

ECOGRAPHY

Research

Ecological memory at millennial time-scales: the importance of data constraints, species longevity and niche features

EDITOR'S
CHOICE

Blas M. Benito, Graciela Gil-Romera and H. John B. Birks

B. M. Benito (<https://orcid.org/0000-0001-5105-7232>) ✉ (blasbenito@gmail.com), Dept of Biological Sciences, Univ. of Bergen, Bergen, Norway. – G. Gil-Romera (<https://orcid.org/0000-0001-5726-2536>), Dept of Geography and Earth Sciences, Aberystwyth Univ., Aberystwyth, UK, and Inst. Pirenaico de Ecología-CSIC, Zaragoza, Spain. – H. J. B. Birks (<https://orcid.org/0000-0002-5891-9859>), Dept of Biological Sciences and Bjerkes Centre for Climate Research, Univ. of Bergen, Bergen, Norway, and Environmental Change Research Centre, Univ. College London, London, UK.

Ecography

43: 1–10, 2020

doi: 10.1111/ecog.04772

Subject Editor: Jessica Blois
Editor-in-Chief: Miguel Araújo
Accepted 24 September 2019



Ecological memory describes how antecedent conditions drive the dynamics of an ecological system. Palaeoecological records are paramount to understand ecological memory at millennial time-scales, but the concept is widely neglected in the literature, and a formal approach is lacking. Here, we fill such a gap by introducing a quantitative framework for ecological memory in palaeoecology, and assessing how data constraints and taxa traits shape ecological memory patterns.

We simulate the population dynamics and pollen abundance of 16 virtual taxa with different life and niche traits as a response to an environmental driver. The data is processed to mimic a realistic sediment deposition and sampled at increasing depth intervals. We quantify ecological memory with Random Forests, and assess how data properties and taxa traits shape ecological memory.

We find that life-span and niche features modulate the relative importance of the antecedent values of the driver and the pollen abundance over periods of 240 yr and longer. Additionally, we find that accumulation rate and decreasing pollen-sampling resolution inflate the importance of antecedent pollen abundance.

Our results suggest that: 1) ecological memory patterns are sensitive to varying accumulation rates. A better understanding on the numerical basis of this effect may enable the assimilation of ecological memory concepts and methods in palaeoecology; 2) incorporating niche theory and models is essential to better understand the nature of ecological memory patterns at millennial time-scales. 3) Long-lived generalist taxa are highly decoupled from the environmental signal. This finding has implications on how we interpret the abundance-environment relationship of real taxa with similar traits, and how we use such knowledge to forecast their distribution or reconstruct past climate.

Keywords: ecological niche, mechanistic models, palaeopalynology, population dynamics, Random Forest, virtual species



www.ecography.org

© 2019 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Introduction

Padisak (1992) defined ecological memory as ‘the capacity of past states or experiences to influence present or future responses of the community’. The ambiguity of this definition has led to different interpretations of what ecological memory actually is and how it can be measured. Schweiger et al. (2019), following the biological legacy concept proposed by Nyström and Folke (2001), define ecological memory at the ecosystem level as the ‘accumulated abiotic and biotic material and information legacies from past dynamics’. Ogle et al. (2015) adopt a quantitative perspective instead, defining ecological memory as ‘the effect of antecedent conditions on current processes’, and propose a quantitative framework to detect the patterns resulting from ecological memory processes.

The framework presented by Ogle et al. (2015) identifies two main components of ecological memory: endogenous memory, which is the influence of antecedent values of a biotic variable of interest on itself, and exogenous memory, which represents the effect of antecedent values of an external driver or drivers on the biotic response. Additionally, the synchronous effect of the driver or drivers on the response (the concurrent effect, Fig. 1) can also be quantified, in order to compare the relative strength of antecedent versus concurrent conditions in driving ecological responses. Two key features of ecological memory components are further defined by Ogle et al. (2015), namely memory length and memory strength, respectively representing the time period over which antecedent conditions have a significant influence on the response and the amount of influence they exert.

Ecological memory analyses reveal that different processes generate memory-driven responses at different biological,

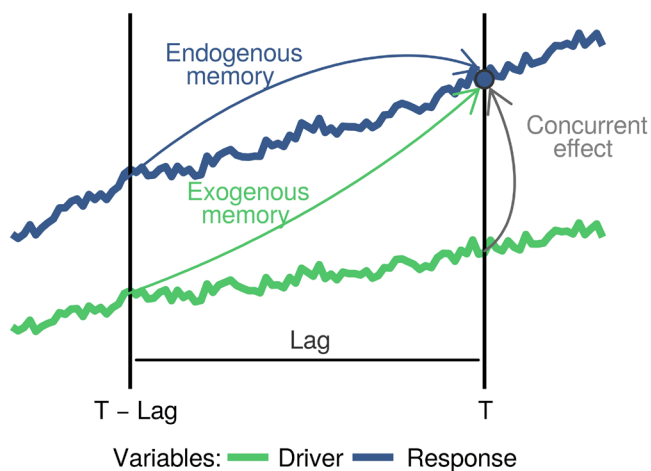


Figure 1. Components of ecological memory. The biotic variable of interest (response) is represented in blue, and the environmental driver in green. Arrows represent the influence of antecedent values of the driver and the response (for a time interval defined by Lag), and concurrent values of the driver over the response at a time of interest T.

spatial and temporal scales. For example, Ogle et al. (2015) found that stomatal conductance in leaves of a desert shrub respond to water content at different depths (30 and 60 cm) with different memory lengths (1 and 7 d). Anderegg et al. (2015) and Schwalm et al. (2017) analysed post-drought responses of tree-growth to find persistent effects after four years. Perring et al. (2018) show how forest management in 1800 still influences plant community composition and functioning today. These studies exemplify very well how analysing time-delayed effects of endogenous and exogenous components of an ecological system provides a way of deepening our understanding of its long-term dynamics.

