrence relation. One was based on RMM and provided interpretability of the inversion process in terms of vital rates. The other three inversion methods aimed to solve for the input term of the recurrence without focusing on the biological interpretation of the inversion process: one was a numerical method without an explicit expression of the solution, while the other two were based on zero- and first-order approximations of the recurrence relation, allowing for an explicit expression of the solution to be computed. When the forward transition matrix has fixed coefficients, these three methods boil down to matrix inversion. Differences among the three methods arise in the more complicate case of density-dependent matrix models.

#### 2.3.1 | Numerical inversion

Given a vector of abundances **O**, a density-dependent matrix model can be numerically inverted to find the inverse image **N** of **O**. It consists in using a numerical algorithm to solve the optimization problem  $\min_{\mathbf{N}} || \mathbf{O} - \mathbf{G}(\mathbf{N})\mathbf{N} - \mathbf{R}(\mathbf{N}) ||$ , where  $|| \cdot ||$  is a norm of  $\mathbb{R}^{K}$  (e.g. the  $\ell^{2}$ -norm). Numerical optimization methods require a starting point for **N**. Any of the solutions presented in the subsequent sections can be used as a starting point for the numerical optimization method.

#### 2.3.2 | Zero-order approximation

A zero-order approximation of the inverse image of the vector of abundances **O** was obtained by replacing **N** by **O** in the expressions of the transition matrix and recruitment vector. In other words, instead of solving the non-linear equation O = G(N)N + R(N) for **N**, we solve the linear equation O = G(O)N + R(O). This approximation makes sense if **N** is close enough to **O**, typically if the time step of the matrix model is small enough (e.g. 1 year) so that small changes are observed between times t and t + 1. Because the determinant of **G** is  $|\mathbf{G}| = q_1 q_2 \dots q_K > 0$ , matrix **G** is invertible. Therefore, the zero-order approximation of the back-projected state is

$$\mathbf{N} = \mathbf{G}(\mathbf{O})^{-1} \left[ \mathbf{O} - \mathbf{R}(\mathbf{O}) \right] \tag{4}$$

If the first element of **R** is a linear function of **N**, the zero-order approximation is rather obtained by solving O = U(O)N, where U is a Usher transition matrix. Under mild conditions (see Supporting Information S2), we can show that an Usher matrix is invertible. The zero-order approximation of the back-projected state thus is  $N = U(O)^{-1}O$ .

#### 2.3.3 | First-order approximation

A first-order approximation of the inverse image of the vector of abundance **O** was obtained by making a first-order Taylor expansion at **O** of the application  $\phi$  that maps **N** onto **G**(**N**)**N** + **R**(**N**). The first-order expansion of  $\phi$  at **O** is (Magnus & Neudecker, 2007):  $\phi(\mathbf{O}) + \mathbf{D}_{\phi}(\mathbf{O})(\mathbf{N} - \mathbf{O})$ , where  $\mathbf{D}_{\phi}$  is the Jacobian matrix of  $\phi$ , that is the

 $K \times K$  matrix whose *ij*th element is the partial derivative of the *i*th element of  $\phi(\mathbf{N})$  with respect to the *j*th element of  $\mathbf{N}$ . Using this first-order expansion, an approximate solution of  $\mathbf{O} = \phi(\mathbf{N})$  thus is

$$\mathbf{N} = \mathbf{O} + \mathbf{D}_{\phi}(\mathbf{O})^{-1} \left[ \mathbf{O} - \mathbf{G}(\mathbf{O})\mathbf{O} - \mathbf{R}(\mathbf{O}) \right]$$
(5)

Using the chain rule of matrix differentiation,  $D_{\phi}(N) = ({}^{t}N \otimes I_{K})J(N) + G(N) + D(N)$ , where J is the Jacobian matrix of G, D is the Jacobian matrix of R,  $I_{K}$  is the  $K \times K$  identity matrix, and  $\otimes$  denotes the Kronecker product (Magnus & Neudecker, 2007, p. 31).

We can readily check that Equation (5) generalizes Equation (4). Indeed, the Jacobian matrix of the application that maps **N** onto G(O)N + R(O) is G(O). Replacing  $D_{\phi}(O)$  by G(O) in Equation (5) brings back Equation (4).

#### 2.3.4 | RMM as an inversion method

By replacing  $M_i$ ,  $Q_i$  and  $P_i$  by  $m_i N_i$ ,  $q_i N_i$  and  $p_i N_i$  respectively in the expressions of  $r_i$  and  $b_i$ , one can express the backward transition rates as a function of the forward transition rates. It brings

$$\begin{cases} r_{1} = (1 - p_{1})/q_{1} \\ r_{i} = (1 - p_{i})/(q_{i} + p_{i-1}/\eta_{i}) & \text{for} \quad i = 2, \dots, K - 1 \\ b_{i} = p_{i-1}/(p_{i-1} + q_{i}\eta_{i}) & \text{for} \quad i = 2, \dots, K \\ r_{K} = 1/(q_{K} + p_{K-1}/\eta_{K}) \end{cases}$$
(6)

where  $\eta_i = N_i / N_{i-1}$ . The expression of  $b_i$  given here is exactly the one obtained by the retrodiction of Markov chains (see Equation 7 in Appendix A), whereas the expression of  $r_i$  slightly differs from that obtained in this manner (see Appendix A). Using these expressions, a reverse matrix model can straightforwardly be built from a given matrix model. In other words, RMM can also be seen as an inversion method. Even when their transition rates are related by Equation (6), the backward transition matrix *C* is not the inverse of the forward matrix *G* (Supporting Information S4). Nonetheless, *GC* and *CG* are close to the identity matrix.

#### 2.4 | Case studies

Two case studies illustrate RMM and the inversion of matrix models. First, we fit a reverse matrix model to a dataset from a tropical rainforest and compare the retrodicted dynamics to past observations. Second, we inverse a density-dependent matrix model and compare the capacity of the four inversion methods to match the predictions of the forward model.

#### 2.4.1 | RMM of a tropical rain forest

We used the data from the Paracou experimental site in French Guiana to fit a reverse matrix model (Gourlet-Fleury et al., 2004). 6

In 1984, twelve 6.25-ha permanent plots were set up at Paracou. Each plot was divided in four 1.5625-ha subplots, which are here the repetition units. All trees with diameter at breast height greater than 10 cm were inventoried. The plots have been monitored since 1984, on an annual basis until 1995, then every 2 years. Diameters have been remeasured, and dead trees and newly recruited trees have been recorded at each survey. The 12 plots were distributed in three blocks. In each block, three silvicultural treatments of increasing intensity, combining selective logging and additional thinning, were applied, each to one plot. The fourth plot in each block was left as control. Thus, for each treatment, there were twelve 1.5625-ha repetitions. Treatments were applied between 1986 and 1988, then no intervention was made and natural dynamics occurred. Treatments removed on average 18.2% (treatment 1), 36.1% (treatment2) and 43.9% (treatment 3) of the basal area of 1986 (Gourlet-Fleury et al., 2004, p. 256). In 2003, the forest on the treated plots was still recovering from the disturbance that had occurred 17 years earlier.

We fitted a RMM for each treatment separately. We used RMMs with fixed transition rates and recruitment proportional to **O**, that is with the transition matrix given by Equation (3). Following Favrichon (1998), we used 11 diameter classes ranging from 10 to 60cm dbh with a constant width of 5cm, except the last class that gathered all trees greater than 60cm dbh. Observations in 2001 and 2003 were used to fit the RMMs. Therefore, the time step of the RMMs was 2 years. Forest dynamics were then retrodicted till 1985.

# 2.4.2 | Inversion of a density-dependent matrix model

We back-projected the density-dependent matrix model developed by Lin et al. (1996) to study the forest tree demography and financial income of the northern hardwoods in the United States. This model relies on K = 12 diameter classes. The time step is 1 year. There are three species groups, denoted I to III. Upgrowth rates depend on the plot basal area:  $p_{gi} = \beta_{gi} + \beta_g B$ , where  $p_{gi}$  is the upgrowth rate for trees in species group g and diameter class *i*, *B* is the plot basal area, and  $\beta_{gi}$  and  $\beta_{g}$  are fixed coefficients. The plot basal area is computed as  $B = \sum_{g=1}^{III} {}^{t}BN_{g}$ where **B** is the vector giving the average basal area of trees in each diameter class and  $N_{g}$  is the vector of tree abundances per diameter class for species group g. Mortality rates  $m_{gi}$  are density-independent. Stasis rates are computed as  $q_{gi} = 1 - m_{gi} - p_{gi}$  (where, by convention,  $p_{gK} = 0$ ) and are thus also density-dependent. The recruitment vector for species group g is  $\mathbf{R}_q = [\mathbf{R}_q, 0, \dots, 0]$ , where the recruitment rate is also densitydependent,  $R_g = \alpha_{g0} + \alpha_{g1}B + \alpha_{g2}N_g$ , where  $N_g$  is the total abundance of species group g and the  $\alpha_{gk}$ s are fixed coefficients. The total abundance of species group g is computed as  $N_g = {}^t \mathbf{1}_K N_g$ , where  $\mathbf{1}_K$  if the vector of length K full of ones. In total, the model has 30 parameters, including 9 for recruitment, 12 for growth and 9 for mortality.

Starting from an empty plot, we projected the state of the forest using Lin et al.'s density-dependent matrix model over a time horizon of *T*, bringing a final vector N(T) of tree abundances in the three species groups. Starting from N(T), we then back-projected

the state of the forest for the same period of time T. We used the four back-projection techniques described in the previous sections: numerical inversion, zero-order approximation, first-order approximation and RMM. For a given horizon T, let N(t) and O(t) be the vectors of tree abundances at time t = 1, ..., T when moving forward and backward, respectively, with N(T) = O(T). We defined divergence between forward projection N(t) and backward projection  $\mathbf{O}(t)$  as an absolute difference in basal area of any of the species group greater than  $0.1 \text{ m}^2 \text{ ha}^{-1}$  and a relative difference in basal area of any of the species group greater than 5%. To assess divergence, we also computed the root mean square error (RMSE) of the difference between N(t) and O(t). The distance to stable structure was defined as the RMSE of the difference between N(t) and  $\lim_{t\to\infty} \mathbf{N}(t)$ . We defined divergence time  $t_{div}$  as the maximum value of t such that O(t) diverges from N(t). When moving backward,  $t_{div}$ actually gives the year when divergence starts. We defined the duration of agreement between the forward and backward projections as  $\tau = T - t_{div}$ . Finally, we defined the elasticity of the duration of agreement to a model parameter  $\theta$  as the relative rate of change of the duration of agreement with respect to the relative rate of change of  $\theta$ , that is  $\partial \ln \tau / \partial \ln \theta$ .

#### 3 | RESULTS

#### 3.1 | Retrodicted dynamics at Paracou

The reverse growth rate increased from control plots to the plots under treatment 3, following the trend in perturbation intensity (Table 1, Supporting Information S5). So did the fecundity rate. The stasis rate showed the opposite trend. To summarize, the more intense the treatment, the faster the reverse dynamics.

RMM extrapolated the dynamics observed in 2001–2003 to the past. For those plots that were not disturbed (control plots) or slightly disturbed in 1986–1988 (treatment 1), the dynamics were quite linear and the RMM retrodiction conformed to observed dynamics (Figure 3a,b and Figure S1a,b in Supporting Information S6). However, for more intense disturbance (treatments 2 and 3), the post-disturbance plot dynamics showed non-linearity. The RMM with fixed transition rates was then

TABLE 1Transition rates and dominant eigenvalue of the reversetransition matrices fitted to forest plots at Paracou, French Guiana.

