



Uncertainties in dating constrain model choice for inferring extinction time from fossil records



Frédéric Saltré ^{a, *}, Barry W. Brook ^{a, 1}, Marta Rodríguez-Rey ^a, Alan Cooper ^{a, b}, Christopher N. Johnson ^c, Chris S.M. Turney ^d, Corey J.A. Bradshaw ^a

^a The Environment Institute, School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

^b Australian Centre for Ancient DNA, School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

^c School of Zoology, Private Bag 5, University of Tasmania, Hobart, Tasmania 7001, Australia

^d School of Biological, Earth and Environmental Sciences, University of New South Wales, Australia

ARTICLE INFO

Article history:

Received 2 August 2014
Received in revised form
23 January 2015
Accepted 24 January 2015
Available online

Keywords:

Model selection key
Extinction
Time series
Sensitivity analysis
Dating

ABSTRACT

Accurate estimates of the timing of extinctions (θ) are critical for understanding the causes of major die-off events and for identifying evolutionary or environmental transitions. Yet many studies have demonstrated that sampling biases and underlying statistical assumptions affect the accuracy of model-based estimates of extinction times ($\hat{\theta}$), and the added uncertainty contributed by inherent (laboratory) dating errors has largely been neglected. Here we provide a general guide (model-selection key) for choosing from among eight alternative 'frequentist sampling' (i.e., non-Bayesian) methods, differentiated by their treatment of both the probability of record occurrence and uncertainties in record dates, the most appropriate for a given record. We first provide a methodological framework to characterize time series of dated records as a function of the number of records, the size of the interval between successive records, and laboratory dating errors. Using both simulated data and dated Australian megafauna remains, we then assess how the characteristic of a dataset's time series dictates model performance and the probability of misclassification (false extant vs. false extinct). Among the four classic frequentist methods providing highest model performance, Marshall's (1997) and McCarthy's (1998) methods have the highest model precision. However, high model performance did not prevent misclassification errors, such that the Gaussian-resampled inverse-weighted McInerney (GRIWM) approach is the only method providing both high model accuracy and no misclassification issues, because of its unique down-weighting interval procedure and its ability to account for uncertainties in record dates. Applying the guideline to three time series of extinct Australian species, we recommend using Marshall's, McCarthy's and/or GRIWM methods to infer θ of both *Thylacinus* sp. and *Genyornis* sp., because each dataset is characterized by many sightings and a low variance of the interval between records, whereas McInerney's method better suits *Diprotodon* sp. due to an even lower interval variance.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Mass extinction events, characterized by palaeontologists as high, planetary-wide species loss within a short geological time frame (e.g., over 75% of species within less than two million years, Barnosky et al., 2011), completely changed the global pattern of species distribution by both removing lineages and triggering

evolutionary opportunities (Jablonski, 2001). However, the causes and mechanisms of mass extinctions, such as the end-Permian mass extinction (Grice et al., 2005; Payne and Clapham, 2012; Sun et al., 2012; Wang et al., 2014) or the late Quaternary megafauna extinction, are still debated by scientists from disciplines spanning palaeontology to archaeology and ecology (Alroy, 2001; Brook and Bowman, 2002; Barnosky et al., 2004; Lorenzen et al., 2011), in large part because of inaccuracy of inference of the timing of a species' extinctions (θ) (Flannery, 2002). Robust and accurate inferences are essential to test, for example, the evidence that the end-Permian transition was abrupt versus having multiple extinction phases (Jin et al., 2000; Song et al., 2013; Wang et al.,

* Corresponding author.

E-mail address: frederik.saltre@adelaide.edu.au (F. Saltré).

¹ Present address: School of Biological Sciences, Faculty of Science, Engineering & Technology, University of Tasmania, Hobart, Tasmania 7001, Australia.

Table 1

Description of the eight methods tested and categorized into five categories as a function of the kind of assumptions they make about sampling intensity over time (p-sampling assumption) and summary dataset characteristics (n , \bar{i} , $\sigma^2 i$, \bar{e} , $\sigma^2 e$; see Table 2 for complete description). For each method, we indicated model constraints (high performance constraints) leading to its best performance from the sensitivity analysis (see Fig. 3 and Fig. A.6). For example, $\uparrow x$ means that a high value of the 'x' parameter leads to high model performance, considering that the number of arrows indicates the relative constraint intensity (i.e., $\uparrow \uparrow > \uparrow$ and $\downarrow \downarrow > \downarrow$).