Ecological memory in a palaeoecological context

Palaeoecological records capture the patterns of continuous ecological processes happening over hundreds to millions of years as discrete samples that are linked together in the time continuum (Rull 2014). If the relativistic principle of causality ‘cause must precede its effect’ is applied, time emerges as the primary axis over which cause-and-effect loops develop. Therefore, palaeoecological time series should be able to provide the means to assess the importance of antecedent states in explaining the system’s dynamics at centennial to millennial time-scales. This idea has long been identified by palaeoecologists, leading to the concept of ‘legacy’ (Herzschuh et al. 2016). However, due to data limitations, this idea has been largely used in a narrative way rather than in a quantitative one, where most efforts rely on the application of cross-correlation analysis and the Granger test.

Cross-correlation analysis relies on computing the mutual correlation between two regular time-series across increasing time lags, and has been used to identify causal links between biomass and fire frequency (Tinner et al. 1999, Gil-Romera et al. 2014), and delayed effects of environmental and anthropogenic drivers on pollen abundance (Kaniewski et al. 2018). Cross-correlation measures the exogenous and concurrent memory components, but ignores the endogenous memory and lacks the means to analyse more than one driver at a time.

The Granger test assumes that a cause (driver) happens prior to its effect (response) and that a cause should have relevant information on the future values of the effect (Granger 1969). Shuman and Marsicek (2016) apply it to analyse causal links between temperature and moisture in the mid-latitudes of North America during the Holocene. Although the authors do not mention ecological memory explicitly, they provide the closest to a quantitative concept of ecological memory that can be found to date in the palaeoecological literature.

In our view, there are three main reasons that prevent the adoption of ecological memory concepts and methods in palaeoecology: 1) the sparse availability of independent proxies of environmental drivers with a direct effect on the presence and abundance of plant taxa, 2) an uneven understanding of the ecology of the taxa represented in the fossil record and

3) the various errors deriving from taphonomic processes and the uncertainty associated with the different quality of chronological models.

Reliable independent palaeoclimate proxies such as oxygen, hydrogen and carbon isotopes exist in different palaeoarchives, although these are often difficult to align with pollen sequences due to chronological uncertainties, even within close geographical locations (Blaauw and Mauquoy 2012). Transient palaeoclimate simulations may help fill this knowledge gap, but still cannot represent high-resolution climate variability (Harrison et al. 2016). However, the increasing number of multi-proxy records (Birks and Birks 2006) offers promising opportunities for the application of novel methods to understand climate-vegetation dynamics.

Simultaneously, the proliferation of public databases on plant traits, taxonomy, species presence, climate and palaeoclimate, remote sensing and repositories of fossil proxies is opening a window of opportunity to improve our empirical knowledge on vegetation responses to environmental change and disturbance regimes, and to bridge the gap between neoecology and palaeoecology through the adoption of state-of-the-art quantitative data, methods and concepts (Svenning et al. 2011).

Sediments deposit at variable accumulation rates, while pollen samples are often taken at fixed depth intervals, resulting in an uneven temporal aggregation that may hide the footprint of ecological processes driving vegetation dynamics (Liu et al. 2012). Conceptual and technical breakthroughs such as pollen accumulation rates (Davis and Deevey 1964) based on high-quality age-depth models are able to standardise pollen counts by units of time, offering a more precise representation of palynological change and, by interpretation, biomass for a given plant taxon.

Improvements in proxy availability, pollen data quality and resolution, and age-depth modelling coupled with a stronger bidirectional flow of knowledge with neo-ecology (especially ecological niche theory and functional ecology) will soon enable higher resolution analyses with a signal-to-noise ratio that allows patterns of ecological memory in palaeoecological datasets to be detected.

Aims

This paper builds on the quantitative momentum that palaeoecology is currently acquiring by introducing the concepts and methods of ecological memory into the discipline. Particularly we aim to assess the suitability of ecological memory as a tool to better understand vegetation dynamics at millennial time-scales by analysing the results of individual-based simulations with a robust machine-learning algorithm to answer two questions. 1) To what extent do sediment accumulation and pollen sampling intervals affect ecological memory patterns? 2) How do life traits and niche features (life span, fecundity, niche breadth, niche optimum) shape ecological memory patterns?

Material and methods

A comprehensive exploration of these questions requires pollen data generated by taxa with known traits as a response to known proxies, and disentangled from the effect of competition with other taxa. Since these properties are absent from most palaeoecological datasets, we have adopted an *in silico* approach relying on a model simulating the processes behind the generation of pollen data that is grounded in ecological theory (Perry et al. 2016).

Simulated taxa

Using the custom 'virtualPollen' package (available in R, R Core Team) we simulate a virtual driver in the range [0, 100] based on a random walk with temporal autocorrelation at annual time-steps over a period of 10 000 years (Supplementary material Appendix 1 section 1). To generate virtual pollen curves with annual resolution based on this driver we used the function 'simulatePopulation' of the same package, which implements an individual-based, mono-specific and non-spatial mechanistic population model, which relies on five assumptions.

1) Environmental niches follow a Gaussian function representing the regeneration niche of the taxa (Grubb 1977). Niche optimum and breadth (tolerance) are represented by the mean and standard deviation of the normal function. This function transforms the driver values into 'environmental suitability' values.

2) Individual biomass growth follows the logistic equation (Birch 1999): $b = (B/1 + B) + e^{(-at)}$, where b is the individual's biomass, B is the maximum biomass, a is the growth rate and t is the age of the individual at the given time.

3) Yearly recruitment is a function of the individual's biomass (Thomas 2011) and the environmental suitability, and limited by maximum fecundity.