Treatment	b	h	ī	λ
то	0.089	0.021	0.939	0.998
T1	0.109	0.042	0.918	0.988
T2	0.152	0.049	0.880	0.971
ТЗ	0.169	0.061	0.863	0.966

Note: These plots underwent silvicultural treatments of increasing intensity, from TO (control plots) to T3.  $\overline{b}$  is the average reverse growth rate across diameter classes, h is the fecundity rate,  $\overline{r}$  is the average stasis rate across diameter classes, and  $\lambda$  is the dominant eigenvalue.

FIGURE 3 The retrodicted basal area of the rainforest at Paracou. French Guiana, matched observations from 2003 back to 1988. Retrodiction was achieved using a reverse matrix model fitted to the observations of 2001 and 2003. The black line is the model retrodiction. The coloured line is the observed average trajectory and the shaded area is the observed envelope of the trajectories of twelve 1.5625-ha plots. These plots underwent silvicultural treatments between 1986 and 1988: (a) control plots (no treatment), (b) treatment 1 (low disturbance intensity). (c) treatment 2 (medium disturbance intensity) and (d) treatment 3 (high disturbance intensity).



less successful at retrodicting dynamics. The agreement between retrodictions and observations was acceptable for basal area (Figure 3c,d) but showed limitations for the density of trees (Figure S1c,d in Supporting Information S6).

As for forward matrix models, the long-term behaviour of RMM at Paracou was determined by the dominant eigenvalue of the reverse transition matrix **V**. The dominant eigenvalue decreased from control plots to the plots under treatment 3, following the opposite trend of the perturbation intensity (Table 1). All eigenvalues were less than one. Therefore, moving backwards from 2003, RMM predicted a long-term origin of the forest that was a void forest. However, this predicted backward decline (= recovery in the chronological order) was slow (almost at equilibrium) for controls plots. To summarize, the more intense the treatment, the faster the long-term backward decline.

#### 3.2 | Inversion of Lin et al.'s model

In Lin et al.'s density-dependent matrix model, the growth transition matrix **G** and the recruitment vector **R** linearly depend on **N**, so the application  $\phi$  that maps **N** onto **G**(**N**)**N** + **R**(**N**) depends quadratically on **N**. Therefore, the Jacobian **J** of **G** and the Jacobian **D** of **R** are constant, whereas the Jacobian **D**<sub> $\phi$ </sub> of  $\phi$  is linear with respect to **N**. The computations of the Jacobian matrices for Lin et al.'s matrix model are given in Supporting Information S3.

Starting from an empty plot, Lin et al.'s matrix projection model predicted a rapid increase in basal area (till  $\approx$ 50 years), followed by slow dampened oscillations that finally converged (after  $\approx$ 500 years)

to a stable diameter structure (see black lines in Figure 4 and Supporting Information S7). Starting from different initial diameter structures, the same stable structure was reached (Figure 5 and Supporting Information S8). The dampened oscillations of basal area that are observed over time correspond to a spiral pattern in the phase space (Figure 5d).

Starting from an empty plot, the backward projections using the first-order approximation succeeded to follow the forward trajectories over at least 100 years (Figure 4, Figures S2–S4 and Table 2). In particular, when projecting the diameter distributions from T = 100 years backwards, the backward model succeeded to go back to the initial state of an empty plot. When initiating the retrodiction from a state closer to the stable structure, the inversion method was able to follow the spiral of the phase space for a bit less than a cycle, then quickly diverged from it.

The numerical inversion method for backward projections had the same performance as the first-order approximation (Figure 4, Figures S2–S4). Backward projections with the numerical inversion method diverged from the forward projections at the same time as with the first-order approximation (Table 2). However, the computing time was 20 times longer with the numerical inversion method than with the first-order approximation.

The backward projections using the zero-order approximation succeeded to follow the forward trajectories over a shorter period of time than with the first-order approximation (Figure 4, Figures S2–S4). Accordingly, the divergence time with this method was closer to the time horizon of the forward projections (Table 2). The backward projections using RMM was able to follow the forward trajectories over an even shorter period of time (Figure 4, 8



FIGURE 4 After 500 years of forward prediction using the density-dependent matrix model by Lin et al. (1996), and starting from the resulting state, backward projections closely tracked the predicted basal area for 62 to 207 years (depending on the inversion method) before quickly diverging. Top row shows the forward (black lines) and backward (coloured lines) projections of basal area; the three lines correspond to the three species groups of Lin et al. (1996). Bottom row shows the root mean square error of the retrodicted diameter structure, where the error is calculated with respect to the structure predicted by the forward model. Backward projections are made using four methods (in columns): Numerical inversion ("num"), first-order approximation ("1-order"), zero-order approximation ("0-order") and reverse matrix model ("RMM").



FIGURE 5 In the forward projection, starting from different initial states, the density-dependent model by Lin et al. (1996) converged to the same equilibrium point. To avoid overloading the figure, only eight different initial states are shown. Projections are shown in the three-dimensional subspace of the phase space corresponding to the basal areas of the three species groups: (a) basal area of group II versus basal area of group I; (b) basal area of group III versus basal area of group I; (c) basal area of group III versus basal area of group I. Panel (d) is a zoom of the grey area shown in panel (b). Colours correspond to the time steps (early time is shown in green and late time in red).

Figures S2–S4). Its divergence time was closer to the time horizon of the forward projections than with the other inversion methods (Table 2).

Starting from an initial empty plot, the duration of agreement  $\tau$  between the forward and backward projections had a tendency to increase as the time horizon T of the forward projection increased (Figure 6). Nonetheless, the duration of agreement also depended on the specific position of the state reached at time T with respect to the stable structure. This dependence of  $\tau$  on the position resulted in a succession of peaks and troughs in the  $\tau$  versus T curve, separated by a duration equal to one cycle of the spiral of the phase space (Figure 6a). When following a forward trajectory, the successive position changed in a continuous way, so the  $\tau$  versus T curve also varied continuously. When plotting  $\tau$  against the distance to the stable structure, this was no longer true (Figure 6b). The succession of peaks and troughs along the cycles then resulted in an hysteresis pattern.

Whatever the inversion method used, the duration of agreement  $\tau$  was less sensitive to the recruitment parameters of the matrix model than to its growth or mortality parameters (Figure 7). RMM was the least sensitive inversion method to the parameters of the matrix model, while the zero-order approximation was the most sensitive one.

TABLE 2 Divergence time between the forward projections of the density-dependent matrix model by Lin et al. (1996) and the backward projections, depending on the method used for model inversion and the time horizon of forward projections. An empty cell means that there was no divergence between the forward and backward projections.

	Time h	orizon T (y	/ear)	
Inversion method	100	300	500	700
Numerical inversion	_	147	293	505
First-order approximation	_	147	293	505
Zero-order approximation	21	200	377	585
Reverse matrix model	50	227	438	640

### 4 | DISCUSSION

Retrodiction allows the matrix models, a modelling approach that has proven effective in predicting future population dynamics, to now be applied for exploring and understanding past dynamics. To retrodict population dynamics using the matrix modelling framework, one has two options: use a reverse matrix model or inverse a forward matrix model. The latter option can be preferred when a matrix model is already available. Inversion methods can thus benefit from the large pool of existing matrix models (Salguero-Gómez et al., 2015, 2016). In contrast, when no forward model a priori exists and observations are to be used to build a model, the former option can be preferred. In both cases, one must remain aware of the functional assumptions made in any modelling work and of the influence that the model choice can have on predictions (Bell et al., 2013; Rueda-Cediel et al., 2015). Even when a forward matrix model already exists, RMM can be seen as a way to re-assess these assumptions.

#### 4.1 | Retrodiction using RMM

#### 4.1.1 | Defining RMM

We defined RMM using the counts of individuals in diameter classes and the proportions relative to the number of living individuals at time t + 1. This approach brought the same expression of the reverse growth rate as the retrodiction of the Markov chain based on the life cycle. Markov chain retrodiction tells us that we may end up with little capacity to reconstruct the past if too much information is lost in the forward direction (Ellison et al., 2009; Rupprecht & Vural, 2018). Because matrix modelling loses the information on dead trees in the forward direction, we can anticipate that the perfect retrodiction of matrix models will not be possible. Nonetheless, matrix models depart from the Markov chain formalism in several aspects, so that results from Markov chain retrodiction cannot be directly applied to matrix modelling.

**FIGURE 6** The duration of agreement  $\tau$  between the forward and backward projections of the density-dependent matrix model by Lin et al. (1996) (a) nonlinearly depended on the time horizon T of the forward projection and (b) nonuniquely depended on the distance to the stable diameter structure. Backward projections are made using four methods: Numerical inversion (red line), first-order approximation (red line too), zero-order approximation (green line) and reverse matrix model (cyan line). In (a), the dotted line is the line y = x.





FIGURE 7 Elasticities of the duration of agreement to the parameters of the density-dependent matrix model by Lin et al. (1996). The duration of agreement is computed using four inversion methods: Numerical inversion (orange), first-order approximation (red line), zero-order approximation (green) and reverse matrix model (cyan). The matrix model has 30 parameters that can be related to three dynamics processes:  $\alpha$ -parameters (9 parameters) relate to recruitment,  $\beta$ -parameters (12) relate to growth, and  $\delta$ -parameters (9) relate to mortality.

Partial differential equations provide another standpoint on matrix model retrodiction. A matrix model can be seen indeed as a finite difference approximation of a partial differential equation (Caswell, 2001, Chap. 8; Takada & Hara, 1994). If the solution of the partial differential equation exists and is unique, it predicts the size structure of the population at any time, whether in the past or in the future (Ackleh et al., 2007). Nevertheless, because the matrix model is a dissipative finite difference scheme (Picard & Liang, 2014), this mathematical standpoint also let us expect some inherent limitation in reconstructing the past from current situation.

#### 4.1.2 | Possible applications of RMM

While this study primarily focuses on case studies with specific matrix models in the context of forest dynamics, the conceptual framework of RMM and its adaptability suggest broader applicability across diverse ecosystems. Our objective was to define the structure of the transition matrix of RMM and the proportion estimators of its transition rates. We reckon that this initial step will make possible a diverse range of methodological developments, paralleling those of matrix modelling (Caswell, 2001).

The ability of the model to retrodict forest demography can be instrumental in ecosystem management (Tester et al., 1997). Effective ecosystem management requires to understand ecosystem dynamics over long time horizons, including past historical dynamics (Beller et al., 2020). RMM can make the link between past, current and future population dynamics, thus contributing to historical ecology. RMM can specifically be employed to assess the success of forest restoration initiatives (Melichar et al., 2022). By comparing retrodicted and observed outcomes, practitioners can refine and adapt restoration strategies, ensuring that efforts align with predesignated ecological goals.

RMM's ability to project backward in time will also be useful for biodiversity conservation (Heppell et al., 2000). Conservationists can combine retrodictions and projections to assess the impact of management practices on different species and evaluate the efficacy of conservation interventions over extended periods (Crone et al., 2013). Populations may be structured by age or stage rather than size, leading to Leslie or Lefkovitch transition matrices instead of Usher matrices. The same reasoning used to define RMM for size-structured populations applies to age- or stage-structured populations. We also reckon that specific matrix modelling methods developed for species-rich ecosystems, like tropical rainforests or coral reefs, will be transferable to RMM (Mortier et al., 2015).

Specifically, the intrinsic rate of increase of a population given by the dominant eigenvalue  $\lambda$  of the transition matrix is a key instrument in population viability analysis (Morris & Doak, 2002, Chap. 7). In the case study at Paracou, the long-term retrodicted dynamics of the population was also determined by the dominant eigenvalue of the reverse transition matrix.  $\lambda < 1$  in RMM corresponds to a growing population in the chronological order. However, the properties of  $\lambda$  in matrix modelling result from the Perron-Frobenius theorem that stands for non-negative matrices. Because the reverse transition matrix **V** has negative terms, this theorem no longer applies to RMM. Further mathematical research will be needed to relate the dominant eigenvalue of **V** to the long-term behaviour of RMM and develop life table response experiments for RMM (Caswell, 1989).

Furthermore, RMM will be useful to assess global change impact. Given the dynamic nature of ecosystems and the influence of external factors such as climate change and land-use/land-cover change (Claeys et al., 2019), RMM provides a valuable tool for retroactively assessing the impacts on forest dynamics. It can also assist in reconstructing past trajectories of forest carbon stocks, thereby enhancing our understanding of the current and future capacity of forests to mitigate climate change (Gasser et al., 2020; Magerl et al., 2019). By incorporating relevant environmental variables into the model as it has been done for matrix models (Liang et al., 2011), RMM can contribute to retrodicting how climate change has altered forest composition, structure and overall health.

In forestry, the application of RMM can aid in optimizing timber harvesting strategies and implement sustainable and economically viable management (Lin et al., 1996). The case study at Paracou illustrated how to fit a RMM to data to retrodict dynamics triggered by logging. The nonlinear post-logging dynamics and the dependence of the transition matrices on disturbance intensity suggested that a density-dependent RMM may be more appropriate than a RMM with fixed coefficients. This conclusion is consistent with Favrichon (1998) who fitted a density-dependent forward matrix model at Paracou. The same techniques used in matrix

Methods in Ecology and Evolution

2041210x, 0, Downloaded from https

//besjournals

.onlinelibrary.wiley

.com/doi/10.1111/2041-210X.14426 by Nicolas Picard

CIRAD, Wiley Online Library on [24/09/2024]. See the Terms

and Co

(http

on Wiley

Online Library

for rule

of use; OA

. articles

are

by the :

applicable Creative Con

modelling to introduce density-dependence can be transferred to RMM (Caswell, 2001, Chap. 16; Solomon et al., 1986).

RMM's capability to retrodict past states and identify potential divergence points enhances its utility in risk assessment (Charles et al., 2009). Forest managers can use this information to evaluate resilience plans, preparing for and mitigating the impact of disturbances such as pests, diseases, or extreme weather events (Martinez-Ramos & Cristiän, 1997). Environmental stochasticity has been introduced in matrix modelling to account for natural disturbances, using matrix resampling, parametric matrix distribution, or stochastic shocks (Dalgleish et al., 2010; Davison et al., 2010). We reckon that the same techniques will be transferable to RMM.

Finally, the transparency of the RMM modelling process and its visualization through forward and backward projections make it an excellent tool for educational purposes. Communicating complex ecological concepts to various stakeholders, including policymakers and the general public, can foster a broader understanding of the importance of sustainable forestry and conservation practices (Arrontes, 2021; Mugasha et al., 2017).

#### 4.2 | Inversion of predictions

When a forward matrix model is available, retrodiction can be achieved by inverting the function that maps the state at time tto the state at time t + 1. Inversion methods are thus analytical methods in the mathematical sense of the term. However, with the exception of RMM that offers interpretable parameters, they tell nothing about the backward transitions. The various inversion methods matched the predictions more or less well. Nonetheless, the numerical error of the inversion should not be confused with the prediction error of the model (Diele & Marangi, 2019; Petrovskii & Petrovskaya, 2012). If a forward matrix model has a poor predictive performance, matching it closely with a good inversion method will still bring a poor retrodiction. For that reason, we still recommend RMM as an inversion method even if it diverged sooner from the forward dynamics than the other inversion methods.

Inversion methods succeeded to retro-project the forest dynamics predicted by Lin et al.'s matrix model over several decades. These results will have to be confirmed with other matrix models. We reckon that these inversion methods will remain efficient as long as the matrix model is not too much nonlinear (i.e. if the  $\phi$ function that maps **N** onto **O** is smooth enough). Machine-learning methods (such as random-forest) are increasingly being used to fit density-dependent relationships in forest dynamics models and bring greater nonlinearity in those relationships (Liu et al., 2018; Ma et al., 2020).

When the starting point of the backward projection gets closer to the stationary state of the matrix model, the back-projection gets less and less able to retrodict when the past dynamics was still transient. As different starting points converge to the same steady state, knowing the ending point (i.e. the steady state) brings no information about the starting point (Rupprecht & Vural, 2018). In other words, moving towards a stationary state results in the loss of some information and then there is an inherent limit to retrodiction due to the information lost on the way forward (Ellison et al., 2009; Rupprecht & Vural, 2018). Practically, before doing retrodiction, it may be useful to map the attractors of a matrix models in its phase space as we did in Figure 5, so as to check where the starting point of back projection is located with respect to the steady state of the model.

When using RMM as an inversion method, the backward transition matrix obtained is not exactly the inverse of the forward transition matrix. The product of **G** and **C** rather is a tridiagonal matrix (Supporting Information S4). Takada and Hara (1994) showed that such tridiagonal transition can be interpreted as the transition matrix of a diffusion process. Consider the *i*th size class that contain a given number of individuals. If **C** was the inverse of **G**, computing one step forward and one step backward would result in all individuals staying in class *i*. With the tridiagonal matrix, most of the individuals are still in class *i* but a small proportion of them has also spread to classes *i* – 1 and *i* + 1. Diffusion terms in transition matrix can be introduced to account for growth variability (Takada & Hara, 1994). It remains to be seen if the off-diagonal terms in **GC** and **CG** are consistent with individual growth variability.

#### 5 | CONCLUSIONS

RMM is an approach to retrodict the matrix model used in ecology to project population dynamics. Its transition rates are similar to the transition probabilities of the retrodicted Markov chain associated to the life cycle of the population dynamics, with a major difference as regards mortality. RMM was able to retrodict the dynamics of a disturbed tropical rainforest and back-project the dynamics predicted by Lin et al.'s density-dependent matrix model over a few decades. Inversion methods based on the inversion of the recurrence relationship of the matrix model were able to back-project predictions over a longer period of time. Because a model is a simplified representation of reality, when having data to fit a model without knowing the true forest dynamics, it makes more sense to fit a RMM than to fit a matrix model and inverse it.

#### AUTHOR CONTRIBUTIONS

All authors conceived the ideas. Nicolas Picard and Jack Andrew Hansen designed the methods and analysed the data. Nicolas Picard and Jingjing Liang led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

#### ACKNOWLEDGEMENTS

This study was supported by Grant Number: 40004269, "Quantifying global forest carbon dynamics using MATRIX model", World Resources Institute and the Department of Forestry and Natural Resources, Purdue University. We thank Guyafor and the Paracou team for giving access to the Paracou data.

#### CONFLICT OF INTEREST STATEMENT

The authors declares no conflict of interests.

#### PEER REVIEW

The peer review history for this article is available at https://www. webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14426.

#### DATA AVAILABILITY STATEMENT

The code used in this paper can be found in Zenodo via https:// zenodo.org/doi/10.5281/zenodo.13351330 (Picard, 2024). The Paracou data are available at Guyafor, DataBase of the French Guyana Permanent Plot Network, Cirad-CNRS-ONF (https://parac ou.cirad.fr/website/resources).

#### ORCID

Nicolas Picard D https://orcid.org/0000-0001-5548-9171

#### REFERENCES

- Ackleh, A. S., Deng, K., & Thibodeaux, J. J. (2007). A monotone approximation for a size-structured population model with a generalized environment. *Journal of Biological Dynamics*, 1(4), 305–319. https:// doi.org/10.1080/17513750701605564
- Alvarez-Buylla, E. R. (1994). Density dependence and patch dynamics in tropical rain forests: Matrix models and applications to a tree species. *The American Naturalist*, 143(1), 155–191. https://doi.org/10. 1086/285599
- Arrontes, J. (2021). Demography\_Lab, an educational application to evaluate population growth: Unstructured and matrix models. *Ecology* and Evolution, 11(5), 1940–1956. https://doi.org/10.1002/ece3. 7170
- Barth, M. A. F., Larson, A. J., & Lutz, J. A. (2015). A forest reconstruction model to assess changes to Sierra Nevada mixed-conifer forest during the fire suppression era. *Forest Ecology and Management*, 354, 104–118. https://doi.org/10.1016/j.foreco.2015.06.030
- Bell, T. J., Powell, K. I., & Bowles, M. L. (2013). Viability model choice affects projection accuracy and reintroduction decisions. *Journal of Wildlife Management*, 77(6), 1104–1113. https://doi.org/10.1002/ jwmg.525
- Beller, E. E., McClenachan, L., Zavaleta, E. S., & Larsen, L. G. (2020). Past forward: Recommendations from historical ecology for ecosystem management. *Global Ecology and Conservation*, 21, e00836. https:// doi.org/10.1016/j.gecco.2019.e00836
- Bonnicksen, T. M., & Stone, E. C. (1982). Reconstrution of a presettlement giant sequoia-mixed conifer forest community using the aggregation approach. *Ecology*, 63(4), 1134–1148. https://doi.org/10. 2307/1937250
- Bruner, H. D., & Moser, J. W. (1973). A Markov chain approach to the prediction of diameter distributions in uneven-aged forest stands. *Canadian Journal of Forest Research*, 3(3), 409–417. https://doi.org/ 10.1139/x73-059
- Caswell, H. (1989). The analysis of life table response experiments. I. Decomposition of effects on population growth rate. *Ecological Modelling*, 46, 221–237. https://doi.org/10.1016/0304-3800(89) 90019-7
- Caswell, H. (2001). Matrix population models: Construction, analysis and interpretation (2nd ed.). Sinauer Associates, Inc.
- Charles, S., Billoir, E., Lopes, C., & Chaumot, A. (2009). Matrix population models as relevant modeling tools in ecotoxicology. In J. Devillers (Ed.), Ecotoxicology modeling, no. 2 in emerging topics in ecotoxicology

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14426 by Nicolas Picard - CIRAD, Wiley Online Library on [24/09/2024]. See the Terms and Conditi (http: ary.wiley on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons

(pp. 261-298). Springer. https://doi.org/10.1007/978-1-4419-0197-2\_10

- Claeys, F., Gourlet-Fleury, S., Picard, N., Ouédraogo, D. Y., Tadesse, M. G., Hérault, B., Baya, F., Bénédet, F., Cornu, G., & Mortier, F. (2019). Climate change would lead to a sharp acceleration of central African forests dynamics by the end of the century. *Environmental Research Letters*, 14(4), 044002. https://doi.org/ 10.1088/1748-9326/aafb81
- Crone, E. E., Ellis, M. M., Morris, W. F., Stanley, A., Bell, T., Bierzychudek, P., Ehrlen, J., Kaye, T. N., Knight, T. M., Lesica, P., Oostermeijer, G., Quintana-Ascencio, P. F., Ticktin, T., Valverde, T., Williams, J. L., Doak, D. F., Ganesan, R., McEachern, K., Thorpe, A. S., & Menges, E. S. (2013). Ability of matrix models to explain the past and predict the future of plant populations. *Conservation Biology*, *27*(5), 968–978. https://doi.org/10.1111/cobi.12049
- Dalgleish, H. J., Koons, D. N., & Adler, P. B. (2010). Can life-history traits predict the response of forb populations to changes in climate variability? *Journal of Ecology*, *98*(1), 209–217. https://doi.org/10. 1111/j.1365-2745.2009.01585.x
- Davison, R., Jacquemyn, H., Adriaens, D., Honnay, O., de Kroon, H., & Tuljapurkar, S. (2010). Demographic effects of extreme weather events on a short-lived calcareous grassland species: Stochastic life table response experiments. *Journal of Ecology*, *98*(2), 255–267. https://doi.org/10.1111/j.1365-2745.2009.01611.x
- Diele, F., & Marangi, C. (2019). Geometric numerical integration in ecological modelling. *Mathematics*, 8(1), 25. https://doi.org/10.3390/ math8010025
- Ellison, C. J., Mahoney, J. R., & Crutchfield, J. P. (2009). Prediction, retrodiction, and the amount of information stored in the present. *Journal of Statistical Physics*, 136, 1005–1034. https://doi.org/10. 1007/s10955-009-9808-z
- Favrichon, V. (1998). Modeling the dynamics and species composition of tropical mixed-species uneven-aged natural forest: Effects of alternative cutting regimes. *Forest Science*, 44(1), 113–124. https://doi. org/10.1093/forestscience/44.1.113
- Gasser, T., Crepin, L., Quilcaille, Y., Houghton, R. A., Ciais, P., & Obersteiner, M. (2020). Historical CO<sub>2</sub> emissions from land use and land cover change and their uncertainty. *Biogeosciences*, 17(15), 4075–4101. https://doi.org/10.5194/bg-17-4075-2020
- Gill, M. S., Lemey, P., Bennett, S. N., Biek, R., & Suchard, M. A. (2016). Understanding past population dynamics: Bayesian coalescentbased modeling with covariates. *Systematic Biology*, 65(6), 1041– 1056. https://doi.org/10.1093/sysbio/syw050
- Gourlet-Fleury, S., Guehl, J. M., & Laroussinie, O. (Eds.). (2004). Ecology and management of a neotropical rainforest. Lessons drawn from Paracou, a long-term experimental research site in French Guiana. Elsevier.
- Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., de Bruin, S., Farina, M., Fatoyinbo, L., Hansen, M. C., Herold, M., Houghton, R. A., Potapov, P. V., Suarez, D. R., Roman-Cuesta, R. M., Saatchi, S. S., Slay, C. M., Turubanova, S. A., & Tyukavina, A. (2021). Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*, 11, 234–240. https://doi.org/10.1038/s41558-020-00976-6
- Heppell, S. S., Crouse, D. T., & Crowder, L. B. (2000). Using matrix models to focus research and management efforts in conservation, (Chap. 10, pp. 148–168). Springer. https://doi.org/10.1007/0-387-22648-6\_ 10
- Hudson, R. R. (1990). Gene genealogies and the coalescent process. Oxford Surveys in Evolutionary Biology, 7(1), 1–44.
- Ingram, D., & Buongiorno, J. (1996). Income and diversity tradeoffs from management of mixed lowland dipterocarps in Malaysia. *Journal of Tropical Forest Science*, 9(2), 242–270. https://www.jstor.org/stable/23616373
- Ingram, D. L. (2012). Life cycle assessment of a field-grown red maple tree to estimate its carbon footprint components. *The International*

Journal of Life Cycle Assessment, 17, 453-462. https://doi.org/10. 1007/s11367-012-0398-7

- Kingman, J. F. C. (1982). The coalescent. Stochastic Processes and their Applications, 13(3), 235–248. https://doi.org/10.1016/0304-4149(82)90011-4
- Leslie, P. H. (1945). On the use of matrices in certain population mathematics. *Biometrika*, 33(3), 183–212. https://doi.org/10.2307/2332297
- Lewis, E. G. (1942). On the generation and growth of a population. *Sankhya*, 6(1), 93–96.
- Liang, J. (2012). Mapping large-scale forest dynamics: A geospatial approach. Landscape Ecology, 27, 1091–1108. https://doi.org/10. 1007/s10980-012-9767-7
- Liang, J., & Picard, N. (2013). Matrix model of forest dynamics: An overview and outlook. Forest Science, 59(3), 359–378. https://doi.org/ 10.5849/forsci.11-123
- Liang, J., Zhou, M., Verbyla, D., Zhang, L., Springsteen, A. L., & Malone, T. (2011). Mapping forest dynamics under climate change: A matrix model. Forest Ecology and Management, 262(12), 2250–2262. https://doi.org/10.1016/j.foreco.2011.08.017
- Lin, C. R., Buongiorno, J., & Vasievich, M. (1996). A multi-species, densitydependent matrix growth model to predict tree diversity and income in northern hardwood stands. *Ecological Modelling*, 91(1-3), 193-211. https://doi.org/10.1016/0304-3800(95)00190-5
- Liu, Z., Peng, C., Work, T., Candau, J.-N., DesRochers, A., & Kneeshaw, D. (2018). Application of machine-learning methods in forest ecology: Recent progress and future challenges. *Environmental Reviews*, 26(4), 339–350. https://doi.org/10.1139/er-2018-0034
- Ma, W., Liang, J., Cumming, J. R., Lee, E., Welsh, A. B., Watson, J. V., & Zhou, M. (2016). Fundamental shifts of central hardwood forests under climate change. *Ecological Modelling*, 332, 28–41. https://doi. org/10.1016/j.ecolmodel.2016.03.021
- Ma, W., Lin, G., & Liang, J. (2020). Estimating dynamics of central hardwood forests using random forests. *Ecological Modelling*, 419, 108947. https://doi.org/10.1016/j.ecolmodel.2020.108947
- Magerl, A., Le Noë, J., Erb, K.-H., Bhan, M., & Gingrich, S. (2019). A comprehensive data-based assessment of forest ecosystem carbon stocks in the US 1907–2012. Environmental Research Letters, 14(12), 125015. https://doi.org/10.1088/1748-9326/ab5cb6
- Magnus, J. R., & Neudecker, H. (2007). Matrix differential calculus with applications in statistics and econometrics (3rd ed.). John Wiley and Sons.
- Manning, A. D., Lindenmayer, D. B., & Fischer, J. (2006). Stretch goals and backcasting: Approaches for overcoming barriers to large-scale ecological restoration. *Restoration Ecology*, 14(4), 487–492. https:// doi.org/10.1111/j.1526-100X.2006.00159.x
- Martinez-Ramos, M., & Cristiän, S. K. (1997). Tree life history patterns and forest dynamics: A conceptual model for the study of plant demography in patchy environments. *Journal of Sustainable Forestry*, 6(1–2), 85–125. https://doi.org/10.1300/J091v06n01\_07
- McDowell, N. G., Allen, C. D., Anderson-Teixeira, K., Aukema, B. H., Bond-Lamberty, B., Chini, L., Clark, J. S., Dietze, M., Grossiord, C., Hanbury-Brown, A., Hurtt, G. C., Jackson, R. B., Johnson, D. J., Kueppers, L., Lichstein, J. W., Ogle, K., Poulter, B., Pugh, T. A. M., Seidl, R., ... Xu, C. (2020). Pervasive shifts in forest dynamics in a changing world. *Science*, 368(6494), eaaz9463. https://doi.org/10. 1126/science.aaz9463
- Melichar, J., Cienciala, E., Albert, J., Braun Kohlová, M., Máca, V., & Pavelčík, P. (2022). Dynamics and management of restored forests in post-mining sites with respect to their recreation value: A matrix growth model. *Forests*, 13(9), 1519. https://doi.org/10.3390/f1309 1519
- Morris, W. F., & Doak, D. F. (2002). Quantitative conservation biology: Theory and practice of population viability analysis. Sinauer Associates, Inc.
- Mortier, F., Ouédraogo, D. Y., Claeys, F., Tadesse, M. G., Cornu, G., Baya, F., Benedet, F., Freycon, V., Gourlet-Fleury, S., & Picard, N. (2015). Mixture of inhomogeneous matrix models for species-rich

ecosystems. Environmetrics, 26(1), 39-51. https://doi.org/10.1002/env.2320

Methods in Ecology and Evolution

- Mugasha, W. A., Bollandsås, O. M., Gobakken, T., Zahabu, E., Katani, J. Z., & Eid, T. (2017). Decision-support tool for management of miombo woodlands: A matrix model approach. *Southern Forests: a Journal* of Forest Science, 79(1), 65–77. https://doi.org/10.2989/20702620. 2016.1233776
- Ohse, B., Compagnoni, A., Farrior, C. E., McMahon, S. M., Salguero-Gómez, R., Rüger, N., & Knight, T. M. (2023). Demographic synthesis for global tree species conservation. *Trends in Ecology & Evolution*, 38(6), 579–590. https://doi.org/10.1016/j.tree.2023.01.013
- Petrovskii, S., & Petrovskaya, N. (2012). Computational ecology as an emerging science. *Interface Focus*, 2(2), 241–254. https://doi.org/ 10.1098/rsfs.2011.0083
- Picard, N. (2024). R code for the retrodiction of population matrix models (v1). Zenodo. https://doi.org/10.5281/zenodo.13351330
- Picard, N., & Liang, J. (2014). Matrix models for size-structured populations: Unrealistic fast growth or simply diffusion? *PLoS One*, *9*(6), e98254. https://doi.org/10.1371/journal.pone.0098254
- Porté, A., & Bartelink, H. H. (2002). Modelling mixed forest growth: A review of models for forest management. *Ecological Modelling*, 150(1–2), 141–188. https://doi.org/10.1016/S0304-3800(01)00476-8
- Pretzsch, H. (2009). Forest dynamics, growth and yield: From measurement to model. Springer-Verlag. https://doi.org/10.1007/978-3-540-88307-4
- Price, O., & Bowman, D. M. J. S. (1994). Fire-stick forestry: A matrix model in support of skilful fire management of *Callitris intratropica* R. T. Baker by North Australian aborigenes. *Journal of Biogeography*, 21(6), 573–580. https://doi.org/10.2307/2846032
- Pugh, T. A. M., Rademacher, T., Shafer, S. L., Steinkamp, J., Barichivich, J., Beckage, B., Haverd, V., Harper, A., Heinke, J., Nishina, K., Rammig, A., Sato, H., Arneth, A., Hantson, S., Hickler, T., Kautz, M., Quesada, B., Smith, B., & Thonicke, K. (2020). Understanding the uncertainty in global forest carbon turnover. *Biogeosciences*, 17(15), 3961–3989. https://doi.org/10.5194/bg-17-3961-2020
- Roberts, M. R., & Hruska, A. L. (1986). Predicting diameter distribution: A test of the stationary Markov model. *Canadian Journal of Forest Research*, 16(1), 130–135.
- Rueda-Cediel, P., Anderson, K. E., Regan, T. J., Franklin, J., & Regan, H. M. (2015). Combined influences of model choice, data quality, and data quantity when estimating population trends. *PLoS One*, 10(7), e0132255. https://doi.org/10.1371/journal.pone.0132255
- Rupprecht, N., & Vural, D. C. (2018). Limits on inferring the past. Physical Review E, 97(6), 062155. https://doi.org/10.1103/PhysRevE.97.062155
- Salguero-Gómez, R., Jones, O. R., Archer, C. R., Bein, C., de Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Römer, G., Ruoff, T., Sommer, V., Wille, J., Voigt, J., Zeh, S., Vieregg, D., Buckley, Y. M., Che-Castaldo, J., ... Vaupel, J. W. (2016). COMADRE: A global data base of animal demography. *Journal of Animal Ecology*, 85(2), 371–384. https://doi.org/10.1111/1365-2656.12482
- Salguero-Gómez, R., Jones, O. R., Archer, C. R., Buckley, Y. M., Che-Castaldo, J., Caswell, H., Hodgson, D., Scheuerlein, A., Conde, D. A., Brinks, E., de Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Römer, G., Runge, J., Ruoff, T., ... Vaupel, J. W. (2015). The COMPADRE plant matrix database: An open online repository for plant demography. *Journal of Ecology*, 103(1), 202–218. https://doi.org/10.1111/1365-2745.12334
- Solomon, D. S., Hosmer, R. A., & Hayslett, H. T. (1986). A two-stage matrix model for predicting growth of forest stands in the northeast. *Canadian Journal of Forest Research*, 16(3), 521–528.
- Surace, J., & Scandi, M. (2023). State retrieval beyond Bayes' retrodiction. Quantum, 7, 990. https://doi.org/10.22331/q-2023-04-27-990
- Takada, T., & Hara, T. (1994). The relationship between the transition matrix model and the diffusion model. *Journal of Mathematical Biology*, 32, 789–807. https://doi.org/10.1007/BF00168798