4) Mortality sources are senescence and self-thinning (Berger et al. 2004). When carrying capacity is reached, individuals are removed according to a risk curve (Supplementary material Appendix 1 Eq. 3) until the population biomass is below the carrying capacity.

5) Yearly pollen production is the biomass sum of all adult individuals multiplied by the environmental suitability of the given year, matching observed pollen-environment relationships (Barnekow et al. 2007).

We define 16 virtual taxa with combinations of life traits and niche features parameters (Supplementary material Appendix 1 Table A7) intended to simulate short and long-lived plant taxa with different fecundities, and having either central or marginal environmental niches with different tolerances (generalist versus specialist taxa). Each simulation output contains one record per year of driver values, suitability computed by the niche function of the virtual taxa, population size, biomass and pollen concentration. Driver values and the pollen output for each simulation is compared through Pearson correlation.

To account for potential distortions produced by uneven sediment accumulation rates on pollen-environment relationships, simulation outputs are aggregated into years per centimetre by following a virtual accumulation rate (range 1–50 yr cm⁻¹; Supplementary material Appendix 1 section 3), and sampled at intervals of 1, 2, 6 and 10 cm (Liu et al. 2012). The data are interpolated at a 20-yr resolution using the R function ‘loess’ (Supplementary material Appendix 1 section 4). The names used hereafter for the aggregation levels are: 1, 2, 6 and 10 cm. The direct output of the simulation interpolated at 20-yr intervals (constant accumulation rate) is named Annual, and represents the true data to be used as a control in the analyses. Each dataset is time-lagged by aligning each pollen sample with its antecedent driver and suitability values at 20-yr intervals, to a maximum of 240 yr.

Ecological memory modelling

Random Forest is a machine-learning algorithm based on an ensemble of regression trees suitable to work with highly collinear data and insensitive to temporal autocorrelation (Breiman 2001). Its variable importance assessment (Supplementary material Appendix 2 section 2 for further details) works under the assumption that if a given variable is not important, then permuting its values does not degrade predictive accuracy. A large number of trees and a small number of variables to find each split has been found to increase the robustness of the variable importance analysis (Grömping 2009). In line with these findings, we configured the R package ‘ranger’ (Wright and Ziegler 2017) to fit forests of 2000 trees, with only two predictors considered at each split. Further details on this method are available in Supplementary material Appendix 2 section 2.

Fitted models were based on the expression:

$$p_t = p_{t-20} + p_{t-40} + \dots + p_{t-240} + d_t + d_{t-20} + d_{t-40} + \dots + d_{t-240} + r \quad (1)$$

where: p_t is the pollen output of the population at time t . p_{t-20} to p_{t-240} are the antecedent pollen values for the lags 20 to 240 (yr), representing the endogenous memory. Lags of 20 to 240 yr taken at 20-yr intervals were selected after preliminary trials showed that memory components of the virtual taxa became insignificant after a lag of 240 yr. d_{t-20} to d_{t-240} are the antecedent values of a driver, or the suitability values derived from the driver by the niche function of the virtual taxa, representing the exogenous memory. d_t is the concurrent effect of the driver or the suitability over the response (at lag 0). r is a variable with a random autocorrelation length (between 1 and 1/4 of the total length of p_t), with no causal link with p_t , used as a null model to assess the statistical significance of the variable importance scores returned by Random Forest.

Considering that temporally autocorrelated variables with no causal link with each other can show high correlation by chance (Dean and Dunsmuir 2016), and that Random Forest does not measure statistical significance for variable

importance scores, we introduce the term r in Eq. 1 to facilitate the identification of spurious links between lagged predictors and the response. Each model is repeated 300 times (this number of repetitions reaches a compromise between accuracy and computation time), each one with an r term with a different autocorrelation length (Supplementary material Appendix 2 section 2.3.1). Percentiles 0.95, 0.5 and 0.05 of the importance scores obtained by each model term across repetitions are computed as descriptors of the ecological memory pattern, and stored for visualisation and further analyses. Values of variable importance below the median of the random component are considered spurious. Pearson correlation between observed and predicted values (pseudo R-squared) is computed to assess model fit. The functions required to perform this analysis are available in the R package ‘memoria’.

We compute three features for each ecological memory component (Fig. 2): 1) strength: defined as the maximum difference in relative importance, scaled to the range [0, 1], between each component (endogenous, exogenous, concurrent) and the median of the random component; 2) length: proportion of lags over which the importance of endogenous and exogenous memory is above the median of the random component; 3) dominance: proportion of the lags above the median of the random term over which endogenous and exogenous memory have a higher importance than the other component.

Ecological memory features of the 16 virtual taxa across datasets (Annual, 1, 2, 6, 10 cm) when using either driver or suitability as the exogenous memory component are computed and used to perform four different analyses.

We compute the differences in ecological memory features between the Annual and the 1 cm dataset (constant versus varying accumulation rates), using driver as the exogenous memory component, to test how sediment accumulation rates can affect ecological memory patterns. To assess how depth intervals between consecutive pollen samples modify

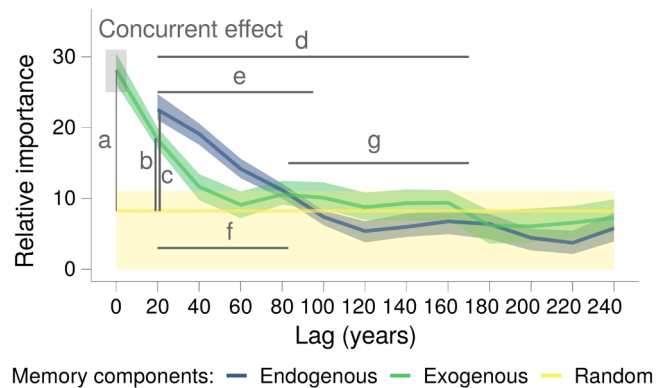


Figure 2. Features of ecological memory components: a: strength of the concurrent effect (highlighted by a grey box); b and c: strength of the exogenous and endogenous memory; d and e: length of the exogenous and endogenous memory; f and g: dominance of the endogenous and exogenous memory.

ecological memory, we compare ecological memory features across the datasets 1, 2, 6 and 10 cm, using driver as the exogenous memory component.

Our model uses a Gaussian function to transform the driver values into environmental suitability. The correlation between driver and suitability decreases when the mean of the niche function is close to the mean of the driver (central niche), and the standard deviation is small (narrow niche). In such a case, ecological memory patterns computed from the driver alone may generate misleading interpretations on the importance of the exogenous memory. We address this potential issue by assessing differences in ecological memory features produced by the Annual dataset when using alternatively driver and suitability as exogenous memory components in Eq. 1.

To understand how different life and niche traits may lead to distinct ecological memory patterns, we averaged the strength, length, and dominance of the endogenous and the exogenous (using suitability as predictor) memory components of the Annual dataset. Life and niche traits of the virtual taxa are plotted on this space to facilitate the interpretation of the links between traits and ecological memory components.

Results

Average Pearson correlation between the driver and the pollen output was 0.384 (standard deviation 0.374) across taxa, 0.031 (standard deviation 0.018) for taxa with central niches and 0.737 (standard deviation 0.121) for taxa with narrow niches.

The concurrent component (effect of the exogenous variable at lag 0 on the response) generally has the highest importance across taxa and datasets. The exogenous and endogenous memory components are generally higher than the median of the random component for lags over 100 yr for most taxa, although taxa with a life-span of 1000 yr yield ecological memory components that remain important beyond the random expectation for most of the complete set of lags (Supplementary material Appendix 3 section 1.2, Fig. 2, 3).

When assessing the differences in ecological memory features between the Annual (constant accumulation rate) and the 1 cm (variable accumulation rate) datasets, we find that sediment accumulation rate generally induces a reduction in the importance of the concurrent and exogenous memory features of most virtual taxa, while it inflates the dominance of the endogenous memory component of half of the virtual taxa (Fig. 3). This effect leads to a switch in the relative importance of the endogenous and exogenous components in the first lags considered for two of the virtual taxa (Supplementary material Appendix 3 Fig. 1).

The analysis of the effect of pollen-sampling intervals on ecological memory patterns shows a general decrease in memory features with increasing sampling intervals, which is especially acute in the exogenous memory features and the strength of the concurrent effect. Endogenous memory

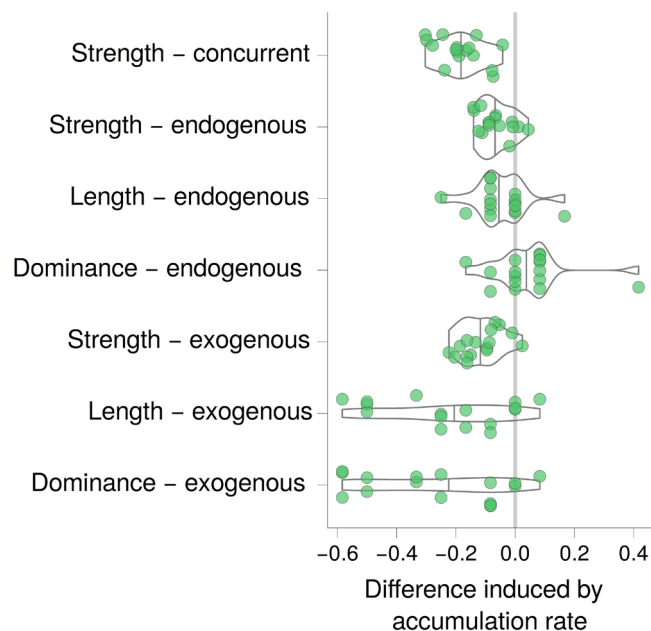


Figure 3. Differences in memory features between the datasets Annual (constant accumulation rate) and 1 cm (variable accumulation rate). Each dot represents a virtual taxon. Central lines in the violin plots represent the median. The x-axis is computed by subtracting the memory features of the 1 cm dataset from the ones of the Annual dataset, and represents the effect of a variable sediment accumulation rate on ecological memory features.

features show gentler but discordant responses, with a moderate increase in the dominance, accompanied by a decrease in the strength and a neutral response in the length (Fig. 4).

We find that using driver instead of suitability as the exogenous memory term in Eq. 1 fitted on the Annual dataset decreases the values of most memory features, although taxa with wider niches are less affected (Fig. 5). This shift is more acute in the strength of the concurrent component, especially for taxa with narrow niches, and in the length of the exogenous memory, irrespective of the niche features of the taxa. Taxa with central niches show a positive shift in the strength of the endogenous memory, while taxa with marginal and narrow niches show neutral or opposite responses. The length of the endogenous component is lower across most taxa.