14

- Tester, J. R., Starfield, A. M., & Frelich, L. E. (1997). Modeling for ecosystem management in Minnesota pine forests. *Biological Conservation*, 80(3), 313–324. https://doi.org/10.1016/S0006-3207(96)00069-9
- Usher, M. B. (1966). A matrix approach to the management of renewable resources, with special reference to selection forests. *Journal of Applied Ecology*, *3*(2), 355–367. https://doi.org/10.2307/2401258
- Usher, M. B. (1969). A matrix model for forest management. *Biometrics*, 25(2), 309–315. https://doi.org/10.2307/2528791
- Zhou, M., Liang, J., & Buongiorno, J. (2008). Adaptive versus fixed policies for economic or ecological objectives in forest management. *Forest Ecology and Management*, 254(2), 178–187. https://doi.org/ 10.1016/j.foreco.2007.07.035

#### SUPPORTING INFORMATION

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

Supporting Information S1. List of symbols.

**Supporting Information S2.** Invertibility of a Usher transition matrix. **Supporting Information S3.** Jacobian of the density-dependent matrix model of Lin et al. (1996).

**Supporting Information S4.** Product of a forward and backward transition matrix.

**Supporting Information S5.** Backward transition matrices at Paracou. **Supporting Information S6.** Retrodicted forest dynamics at Paracou, French Guiana.

**Supporting Information S7.** Inversion of Lin et al.'s densitydependent matrix model.

**Supporting Information S8.** Interactive three-dimensional view of the projections of the model by Lin et al. (1996).

Supporting Information S9. Comments on the R code.

Appendix A. RMM using retrodiction of Markov chains.

How to cite this article: Picard, N., Hansen, J. A., Harris, N., & Liang, J. (2024). Retrodiction of forest demography: Backward simulation with reverse matrix models. *Methods in Ecology and Evolution*, 00, 1–14. <u>https://doi.org/10.1111/2041-210X.14426</u>

# Appendix A. RMM using retrodiction of Markov chains

# Retrodiction of forest demography: backward simulation with reverse matrix models

The life cycle interpretation of matrix models does not fully comply with a Markov chain because of reproduction or regeneration. At least, the growth and death transitions of the life cycle can be interpreted as a Markov chain. Therefore, let us consider a cohort of trees that grow or die without regenerating. The 'dead' state is added to the diameter classes so that the set of possible Markov chain transitions is complete. Referring to Figure 1, let  $X_t$  be the state of a tree at time t, which can be the *i*th diameter class  $(X_t = i)$  or death  $(X_t = \dagger)$ . Retrodiction of Markov chains commonly relies on Bayes' theorem (Surace & Scandi, 2023). For the reverse growth rate  $b_i$ , Bayes's theorem gives:

$$b_i \equiv \Pr(X_t = i - 1 | X_{t+1} = i) = \frac{\Pr(X_{t+1} = i | X_t = i - 1) \,\Pr(X_t = i - 1)}{\Pr(X_{t+1} = i)}$$

By definition, the probability  $\Pr(X_{t+1} = i | X_t = i - 1)$  identifies with  $p_{i-1}$ , and  $\Pr(X_t = j)$  for any diameter class j is proportional to  $N_j$ . Moreover, by the law of total probability:

$$\Pr(X_{t+1} = i) = \sum_{j} \Pr(X_{t+1} = i | X_t = j) \, \Pr(X_t = j)$$

For  $i \ge 2$ , the life cycle (Figure 1) shows that there are only two possible transitions towards diameter class i, so:

$$\Pr(X_{t+1} = i) = \Pr(X_{t+1} = i | X_t = i - 1) \ \Pr(X_t = i - 1) + \Pr(X_{t+1} = i | X_t = i) \ \Pr(X_t = i)$$

Therefore:

$$b_i = \frac{p_{i-1}N_{i-1}}{p_{i-1}N_{i-1} + q_iN_i} = \frac{p_{i-1}}{p_{i-1} + q_i\eta_i} \tag{1}$$

where  $\eta_i = N_i/N_{i-1}$ . Similar calculations for the stasis rate bring:  $r_i \equiv \Pr(X_t = i | X_{t+1} = i) = q_i/(q_i + p_{i-1}/\eta_i)$  for  $i \geq 2$  and  $r_1 = 1$ . The two transition rates sum to one:  $r_i + b_i = 1$ . The difference between RMM and the retrodiction of the Markov chain lies in the mortality process. The stochastic matrix of the Markov chain associated to the life cycle shown in Figure 1 and the stochastic matrix of the retrodicted Markov chain are respectively:

$$\begin{bmatrix} & & 0 \\ G & \vdots \\ 0 \\ \hline m_1 & \dots & m_K & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} & & d_1 \\ \vdots \\ 0 \\ \hline 0 & \dots & 0 & d_{\dagger} \end{bmatrix}$$

where  $d_i \equiv \Pr(X_t = i | X_{t+1} = \dagger)$  for  $i \leq K$  are the revival rates and  $d_{\dagger} \equiv \Pr(X_t = \dagger | X_{t+1} = \dagger)$ . Again, the application of Bayes' theorem gives:  $d_i = m_i w_i / (w_{\dagger} + \sum_{j=1}^K m_j w_j)$  and  $d_{\dagger} = w_{\dagger} / (w_{\dagger} + \sum_{j=1}^K m_j w_j)$ , where  $w_i = N_i / (N_{\dagger} + \sum_{j=1}^K N_j)$  and  $w_{\dagger} = N_{\dagger} / (N_{\dagger} + \sum_{j=1}^K N_j)$ . In the matrix model, the projection of the number  $N_i$  of trees in each of the K diameter classes does not require to know the number  $N_{\dagger}$  of dead trees. This is no longer true with the retrodiction of the Markov chain. RMM can thus be seen as an adjustment of the retrodiction of the Markov chain so that the number of dead trees is not needed.

# **Supporting Information**

Retrodiction of forest demography: backward simulation with reverse matrix models

# S1. List of symbols

The length of vectors and dimensions of matrices are given for a single population with K size classes. If there are S populations, as in the matrix model by Lin et al. (1996) where S = 3, K has to be replaced by SK.

- $\otimes$ : Kronecker product of two matrices
- $\odot$ : Hadamard product of two matrices
- †: dead state of an individual in its life cycle
- $\mathbf{0}_n$ : vector of length *n* full of zeros
- $\mathbf{1}_n$ : vector of length *n* full of ones
- $\alpha_{gk}$ : coefficients of the linear relationship between the recruitment rates, basal area and tree density in the matrix model of Lin et al. (1996)
- $\boldsymbol{\alpha}_k$ : vector that collates the  $\alpha_{gk}$ s
- $\beta_g$ : slope of the linear relationship between the upgrowth rates and basal area in the matrix model of Lin et al. (1996)
- $\beta_{gi}$ : intercept of the linear relationship between the upgrowth rates and basal area in the matrix model of Lin et al. (1996)
- $\beta_0$ : vector that collates the  $\beta_{gi}$ s
- $\beta_1$ : vector that collates the  $\beta_g$ s
- $\delta$ : parameters related to mortality in the matrix model of Lin et al. (1996)
- $\eta_i$ : ratio of the number of individuals in size class i at time t over the number of individuals in size class i 1 at time t
- $\lambda$ : dominant eigenvalue of a transition matrix
- $\tau$ : duration of agreement
- $\phi$ : recurrence relationship of the matrix model, i.e. application that maps N onto G(N) N + R(N)
- $A_K$ : square matrix of dimension K obtained from a Usher transition matrix of dimension K + 1 by removing its first row and first column
  - $b_i$ : reverse growth rate of the *i*th size class of the reverse matrix model, i.e. conditional probability that an individual was in the (i-1)th size class at time t knowing that it is in the *i*th size class at time t+1
  - B: basal area of a forest stand
  - B: vector of length K giving the average basal area of trees in each diameter class
- $\boldsymbol{B}_n$ : vector of length nK that repeats  $\boldsymbol{B}$  n times
- C: transition matrix of the reverse matrix model ( $K \times K$  matrix)
- $d_i$ : revival rate of the *i*th size class of the reverse matrix model, i.e. conditional probability that an individual was in the *i*th size class at time t knowing that it is dead at time t + 1

- **D**: Jacobian matrix of the application that maps N onto R(N), i.e.  $K \times K$  matrix whose *ij*-th term is the partial derivative of the *i*th element of R with respect to the *j*th element of N
- $D_{\phi}$ : Jacobian matrix of  $\phi$ , i.e.  $K \times K$  matrix whose *ij*-th element is the partial derivative of the *i*th element of  $\phi(N)$  with respect to the *j*th element of N
  - **E**: vector of length K composed of K 1 ones followed by zero
  - f: average fecundity rate, i.e. ratio of the total number of offspring over the total number of living trees
  - $f_i$ : fecundity rate in the *i*th size class of the matrix model, i.e. expected number of offspring generated by an individual in the *i*th size class between times t and t + 1
  - f: vector of fecundities (vector of length K whose *i*th element is  $f_i$ )
  - g: index of species group in the matrix model of Lin et al. (1996) G: transition matrix of the matrix model  $(K \times K \text{ matrix})$
  - h: average fecundity rate of the reverse matrix model
  - i: index of a size class = one of the states of the life cycle of an individual in a size-structured population
- $I_n: n \times n$  identity matrix
- J: Jacobian matrix of the application that maps N onto G(N), i.e.  $K^2 \times K$  matrix whose ij-th term is the partial derivative of the *i*th element of vec G with respect to the *j*th element of N
- K: number of size classes
- $m_i$ : mortality rate of the *i*th size class of the matrix model, i.e. conditional probability that an individual is dead at time t + 1 knowing that it was in the *i*th size class at time t
- m: vector that collates the  $m_i$ s
- $M_i$ : given a dataset of observed individuals at times t and t + 1, number of individuals in size class i that die between t and t + 1
- $N_i$ : number of living individuals in size class *i* at time *t*
- N: vector of the abundances of individuals in the size classes at time t (vector of length K)
- $O_i$ : number of living individuals in size class *i* at time t + 1
- **O**: vector of the abundances of individuals in the size classes at time t + 1 (vector of length K)
- $p_i$ : upgrowth rate of the *i*th size class of the matrix model, i.e. conditional probability that an individual is in the (i + 1)th size class at time t + 1 knowing that it was in the *i*th size class at time t
- p: vector that collates the  $p_i$ s
- $P_i$ : given a dataset of observed individuals at times t and t + 1, number of individuals that move up from size class i to i + 1 between t and t + 1
- $q_i$ : stasis rate of the *i*th size class of the matrix model, i.e. conditional probability that an individual is in the *i*th size class at time t + 1 knowing that it was in the *i*th size class at time t
- $Q_i$ : given a dataset of observed individuals at times t and t + 1, number of individuals that stay alive in size class i between t and t + 1
- $r_i$ : stasis rate of the *i*th size class of the reverse matrix model, i.e. conditional probability that an individual was in the *i*th size class at time t knowing that it is in the *i*th size class at time t + 1
- R: number of recruited trees between t and t+1