The analysis of the relationship between species traits and ecological memory components shows that the exogenous memory and the concurrent effect (representing the effect of past and concurrent environments) are stronger than the endogenous memory for all taxa. It also reveals how life-span interacts heavily with niche position to generate memory patterns with different relative contributions of each memory component. In particular, taxa with longer life-spans show a stronger endogenous memory and a weaker exogenous memory when their niche position is not central. Short-lived taxa yield the weakest endogenous memory, except for those with central and narrow niches that simultaneously have the highest values of ecological memory and concurrent effect

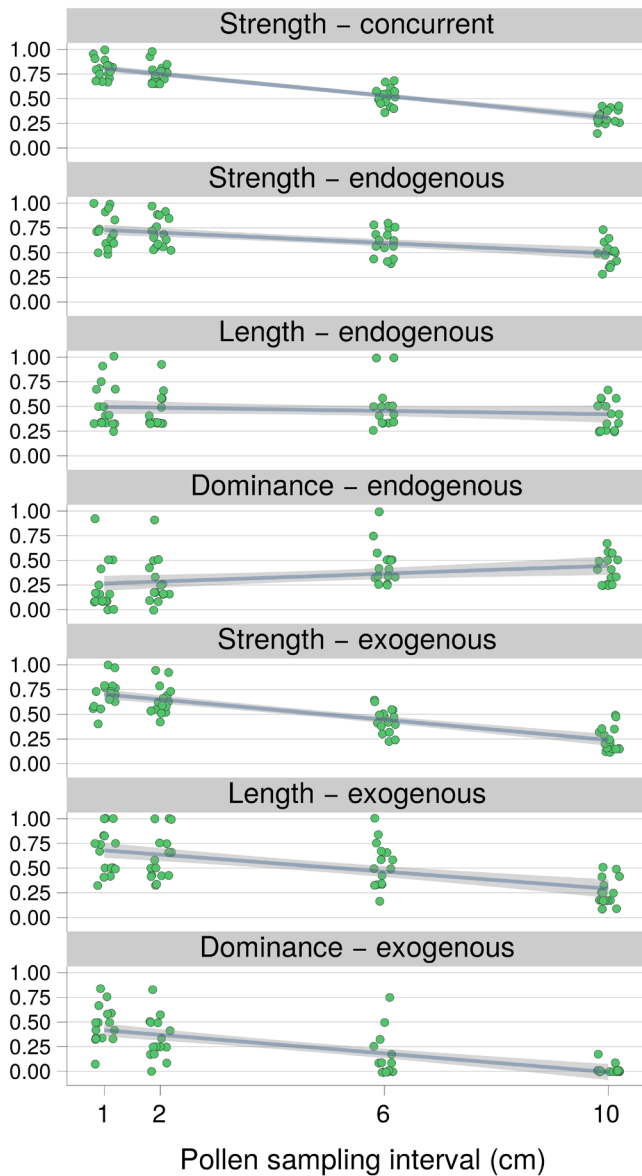


Figure 4. Ecological memory features across increasing sampling intervals. Dots represent virtual taxa. Horizontal jitter was applied to the points for ease of examination. Linear models across groups are fitted with the ‘lm’ function in the R software.

(Fig. 6). The fecundity trait does not seem to play an important role in defining the memory patterns of the virtual taxa.

Discussion

In this paper we assess how data-generation processes, such as sediment accumulation rate and pollen-sampling intervals, and taxa features such as the properties of their environmental niche and life traits, may shape the expression of ecological memory patterns *sensu* Ogle et al. (2015) in palaeoecological time-series. We find that: 1) sediment accumulation rate and long depth intervals between consecutive

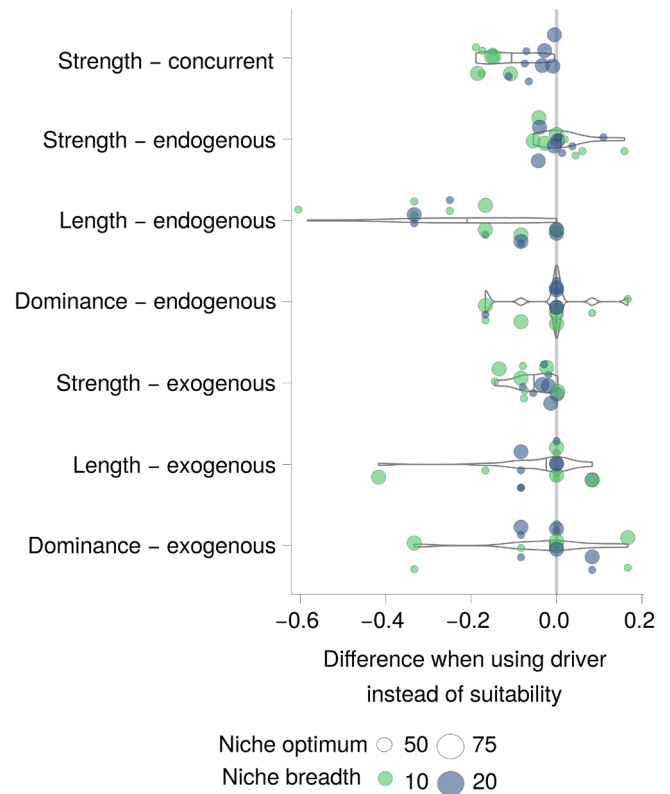


Figure 5. Change in ecological memory features when using driver instead of suitability as exogenous component in Eq. 1 fitted on the Annual dataset. Each dot represents a virtual taxon, size represents niche optimum (position), and colours represent niche breadth.

pollen samples act in synergy to distort ecological memory patterns by inflating the importance of endogenous memory and downplaying the importance of concurrent and antecedent climate; 2) ignoring the niche functions that mediate pollen–climate relationships can lead to an under-estimation of the influence of antecedent endogenous and exogenous conditions, especially when the environmental niche of the taxon is narrow; 3) life-span interacts strongly with niche position to generate distinct ecological memory patterns. Long life-span leads to an increased strength of the endogenous memory and a decreased influence of the concurrent and antecedent climate when the niche position is marginal.

With this study we formally introduce the concept of ecological memory into palaeoecology, evaluate its usefulness as a quantitative tool in the discipline, assess how data and taxa properties influence the interpretation of ecological memory patterns, and discuss its potential to improve our understanding of time-delayed responses of biotic variables to environmental change at millennial time-scales.