- $R_g$ : recruitment rate of species group g in the matrix model of Lin et al. (1996)
- $\mathbf{R}$ : recruitment vector (vector of length K)
- s: discrete-time index
- S: number of populations in a multi-population matrix model (e.g. S = 3 in the matrix model of Lin et al., 1996)
- t: discrete-time index
- $t_{\rm div}$ : divergence time
  - T: time horizon of the forward projection = starting time of the backward projection
- U: Usher transition matrix ( $K \times K$  matrix)
- V: reverse transition matrix that is analogue of the Usher transition matrix in the reverse case ( $K \times K$  matrix)
- $w_i$ : proportion of individuals in the *i*th size class at time t
- $W_n$ : square  $n \times n$  matrix full of zeros with the exception of its subdiagonal that is full of ones
- $X_t$ : random state at time t of an individual of a size-structured population when describing its life cycle by a Markov chain

## S2. Invertibility of a Usher transition matrix

Let U be a Usher transition matrix:

$$\boldsymbol{U} = \begin{bmatrix} q_1 + f_1 & f_2 & \dots & f_K \\ p_1 & q_2 & \boldsymbol{0} \\ & \ddots & \ddots \\ \boldsymbol{0} & p_{K-1} & q_K \end{bmatrix}$$

Using Laplace expansion along its first column, the determinant of  $\boldsymbol{U}$  can be computed as:

$$U| = (q_1 + f_1)q_2 \dots q_K - p_1 |A_{K-1}|$$
(1)

where  $A_{K-1}$  is the square matrix of dimension K-1:

$$\boldsymbol{A}_{K-1} = \begin{bmatrix} f_2 & f_3 & \dots & f_K \\ p_2 & q_3 & \boldsymbol{0} \\ & \ddots & \ddots \\ \boldsymbol{0} & p_{K-1} & q_K \end{bmatrix}$$

The following proposition specifies a sufficient condition to obtain a range for  $|A_{K-1}|$ .

**Proposition 1**: Let  $A_K$  be a square matrix of dimension K defined by

$$\boldsymbol{A}_{K} = \begin{bmatrix} f_{1} & f_{2} & \dots & f_{K} \\ p_{1} & q_{2} & & \boldsymbol{0} \\ & \ddots & \ddots & \\ \boldsymbol{0} & & p_{K-1} & q_{K} \end{bmatrix}$$

with  $0 < p_i < 1$ ,  $0 < q_i < 1$  and  $f_i > 0$ . Let  $f_{\min} = \min\{f_1, \ldots, f_K\}$ ,  $f_{\max} = \max\{f_1, \ldots, f_K\}$ ,  $q_{\min} = \min\{q_1, \ldots, q_K\}$  and  $p_{\min} = \min\{p_1, \ldots, p_{K-1}\}$ . If  $f_{\min}q_{\min} > f_{\max}p_{\max}$ , then  $0 < |\mathbf{A}_K| < f_{\max}q_2 \ldots q_K$ .

**Proof.** We prove the proposition by induction. For K = 2,

$$|\boldsymbol{A}_2| = \left| \begin{array}{cc} f_1 & f_2 \\ p_1 & q_2 \end{array} \right| = f_1 q_2 - f_2 p_1$$

Because  $f_2p_1 > 0$ ,  $|\mathbf{A}_2| < f_1q_2 \le f_{\max}q_2$ . Moreover,  $f_1q_2 \ge f_{\min}q_{\min}$  and  $f_{\max}p_{\max} \ge f_2p_1$ , therefore  $f_{\min}q_{\min} > f_{\max}p_{\max}$  implies  $f_1q_2 > f_2q_1$ , that is  $|\mathbf{A}_2| > 0$ .

Now assume that the proposition holds for 1, 2, ..., K - 1, and let's demonstrate it for K. Using Laplace expansion along the first column of  $A_K$ , the determinant of  $A_K$  is:

$$|\boldsymbol{A}_K| = f_1 q_2 \dots q_K - p_1 |\boldsymbol{A}_{K-1}|$$

Because  $|\mathbf{A}_{K-1}| > 0$ ,  $|\mathbf{A}_{K}| < f_1 q_2 \dots q_K \le f_{\max} q_2 \dots q_K$ . Moreover,  $|\mathbf{A}_{K-1}| < f_{\max} q_3 \dots q_K$ implies:  $-p_1 |\mathbf{A}_{K-1}| > -p_1 f_{\max} q_3 \dots q_K$ . Therefore,

$$\begin{aligned} |\mathbf{A}_K| &> f_1 q_2 \dots q_k - p_1 f_{\max} q_3 \dots q_K \\ |\mathbf{A}_K| &> (f_1 q_2 - p_1 f_{\max}) q_3 \dots q_k \end{aligned}$$

The condition  $f_{\min}q_{\min} > f_{\max}p_{\max}$  implies  $f_1q_2 - p_1f_{\max} > 0$ , and therefore  $|\mathbf{A}_K| > 0$ .

Turning back to eq. (1), we see that  $|\boldsymbol{U}| = q_1 \dots q_K + |\boldsymbol{A}_K|$ , so  $|\boldsymbol{A}_K| > 0$  readily implies  $|\boldsymbol{U}| > 0$ . Therefore,  $f_{\min}q_{\min} > f_{\max}p_{\max}$  is a sufficient condition for  $\boldsymbol{U}$  to be invertible. This condition is mild in real-life situations. Most of the time, it is not possible to estimate the fecundity rate of each class separately because it is not possible to trace back the parent tree of a newly recruited tree, so a common fecundity rate is used:  $f_1 = f_2 = \dots = f_K$ . Then  $f_{\min} = f_{\max}$ , and the condition boils down to comparing  $q_{\min}$  and  $p_{\max}$ . Because tree growth is slow, the probability to stay in a class is often close to 1, while the probability to move up from one class to the next one is often close to 0. Therefore, we naturally almost always have:  $q_i > p_j$  for all i and j, so that the condition is met.

In some cases, a distinction is made between reproducing and non-reproducing size classes, i.e. there is a threshold  $s \in [1, \ldots, K]$  such that  $f_i = 0$  for i < s. In this case,  $f_{\min} = 0$  and the condition  $f_{\min}q_{\min} > f_{\max}p_{\max}$  is not met. Nonetheless, U can still be invertible in this case.

# S3. Jacobian of the density-dependent matrix model of Lin et al. (1996)

We concatenate the vectors of the three species groups into a single vector of length 3K:  $N = [N_{\rm I}, N_{\rm II}, N_{\rm III}]$ . Given the vector B of length K that gives the average basal area of trees in each diameter class, let  $B_3 = \mathbf{1}_3 \otimes B$  the vector of length 3K that repeats Bthree times. Thus, the plot basal area writes  $B = {}^{\mathrm{t}}B_3N$ .

As regards the growth transition matrix, let  $\beta_g^0 = [\beta_{g1}, \ldots, \beta_{gK}]$  (with  $\beta_{gK} = 0$ ) and  $\beta_0 = [\beta_1^0, \beta_{II}^0, \beta_{III}^0]$ , and similarly  $p_g = [p_{g1}, \ldots, p_{gK}]$  and  $p = [p_I, p_{II}, p_{III}]$ , and  $m_g = [m_{g1}, \ldots, m_{gK}]$  and  $m = [m_I, m_{III}, m_{III}]$ . Let  $E = [\mathbf{1}_{K-1}, 0]$  and  $\beta_1 = [\beta_I E, \beta_{II} E, \beta_{III} E]$ , so that:  $p = \beta_0 + B\beta_1$ . The growth transition matrix G for the three species groups is the  $3K \times 3K$  matrix whose diagonal vector is  $\mathbf{1}_{3K} - m - p$  and whose subdiagonal vector is p (disregarding the last element of this vector). In other words:

$$m{G}(m{N}) = m{I}_{3K} \odot \{m{1}_{3K}{}^{ ext{t}}[m{1}_{3K} - m{m} - m{p}]\} + m{W}_{3K} \odot [m{1}_{3K}{}^{ ext{t}}m{p}]$$

where  $\odot$  denotes the Hadamard product (Magnus & Neudecker, 2007, p.53) and  $W_n$  is the square  $n \times n$  matrix full of zeros with the exception of its subdiagonal that is full of ones. Replacing p by  $\beta_0 + B\beta_1$  gives:

$$\begin{array}{lll} \boldsymbol{G}(\boldsymbol{N}) &=& \boldsymbol{I}_{3K} \odot \{ \boldsymbol{1}_{3K}^{\phantom{\dagger} \mathrm{t}} [ \boldsymbol{1}_{3K} - \boldsymbol{m} - \boldsymbol{\beta}_{0} ] \} + \boldsymbol{W}_{3K} \odot [ \boldsymbol{1}_{3K}^{\phantom{\dagger} \mathrm{t}} \boldsymbol{\beta}_{0} ] \\ &+ B\{ [ \boldsymbol{W}_{3K} - \boldsymbol{I}_{3K} ] \odot [ \boldsymbol{1}_{3K}^{\phantom{\dagger} \mathrm{t}} \boldsymbol{\beta}_{1} ] \} \end{array}$$

By definition, the Jacobian matrix J of G is the  $(3K)^2 \times 3K$  matrix whose ij-th term is the partial derivative of the *i*th element of vec G with respect to the *j*th element of N, where vec is the operator that stacks the columns of a matrix one underneath the other (Magnus & Neudecker, 2007, p.34). The first two terms in the above expression of G do not depend on N. The third term linearly depends on N. Therefore, the Jacobian matrix of G is constant:

$$oldsymbol{J} = \mathrm{vec}\left\{ \left[oldsymbol{W}_{3K} - oldsymbol{I}_{3K}
ight] \odot \left[oldsymbol{1}_{3K}{}^{\mathrm{t}}oldsymbol{eta}_{1}
ight]
ight\}{}^{\mathrm{t}}oldsymbol{B}_{3}$$

In the expression of J, the matrix within the vec operator simply is the square  $3K \times 3K$  matrix with diagonal  $-\beta_1$ , subdiagonal  $\beta_1$  (dropping the last element of this vector) and zero elsewhere.