Data generation processes and ecological memory

Our results show that a variable sediment accumulation rate leads to consistent changes in the importance of the different

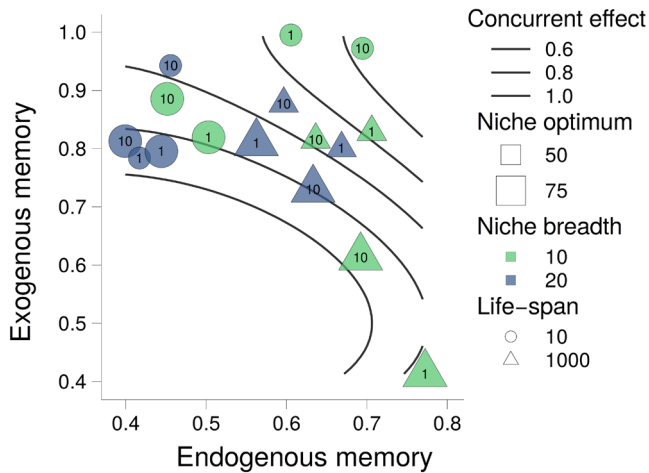


Figure 6. Relationship between species traits and ecological memory components as found in the Annual dataset when using suitability as the exogenous memory term. Axes x and y are the average of all memory features for endogenous and exogenous memory, respectively. Contour lines (concurrent effect) are the result of a polynomial generalised linear model (GLM) using endogenous and exogenous memory as predictors (pseudo R-squared = 0.95).

ecological memory components, and to an underestimation of the influence of the exogenous memory and the concurrent effect. Random differences in memory patterns from the same data under constant and variable accumulation rates are expected, because the average pollen abundance of a section with high inter-annual pollen variability and a steep sediment accumulation rate might not be very representative of the actual pollen abundance produced by the simulation at a given time. However, and according to our results, there is directionality in the observed changes in ecological memory features. We hypothesise that the uneven smoothing produced by the accumulation rate is increasing the temporal autocorrelation of the pollen abundance, and cancelling the fine-scale correlation between the driver and the pollen curve, ultimately raising endogenous memory features and reducing the exogenous ones.

The effect of increasing depth intervals between consecutive pollen samples perfectly mirrors the pattern found when analysing the effect of a varying accumulation rate on ecological memory features. Even though all datasets are interpolated to the same regular time grid, decreasing sampling resolution only increases the smoothing of the data and enhances the decoupling of fine-scale details of the driver and pollen abundance. This effect is largely linear across sampling intervals, and clearly indicates that short intervals between consecutive samples, or no intervals at all, may provide the data conditions required to assess time-delayed climate–vegetation relationships.

However, it is worth noting that we test a single accumulation rate curve (Supplementary material Appendix 1 Fig. 14) and that the number of virtual taxa and traits used is limited. Nevertheless, our results suggest that further research is needed on how sediment accumulation rates decouple

fine-scale relationships between climate proxies and biotic responses. Such analyses can be either carried out in a simulation like the one we designed and used in this study, or by applying virtual accumulation rates to high-resolution pollen datasets, such as from annually resolved varved sediments (Vandergoes et al. 2018).

Ecological memory, traits and niches

Interesting properties of ecological memory patterns emerge when considering ecological niche theory within a palaeoecological context and assuming that the link between a driver and a biological response is mediated by a niche function (i.e. Gaussian). Strong non-linearities between driver and response may arise, more so when the niche optimum of the given taxon is close to the average of the driver (central niche), and niche breadth is wide, as happens with generalist taxa (Sexton et al. 2017). This signal is consistent with other effects found in species with wide niche breadths, such as a poor predictability of their distributions in relation to climate (Connor et al. 2018), or lower potential impact of ongoing climate warming on their present (Herrera et al. 2018) and future distributions (Thuiller et al. 2005). The practical consequence of this issue is not only that a low correlation between driver and response might hinder our ability to unveil ecological memory patterns with statistical methods, but that the driver itself might actually have a minor role in driving the population dynamics of a taxon when the niche function, and its result, namely suitability as perceived by the taxon, is ignored. Our results show that this omission might come with a price, since the relative contribution of different ecological memory features obtained when using a driver as the exogenous memory component might differ from the truth, and can lead to an underestimation of the actual importance of ecological memory processes.

The effects of niche properties on ecological memory are largely amplified by life-spans. In particular, long-lived taxa can further decouple their population dynamics from climate variability, because they require longer time to reach maturity (Petit and Hampe 2006), and are therefore unable to track fast climate amelioration. Furthermore, long life-span also promotes the persistence of biomass under low climate suitability, as observed in relict plant populations (Hampe and Arroyo 2002), facilitating a fast recovery of the reproductive function and maximum pollen production when conditions become suitable, without requiring a long process of population build-up.

Our results align with these ideas in suggesting that the population dynamics (and hence pollen productivity) of long-lived taxa is often less restricted by climatic drivers and more driven by their own life-span. In consequence, pollen concentrations yielded by these virtual taxa are mostly decoupled from antecedent driver values and, in consequence, their endogenous memory is stronger than in other taxa with shorter life-spans.

We acknowledge that considering niche features and life-traits might be seem impractical, as often palaeopalynological data may represent taxonomic entities as coarse as a plant family, with unknown ecological niches, and that available climate proxies represent a combination of climate factors rather than particular climatic variables. However, we provide evidence that it is critical to understand the shape of taxa responses to environmental drivers in order to interpret correctly the observed ecological memory patterns. An open question is to what extent this limitation applies to the interpretation of other patterns observed in palaeoecological time series, or how does it contribute to observed uncertainty in climate reconstructions based on modern analogues (Birks 2011).