As regards recruitment, let  $\mathbf{R} = [\mathbf{R}_{\text{I}}, \mathbf{R}_{\text{II}}, \mathbf{R}_{\text{III}}]$  be the vector of length 3K that concatenates the recruitment vectors of the three species groups. Let  $\boldsymbol{\alpha}_{gk}$  be the vector of length K defined as  $\alpha_{gk} [1, \mathbf{0}_{K-1}]$ , where  $\mathbf{0}_n$  is the vector of length n full of zeros, and let  $\boldsymbol{\alpha}_k = [\boldsymbol{\alpha}_{\text{Ik}}, \boldsymbol{\alpha}_{\text{IIk}}, \boldsymbol{\alpha}_{\text{IIIk}}]$  be the vector of length 3K that concatenates the  $\boldsymbol{\alpha}_{gk}$ s for k = 0, 1, 2. The recruitment vector thus is written as:

$$oldsymbol{R}(oldsymbol{N}) = oldsymbol{lpha}_0 + ({}^{\mathrm{t}}oldsymbol{B}_3oldsymbol{N}) \,oldsymbol{lpha}_1 + oldsymbol{lpha}_2 \odot \left[(oldsymbol{I}_3 \otimes {}^{\mathrm{t}}oldsymbol{1}_K \otimes oldsymbol{1}_K)oldsymbol{N}
ight]$$

By definition, the Jacobian matrix D of R is the  $3K \times 3K$  matrix whose *ij*-th term is the partial derivative of the *i*th element of R with respect to the *j*th element of N. The first term in the above expression of R does not depend on N. The second and third terms linearly depend on N. Therefore, the Jacobian matrix of R is constant:

$$oldsymbol{D} = oldsymbol{lpha}_1\,{}^{\mathrm{t}}oldsymbol{B}_3 + ({}^{\mathrm{t}}oldsymbol{1}_{3K}\otimesoldsymbol{lpha}_2)\odot(oldsymbol{I}_3\otimes{}^{\mathrm{t}}oldsymbol{1}_K\otimesoldsymbol{1}_K)$$

### S4. Product of a forward and backward transition matrix

Let G be a forward transition matrix as given by eq. (1) and let C be a backward transition matrix as given by eq. (2). Their product is:

$$\boldsymbol{GC} = \begin{bmatrix} q_1 r_1 & q_1 b_2 & 0 & \mathbf{0} \\ r_1 p_1 & (q_2 r_2 + p_1 b_2) & q_2 b_3 & & \\ 0 & r_2 p_2 & (q_3 r_3 + p_2 b_3) & \ddots & \\ & \ddots & \ddots & q_{K-1} b_K \\ \mathbf{0} & & & r_{K-1} p_{K-1} & (q_K r_K + p_{K-1} b_K) \end{bmatrix}$$
(2)

and

$$\boldsymbol{C}\boldsymbol{G} = \begin{bmatrix} (q_1r_1 + p_1b_2) & q_2b_2 & 0 & \boldsymbol{0} \\ r_2p_1 & (q_2r_2 + p_2b_3) & q_3b_3 & & & \\ 0 & r_3p_2 & \ddots & \ddots & & \\ & & \ddots & (q_{K-1}r_{K-1} + p_{K-1}b_K) & q_Kb_K \\ \boldsymbol{0} & & & r_Kp_{K-1} & q_Kr_K \end{bmatrix}$$
(3)

Matrix GC has a diagonal whose *i*th term is  $q_ir_i + p_{i-1}b_i$  (except the first term that is  $q_1r_1$ ), a subdiagonal whose *i*th term is  $r_ip_i$ , and a superdiagonal whose *i*th term is  $q_ib_{i+1}$ . The rest of the matrix is null. Matrix CG has a diagonal whose *i*th term is  $q_ir_i + p_ib_{i+1}$  (except the last term that is  $q_Kr_K$ ), a subdiagonal whose *i*th term is  $r_{i+1}p_i$ , and a superdiagonal whose *i*th term is  $q_ib_i$ . Using the expressions (6), one can show that  $q_1r_1 = 1 - p_1$ ,  $q_ir_i = (1 - p_i)(1 - b_i)$  for  $i = 2, \ldots, K - 1$  and  $q_Kr_K = 1 - b_K$ . Considering that  $p_i \ll 1$  and  $b_i \ll 1$ , these equations show that GC and CG are actually close to the identity matrix.

## S5. Backward transition matrices at Paracou

For each treatment, we give the observed number of trees in the 11 diameter class in 2001 ( $N_{obs}$ , on the left-hand side of the matrix relationship), the observed number of trees in the 11 diameter class in 2003 ( $O_{obs}$ , on the right-hand side), and the fitted backward transition matrix ( $V_{fitted}$ ). The estimators of the reverse transition rates given in Section 2.2 ensure that:  $N_{obs} = V_{fitted} O_{obs}$ ).

• Control plots:

4461 -		0.9976	0.0388	-0.0206	-0.0206	-0.0206	-0.0206	-0.0206	-0.0206	-0.0206	-0.0206	-0.0206	1 Г	4475 -
2386		0	0.9634	0.0690	0	0	0	0	0	0	0	0		2374
1450		0	0	0.9526	0.0884	0	0	0	0	0	0	0		1435
925		0	0	0	0.9276	0.0854	0	0	0	0	0	0		939
638		0	0	0	0	0.9335	0.0899	0	0	0	0	0		632
530	=	0	0	0	0	0	0.9307	0.0965	0	0	0	0		534
336		0	0	0	0	0	0	0.9211	0.0925	0	0	0		342
236		0	0	0	0	0	0	0	0.9207	0.1588	0	0		227
162		0	0	0	0	0	0	0	0	0.8765	0.1161	0		170
110		0	0	0	0	0	0	0	0	0	0.9107	0.0339		112
231		0	0	0	0	0	0	0	0	0	0	0.9788	ΙL	236 _

#### • Treatment 1:

F 4570 -	1	F 0.9771	0.0401	-0.0423	-0.0423	-0.0423	-0.0423	-0.0423	-0.0423	-0.0423	-0.0423	-0.0423 J	F 4772
2411		0	0.9378	0.1193	0	0	0	0	0	0	0	0	2380
1434		0	0	0.8934	0.1040	0	0	0	0	0	0	0	1501
884		0	0	0	0.9116	0.1087	0	0	0	0	0	0	894
640		0	0	0	0	0.9024	0.1354	0	0	0	0	0	635
473	=	0	0	0	0	0	0.8788	0.1271	0	0	0	0	495
295		0	0	0	0	0	0	0.8829	0.1372	0	0	0	299
218		0	0	0	0	0	0	0	0.8805	0.1348	0	0	226
133		0	0	0	0	0	0	0	0	0.8794	0.1047	0	141
85		0	0	0	0	0	0	0	0	0	0.9186	0.0357	86
L 166 _	]	L O	0	0	0	0	0	0	0	0	0	0.9881	L 168

• Treatment 2:

F 5443 <sup>.</sup>	1	0.9714	0.0656	-0.0491	-0.0491	-0.0491	-0.0491	-0.0491	-0.0491	-0.0491	-0.0491	-0.0491	Г	5633 .
2784		0	0.9038	0.1190	0	0	0	0	0	0	0	0		2859
1673		0	0	0.9024	0.1587	0	0	0	0	0	0	0		1680
950		0	0	0	0.8524	0.1524	0	0	0	0	0	0		989
679		0	0	0	0	0.8675	0.1580	0	0	0	0	0		702
423	=	0	0	0	0	0	0.8510	0.1673	0	0	0	0		443
262		0	0	0	0	0	0	0.8400	0.2109	0	0	0		275
130		0	0	0	0	0	0	0	0.8027	0.1558	0	0		147
72		0	0	0	0	0	0	0	0	0.8571	0.1429	0		77
44		0	0	0	0	0	0	0	0	0	0.8810	0.1429		42
L 44 .		L 0	0	0	0	0	0	0	0	0	0	0.8980	L	49

• Treatment 3:

F 6397 <sup>·</sup>	1	F 0.9665	0.0879	-0.0610	-0.0610	-0.0610	-0.0610	-0.0610	-0.0610	-0.0610	-0.0610	-0.0610 J	F 6600
2863		0	0.8683	0.1501	0	0	0	0	0	0	0	0	3022
1509		0	0	0.8662	0.1433	0	0	0	0	0	0	0	1592
877		0	0	0	0.8666	0.1527	0	0	0	0	0	0	907
560		0	0	0	0	0.8641	0.0939	0	0	0	0	0	596
488	=	0	0	0	0	0	0.9207	0.1950	0	0	0	0	479
212		0	0	0	0	0	0	0.8216	0.1429	0	0	0	241
100		0	0	0	0	0	0	0	0.8980	0.1905	0	0	98
61		0	0	0	0	0	0	0	0	0.8095	0.2326	0	63
42		0	0	0	0	0	0	0	0	0	0.7674	0.2432	43
L 29		L 0	0	0	0	0	0	0	0	0	0	0.7838	L 37

### S6. Retrodicted forest dynamics at Paracou, French Guiana



Fig. S1: Retrodiction of the tree density of the rainforest at Paracou, French Guiana, using a reverse matrix model fitted to the observations of 2001 and 2003. The black line is the model prediction. The colored line is the observed average trajectory and the shaded area is the observed envelope of the trajectories of twelve 1.5625-ha plots. These plots underwent silvicultural treatments between 1986 and 1988: (a) control plots (no treatment), (b) treatment 1 (low disturbance intensity), (c) treatment 2 (medium disturbance intensity), (d) treatment 3 (high disturbance intensity).

## S7. Inversion of Lin et al.'s density-dependent matrix model

# References

Lin, C. R., Buongiorno, J., & Vasievich, M. (1996). A multi-species, density-dependent matrix growth model to predict tree diversity and income in northern hardwood stands. *Ecological Modelling*, 91(1-3), 193–211.



Fig. S2: Forward and backward projections of the basal area of a forest using the densitydependent matrix model by Lin et al. (1996). The time horizon of forward projections varies from T = 100 years (leftmost column) to T = 700 years (rightmost column). Backward projections are made using four methods: numerical inversion (row denoted "num"), first-order approximation ("1-order"), zero-order approximation ("0-order"), Reverse Matrix Model ("RMM"). In each panel, the three lines correspond to the three species groups of Lin et al. (1996). Black lines are forward projections, colored lines are backward projections.



Fig. S3: Forward and backward projections of the tree density of a forest using the density-dependent matrix model by Lin et al. (1996). The time horizon of forward projections varies from T = 100 years (leftmost column) to T = 700 years (rightmost column). Backward projections are made using four methods: numerical inversion (row denoted "num"), first-order approximation ("1-order"), zero-order approximation ("0-order"), Reverse Matrix Model ("RMM"). In each panel, the three lines correspond to the three species groups of Lin et al. (1996). Black lines are forward projections, colored lines are backward projections.



Fig. S4: Root mean square error of the retrodicted diameter structure of a forest, where the error of the retrodicted structure is calculated with respect to the structure predicted by the density-dependent matrix model by Lin et al. (1996). The time horizon of forward projections varies from T = 100 years (leftmost column) to T = 700 years (rightmost column). Backward projections are made using four methods: numerical inversion (row denoted "num"), first-order approximation ("1-order"), zero-order approximation ("0-order"), Reverse Matrix Model ("RMM").

Magnus, J. R., & Neudecker, H. (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley series in probability and statistics. Chichester, UK: John Wiley and Sons, 3rd ed.