Considerations about our theoretical approach

The simulation model we present allows us to design and perform *in silico* experiments without most of the uncertainties found in real palaeopalynological datasets (Paus 2013). However, these unknowns are replaced by the limitations introduced by the assumptions of the model and its numerical mechanics. In order to achieve maximum simplification by omitting interspecific interactions, our model assumes mono-specific communities, which are relatively rare in nature, albeit more frequent in northern latitudes. Furthermore, real pollen datasets are often expressed as pollen percentages or proportions (Prentice and Webb 1986), making it difficult to consider any given pollen type in isolation from the others. A possible work-around to this particular limitation is to estimate pollen accumulation rates (Davis and Deevey 1964), which are intended to represent independent pollen influx of a given taxon, and should be useful to compute ecological memory without the interference arising from the abundances of other taxa.

Additionally, we assume that taxa responses to the environment follow Gaussian functions, even though empirical data show that response curves can be asymmetric, particularly when close to the extremes of the environmental gradient (Austin and Gaywood 1994). Nevertheless, this limitation can be overcome by introducing asymmetric niche functions into the simulation, such as the Beta distribution, among others.

In our results, the concurrent component (synchronous effect of the driver over the concentration of virtual pollen) has the highest importance across virtual taxa and pollen-sampling intervals. Such an outcome is a direct result of the simulation design, which computes pollen productivity at any given time by multiplying biomass (range [0, 100]) and suitability (range [0, 1]), producing pollen maxima coupled with peaks of high suitability. This design matches observed patterns of climate–pollen relationships (Barnekow et al. 2007), but research on this topic is still scarce, and it is possible that different real pollen types might show different levels of dependence between concurrent climate and pollen productivity.

From what is outlined above it could be concluded that our simulation model is not a comprehensive theoretical representation of every process contributing to the generation of pollen curves, but a simplified one, useful to test ideas on how ecological memory patterns may emerge from the interaction between taxa with different features and a changing environment. Undoubtedly, proper assessments of time-delayed responses to environmental factors must be tested on real multi-proxy datasets and on taxa with known traits and niche features.

Moving forward

Our study shows that there is potential for the application of ecological memory concepts in palaeoecology. It also highlights how data quality and resolution, along with the availability of proxies representing meaningful drivers, and knowledge about the ecological niche of target taxa, are critical to advance from a theoretical stage into a more practical application of the concept. Under these ideal circumstances, we identify several exciting but challenging questions lying ahead. 1) Are ecological memory patterns conserved across time and sites? Finding whether ecological memory patterns for given taxa are constant or not under different climate conditions and accompanying communities opens the door to a better understanding of the mechanisms behind idiosyncratic responses to environmental change (Rumpf et al. 2014). 2) Do population growth and collapse show the same ecological memory footprint? Different ecological mechanisms are behind population growth and collapse under changing conditions, but these differences are rarely approached from a quantitative perspective in palaeoecology. Assessing ecological memory separately for population build-up and decline phases in pollen curves has the potential to detect links between species traits and climate trends not examined before. 3) Are ecological memory patterns found in palaeopalynological datasets useful to understand future responses of vegetation to the ongoing climate warming? If ecological memory patterns reveal time-delayed responses of biotic variables to antecedent climate, a comprehensive understanding of past ecological memory patterns might prove to be valuable in improving current ecological forecasting practices, moving these beyond correlations with concurrent climate, and bridging the gap existing between palaeoecology and neo-ecology in their efforts to address responses of biota to global warming.

Data availability statement

Supplementary materials with the code and simulated data used in this paper are available at <<https://github.com/BlasBenito/EcologicalMemory>>, doi: 10.5281/zenodo.3236128. R packages ‘virtualPollen’ and ‘memoria’, authored by Blas M. Benito, are available in CRAN and GitHub (<<https://github.com/BlasBenito>>).

Acknowledgements – This paper is a contribution to the IGNEX Project.

Funding – BMB and HJBB were supported by FRIMEDBIO (Research Council of Norway) through IGNEX (project 249894). GGM was supported by the DFG funded action FOR 2358.

Author contributions – BMB, GGR and HJBB conceived the ideas; BMB designed the methodology, and performed the analyses; BMB led the writing of the manuscript, with feedback from GGR and HJBB. All authors contributed critically to the drafts and gave final approval for publication.