# **Supporting Information**

Retrodiction of forest demography: backward simulation with reverse matrix models

# S9. Comments on the R code

**backtrans.** The backtrans function computes a backward growth transition matrix from a forward growth transition matrix and the current diameter structure. For instance, a growth transition matrix for *Entandrophragma cylindricum* is (COMPADRE matrix #247558, https://compadre-db.org/Species/47283):

x <- matrix(c(0.972117, 0, 0, 0, 0, 0, 0, 0.026148, 0.949314, 0, 0, 0, 0, 0, 0, 0, 0.041553, 0.911738, 0, 0, 0, 0, 0, 0, 0.07913, 0.819965, 0, 0, 0, 0, 0, 0, 0.170903, 0.814529, 0, 0, 0, 0, 0, 0, 0.176338, 0.804234, 0, 0, 0, 0, 0, 0, 0, 0.186634, 0.990868), 7, byrow=TRUE)</pre>

For a uniform size structure, the backward growth transition obtained from  $\mathbf{x}$  is:

backtrans(x, rep(1, 7))

[,1][,2][,3][,4][,5][,6][,7][1,]1.0017850.026805760.0000000.00000000.00000000.00000000.0000000[2,]0.0000000.982556980.04358900.00000000.00000000.00000000.0000000[3,]0.0000000.00000000.96599040.088010720.00000000.00000000.0000000[4,]0.0000000.00000000.922146160.17342950.00000000.0000000[5,]0.0000000.00000000.00000000.83583850.17983180.0000000[6,]0.0000000.00000000.00000000.00000000.82948120.1584999[7,]0.0000000.00000000.00000000.00000000.00000000.8492555

For a geometrically decreasing size structure, it becomes:

backtrans(x, 0.5<sup>(0:6)</sup>)

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]
[1,]	1.001785	0.05221194	0.0000000	0.000000	0.000000	0.000000	0.000000
[2,]	0.00000	0.95690638	0.08353672	0.000000	0.000000	0.000000	0.000000
[3,]	0.00000	0.0000000	0.92564261	0.1617828	0.000000	0.000000	0.000000
[4,]	0.00000	0.0000000	0.0000000	0.8475525	0.2955943	0.000000	0.000000
[5,]	0.00000	0.0000000	0.0000000	0.000000	0.7123040	0.3048431	0.000000
[6,]	0.00000	0.0000000	0.0000000	0.000000	0.000000	0.7030504	0.2736296
[7,]	0.000000	0.0000000	0.0000000	0.000000	0.000000	0.000000	0.7330647

**lin96.** The lin96 function performs forward projection and backward retrodiction of Lin et al.'s density-dependent matrix model. Retrodiction can be achieved with four inversion methods. By default, the initial state is an empty plot. To predict the state of the forest at year 300, the command is: x <- lin96(tmax=300, direction="forward"). The final size structure is:

c(x\$N)

```
[1] 440.5134424 145.6140188
                               71.2230832
                                                         29.3079380
                                            42.8568124
                                                                      21.7987933
 [7]
      17.1198835
                   13.8533728
                               11.2661934
                                             8.9359088
                                                          6.6365081
                                                                       8.0816202
[13]
      36.1248587
                    9.3961047
                                 3.7339482
                                              1.9484527
                                                          1.2746403
                                                                       1.0236757
[19]
                                              1.0909483
       0.9682634
                    1.0066783
                                 1.0691262
                                                          1.0172557
                                                                       2.1957118
[25]
      47.0053753
                   10.3408346
                                 3.1369528
                                              1.0878761
                                                          0.4369080
                                                                       0.2840420
[31]
       0.3477773
                    0.5133974
                                 0.7107542
                                              0.8750458
                                                          0.9398427
                                                                       2.1998746
```

Now let's predict the state of the forest at year 350 and retrodict it by 50 years, so as to go back to the state of the forest at year 300. The state of the forest at year 350 is given as before: y <- lin96(tmax=350, direction="forward"). When using the zero-order linear approximation as the inversion method, retrodiction over 50 years is obtained as: z <- lin96(y\$N, tmax=50, direction="backward", method="lin0"). The resulting size structure is:

c(z\$N)

[1]	440.9099975	145.7083961	71.4056396	42.9234726	29.3581548	21.8220650
[7]	17.1207081	13.8431935	11.2468742	8.9160017	6.6210297	8.0784638
[13]	36.2086037	9.4415105	3.7866975	1.9487815	1.2860456	1.0186158
[19]	0.9628681	0.9962843	1.0577456	1.0797272	1.0083329	2.1893031
[25]	47.0566714	10.5618776	3.3150239	1.0490887	0.5099222	0.2549266
[31]	0.3547649	0.4916455	0.6942365	0.8547708	0.9231371	2.1799661

To use the first-order linear approximation as the inversion method, the command becomes: z <- lin96(y\$N, tmax=50, direction="backward", method="lin1"), and the resulting size structure is:

c(z\$N)

```
[1] 440.8711115 145.8431489
                                71.3623019
                                                         29.3567663
                                                                      21.8213016
                                            42.9418675
 [7]
      17.1227573
                   13.8424333
                                11.2475331
                                             8.9160101
                                                          6.6212376
                                                                       8.0789789
[13]
      36.3006995
                                 3.7649258
                                              1.9626233
                                                          1.2792008
                                                                       1.0217545
                    9.4638429
[19]
       0.9616082
                    0.9967436
                                 1.0575836
                                              1.0797666
                                                          1.0083257
                                                                       2.1894161
[25]
      47.6033913
                   10.5477067
                                 3.2288341
                                              1.1312102
                                                          0.4546689
                                                                       0.2861415
[31]
       0.3394541
                    0.4981946
                                 0.6917965
                                             0.8555298
                                                          0.9229339
                                                                       2.1800416
```

To use the numerical inversion as the inversion method, the command becomes: z <- lin96(y\$N, tmax=50, direction="backward", method="num"), and the resulting size structure is:

c(z\$N)

[1]	440.8711115	145.8431489	71.3623019	42.9418675	29.3567663	21.8213016
[7]	17.1227573	13.8424333	11.2475331	8.9160101	6.6212376	8.0789789
[13]	36.3006995	9.4638429	3.7649258	1.9626233	1.2792008	1.0217545
[19]	0.9616082	0.9967436	1.0575836	1.0797666	1.0083257	2.1894161
[25]	47.6033913	10.5477067	3.2288341	1.1312102	0.4546689	0.2861415
[31]	0.3394541	0.4981946	0.6917965	0.8555298	0.9229339	2.1800416

Finally, to use RMM as the inversion method, the command becomes: z <- lin96(y\$N, tmax=50, direction="backward", method="trans"), and the resulting size structure is:

c(z\$N)

	00010 11.2101	43.0195	29.351099	/ 21.8309586
60414 13.830	04012 11.2309	9566 8.8988	737 6.588355	4 8.1148316
41202 9.644	40783 3.7241	1.9891	733 1.276687	1 1.0217430
<b>59211</b> 0.99	13534 1.0539	9252 1.07674	442 0.988259	0 2.2168478
19329 11.28	91339 3.0125	5472 1.26196	642 0.454678	2 0.3100076
48011 0.464	42630 0.6935	5197 0.86262	0.893427	8 2.2180646
	60414       13.83         41202       9.64         59211       0.99         19329       11.28         48011       0.46	60414       13.8304012       11.2309         41202       9.6440783       3.7241         59211       0.9913534       1.0539         19329       11.2891339       3.0125         48011       0.4642630       0.6935	60414       13.8304012       11.2309566       8.8988'         41202       9.6440783       3.7241585       1.9891'         59211       0.9913534       1.0539252       1.07674'         19329       11.2891339       3.0125472       1.26196'         48011       0.4642630       0.6935197       0.86262'	6041413.830401211.23095668.89887376.588355412029.64407833.72415851.98917331.276687592110.99135341.05392521.07674420.9882591932911.28913393.01254721.26196420.454678480110.46426300.69351970.86262540.893427

**RMM.** The RMM function fits a reverse matrix model to observations. Let's first generate a virtual dataset. We generate 5,000 observations that correspond to trees with a uniform diameter distribution in the range 10 cm–80 m. The initial size of individuals is given a column named initial:

set.seed(0)
data <- data.frame(initial=runif(5000, 10, 80))</pre>

We assume that all trees grow by 1 cm per year. The final size of individuals is given in a column named final:

#### data <- cbind(data, final=data\$initial+1)</pre>

We finally assume that 2% of individuals die and that 2% of individuals are newly recruited individuals. Dead individuals are indicated by a NA value in the final column while newly recruited individuals are indicated by a NA value in the initial column:

```
data[sample(5000, 5000*0.02), 1] <- NA
data[sample(5000, 5000*0.02), 2] <- NA
data$final[is.na(data$initial)] <- 10</pre>
```

We use diameter classes 10 cm wide. Therefore, on average, 10% of trees will move up from one class to the next one. To fit RMM, the command is:

```
r <- RMM(data, breaks=seq(10, 70, by=10))[[1]]
```

The output is a list with the three elements: r gives the initial size structure, r gives the final size structure, and the backward transition matrix is given by

r\$C

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]
[1,]	0.9986031	0.07859769	-0.0207999	-0.02079990	-0.0207999	-0.02079990	-0.02079990
[2,]	0.000000	0.92168675	0.1042524	0.0000000	0.000000	0.0000000	0.0000000
[3,]	0.000000	0.0000000	0.9135802	0.09745127	0.000000	0.0000000	0.0000000
[4,]	0.000000	0.0000000	0.000000	0.92053973	0.1062092	0.0000000	0.0000000
[5,]	0.000000	0.0000000	0.000000	0.0000000	0.9150327	0.08545727	0.0000000
[6,]	0.000000	0.0000000	0.000000	0.0000000	0.000000	0.94002999	0.07449495
[7,]	0.000000	0.0000000	0.000000	0.0000000	0.000000	0.0000000	0.94696970

We can check that the initial size structure is exactly given by the final size structure and the backward transition matrix:

c(r\$C %\*% r\$O) [1] 749 688 731 679 617 686 750 c(r\$N) 1 2 3 4 5 6 7 749 688 731 679 617 686 750

In this default setting, recruitment appears a fecundity rate on the first row of the transition matrix (with a negative sign because recruitment is removed from the first size class when moving backward). An alternative is to consider recruitment as an additional term. The fitting of RMM is then performed by the following command:

```
r <- RMM(data, breaks=seq(10, 70, by=10), recruitment="additional")[[1]]
```

The output now has an additional term r\$R that gives the number of recruited trees:

r\$C

### [1] 100

The backward transition matrix now is:

r\$C

[,1][,2][,3][,4][,5][,6][,7][1,]1.0194030.099397590.0000000.00000000.00000000.00000000.0000000[2,]0.0000000.921686750.10425240.00000000.00000000.00000000.0000000[3,]0.0000000.00000000.91358020.097451270.00000000.00000000.0000000[4,]0.0000000.00000000.920539730.10620920.00000000.0000000[5,]0.0000000.00000000.00000000.91503270.085457270.0000000[6,]0.0000000.00000000.00000000.00000000.940029990.07449495[7,]0.0000000.00000000.00000000.00000000.00000000.94696970

The initial size structure is again exactly given by the backward transition matrix and the final size structure from which the recruitment is now removed:

c(r\$C %\*% c(r\$O-c(r\$R,rep(0,6))))

[1] 749 688 731 679 617 686 750