References

- Anderegg, W. R. L. et al. 2015. Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. – *Science* 349: 528–532.
- Austin, M. P. and Gaywood, M. J. 1994. Current problems of environmental gradients and species response curves in relation to continuum theory. – *J. Veg. Sci.* 5: 473–482.
- Barnekow, L. et al. 2007. Strong correlation between summer temperature and pollen accumulation rates for *Pinus sylvestris*, *Picea abies* and *Betula* spp. in a high-resolution record from northern Sweden. – *J. Quat. Sci.* 22: 653–658.
- Berger, U. et al. 2004. Age-related decline in forest production: modelling the effects of growth limitation, neighbourhood competition and self-thinning. – *J. Ecol.* 92: 846–853.
- Birch, C. P. D. 1999. A new generalized logistic sigmoid growth equation compared with the Richards growth equation. – *Ann. Bot.* 83: 713–723.
- Birks, H. H. and Birks, H. J. B. 2006. Multi-proxy studies in palaeolimnology. – *Veg. Hist. Archaeobot.* 15: 235–251.
- Birks, H. J. B. et al. 2011. Strengths and weaknesses of quantitative climate reconstructions based on Late-Quaternary biological proxies. – *Open Ecol. J.* 3: 68–110.
- Blaauw, M. and Mauquoy, D. 2012. Signal and variability within a Holocene peat bog – chronological uncertainties of pollen, macrofossil and fungal proxies. – *Rev. Palaeobot. Palynol.* 186: 5–15.
- Breiman, L. 2001. Random forests. – *Mach. Learn.* 45: 5–32.
- Connor, T. et al. 2018. Effects of grain size and niche breadth on species distribution modeling. – *Ecography* 41: 1270–1282.
- Davis, M. B. and Deevey, E. S. 1964. Pollen accumulation rates: estimates from late-glacial sediment of Rogers Lake. – *Science* 145: 1293–1295.
- Dean, R. T. and Dunsmuir, W. T. M. 2016. Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: the importance of constructing transfer function autoregressive models. – *Behav. Res. Methods* 48: 783–802.
- Gil-Romera, G. et al. 2014. Biomass-modulated fire dynamics during the last glacial–interglacial transition at the Central Pyrenees (Spain). – *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 402: 113–124.
- Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. – *Econometrica* 37: 424–438.
- Grömping, U. 2009. Variable importance assessment in regression: linear regression versus random forest. – *Am. Stat.* 63: 308–319.
- Grubb, P. J. 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. – *Biol. Rev.* 52: 107–145.
- Hampe, A. and Arroyo, J. 2002. Recruitment and regeneration in populations of an endangered South Iberian Tertiary relict tree. – *Biol. Conserv.* 107: 263–271.
- Harrison, S. P. et al. 2016. What have we learnt from palaeoclimate simulations? – *J. Quat. Sci.* 31: 363–385.
- Herrera, J. M. et al. 2018. Climatic niche breadth determines the response of bumblebees (*Bombus* spp.) to climate warming in mountain areas of the northern Iberian Peninsula. – *J. Insect Conserv.* 22: 771–779.
- Herzschuh, U. et al. 2016. Glacial legacies on interglacial vegetation at the Pliocene–Pleistocene transition in NE Asia. – *Nat. Commun.* 7: 11967.
- Kaniewski, D. et al. 2018. Croatia's mid-Late Holocene (5200–3200 BP) coastal vegetation shaped by human societies. – *Quat. Sci. Rev.* 200: 334–350.
- Liu, Y. et al. 2012. Temporal density of pollen sampling affects age determination of the mid-Holocene hemlock (*Tsuga*) decline. – *Quat. Sci. Rev.* 45: 54–59.
- Nyström, M. and Folke, C. 2001. Spatial resilience of coral reefs. – *Ecosystems* 4: 406–417.
- Ogle, K. et al. 2015. Quantifying ecological memory in plant and ecosystem processes. – *Ecol. Lett.* 18: 221–235.
- Padisak, J. 1992. Seasonal succession of phytoplankton in a large shallow lake (Balaton, Hungary) – a dynamic approach to ecological memory, its possible role and mechanisms. – *J. Ecol.* 80: 217–230.
- Paus, A. 2013. Human impact, soil erosion, and vegetation response lags to climate change: challenges for the mid-Scandinavian pollen-based transfer-function temperature reconstructions. – *Veg. Hist. Archaeobot.* 22: 269–284.
- Perring, M. P. et al. 2018. Global environmental change effects on plant community composition trajectories depend upon management legacies. – *Global Change Biol.* 24: 1722–1740.
- Perry, G. L. W. et al. 2016. Experimental simulation: using generative modeling and palaeoecological data to understand human-environment interactions. – *Front. Ecol. Evol.* 4: 109.
- Petit, R. J. and Hampe, A. 2006. Some evolutionary consequences of being a tree. – *Annu. Rev. Ecol. Evol. Syst.* 37: 187–214.
- Prentice, I. C. and Webb, T. 1986. Pollen percentages, tree abundances and the Fagerlind effect. – *J. Quat. Sci.* 1: 35–43.
- Rull, V. 2014. Time continuum and true long-term ecology: from theory to practice. – *Front. Ecol. Evol.* 2: 75.
- Rumpf, S. B. et al. 2014. Idiosyncratic responses of high arctic plants to changing snow regimes. – *PLoS One* 9: e86281.
- Schwalm, C. R. et al. 2017. Global patterns of drought recovery. – *Nature* 548: 202–205.
- Schweiger, A. H. et al. 2019. The importance of ecological memory for trophic rewinding as an ecosystem restoration approach. – *Biol. Rev.* 94: 1–15.
- Sexton, J. P. et al. 2017. Evolution of ecological niche breadth. – *Annu. Rev. Ecol. Evol. Syst.* 48: 183–206.
- Shuman, B. N. and Marsicek, J. 2016. The structure of Holocene climate change in mid-latitude North America. – *Quat. Sci. Rev.* 141: 38–51.
- Svenning, J.-C. et al. 2011. Applications of species distribution modeling to paleobiology. – *Quat. Sci. Rev.* 30: 2930–2947.
- Thomas, S. C. 2011. Age-related changes in tree growth and functional biology: the role of reproduction. – In: Meinzer,

- F. C. et al. (eds), Size- and age-related changes in tree structure and function. *Tree physiology*. Springer, pp. 33–64.
- Thuiller, W. et al. 2005. Niche properties and geographical extent as predictors of species sensitivity to climate change. – *Global Ecol. Biogeogr.* 14: 347–357.
- Tinner, W. et al. 1999. Long-term forest fire ecology and dynamics in southern Switzerland. – *J. Ecol.* 87: 273–289.
- Vandergoes, M. J. et al. 2018. Integrating chronological uncertainties for annually laminated lake sediments using layer counting, independent chronologies and Bayesian age modelling (Lake Ohau, South Island, New Zealand). – *Quat. Sci. Rev.* 188: 104–120.
- Wright, M. N. and Ziegler, A. 2017. ranger: a fast implementation of random forests for high dimensional data in C++ and R. – *J. Stat. Softw.* 77: 1–17.

Supplementary material (Appendix ECOG-04772 at <www.ecography.org/appendix/ecog-04772>). Appendix 1–3.