Method	p-Sampling assumption	n	\bar{i}	$\sigma^2 i$	\bar{e}	$\sigma^2 e$	High performance constraints
Strauss and Sadler (1989)							$\uparrow \uparrow n$
McInerney et al. (2006)		x	x	x	–	–	$\uparrow n, \downarrow \downarrow \sigma^2 i$
BRIWM	Poisson stationary process						$\uparrow n, \downarrow \bar{i}$
Solow et al. (2006)		x	x	x	x	–	$\downarrow \downarrow \bar{e}, \downarrow \sigma^2 i$
McCarthy (1998) Marshall (1997)	Recovery potential	x	x	x	–	–	$\uparrow n, \downarrow \sigma^2 i$
Roberts and Solow (2003)	No assumptions	x	x	x	–	–	$\uparrow \uparrow n, \downarrow \bar{i}$
GRIWM (Bradshaw et al., 2012a)		x	x	x	x	x	$\uparrow n, \downarrow \bar{i}, \downarrow \bar{e}$

2014), or that megafauna extinction was primarily climate- or human-driven in South America (Johnson et al., 2013; Lima-Ribeiro and Felizola Diniz-Filho, 2013) and Australia (Brook and Bowman, 2002; Wroe et al., 2013).

The megafauna extinction stalemate in particular persists primarily because the estimated timing of these species' extinctions ($\hat{\theta}$) is uncertain due to the variable quality of the dated precursor fossil specimens, meaning that debates digress to matters of opinion rather than accurately measured phenomena and scientific hypothesis testing (Brook et al., 2013). Although quality fossil data are essential to improve our inferences of past extinctions, palaeo-ecological archives are inherently incomplete and geochronological dating methods are characterized by errors of centuries to millennia, so the reliability of θ inference based only on their scant information remains a major challenge. The absence of a species in a particular site or temporal window does not necessarily mean it was not present, so apparent declines of taxa in these records might simply reflect sampling artefacts rather than real trends in diversity (Prideaux et al., 2007). Such absences might also arise for taphonomic reasons (i.e., type of facies and sedimentary environments that can prevent the preservation of remains), life-history traits (e.g., taxa from lower trophic levels, because they are more abundant, have a relatively higher potential for fossilization) and ecological specialization (i.e., specialists living in a specific habitat will have their remains fossilised only there, whereas generalists will have an overall higher probability of being recorded). Evidence from extinctions observed in modern times suggests that as a doomed species approaches its final extinction date, population size tends to decrease exponentially due to the synergistic feedbacks (Brook et al., 2008) that lead to the extinction vortex (Fagan and Holmes, 2006), which reduces the probability of discovering fossil records near the terminal date and artificially truncates the true temporal range of a species' persistence window (Signor-Lipps effect; Signor and Lipps, 1982). Moreover, fossil records – retrieved from specific sites where the rare phenomenon of preservation was possible – only describe local losses of species such that the last date known cannot necessarily testify to a global extinction. Indeed, in some cases apparent disappearances can be followed by the subsequent reappearance of the species after further sampling (the 'Lazarus' effect; Keith and Burgman, 2004).

As population size tends to decline to incrementally lower densities prior to extinction (Fagan and Holmes, 2006), it is logical to assume that the last dated record of a species occurs sometime before its true extinction (i.e., the death of the last individual). Based on this assumption, many probabilistic methods (also called "classical frequentist methods", Alroy, 2014) have been developed to provide a confidence interval around $\hat{\theta}$ given a particular time series of occurrence records, but uncertainties in dating techniques (e.g., inherent laboratory errors in radiometric dating), and the probability of sampling reliably dated specimens (i.e., sampling rate and location) make inference complex. For example, Roberts and

Solow (2003) applied an optimal linear estimation method based on a record of historical sightings of the dodo (*Raphus cucullatus*) to determine the confidence interval surrounding its true extinction year. That method was extended to account explicitly for error in estimates of the record date for fossils (Solow et al., 2006), but comparisons within and among species were still difficult due to variation in sampling rates that can affect model performance (Rivadeneira et al., 2009). McInerney et al. (2006) proposed another frequentist-probabilistic method that incorporates sampling rate, which was further modified by Bradshaw et al. (2012a) to take into account the number and uncertainty of dates in the time series.

Each method is characterized by a set of statistical assumptions conditioning its adequate application to a given time series (e.g., sampling probability uniformly distributed and independent, or dating error being constant; Table 1 and Solow et al., 2006), which if violated, can lead to the misclassification of a species as extinct or extant (so-called Type I and II statistical inference errors, respectively; Brosi and Biber, 2008; Jarić and Ebenhard, 2010; Fisher and Blomberg, 2012). In addition to methodological issues, the quality (number of records, record interval, variation in dating error over time) and the reliability of the datasets used to infer θ (e.g., species misidentification – Rasmussen and Prys-Jones, 2003; an erroneous ceiling on apparent dates due to the time limit of radiocarbon [^{14}C] dating validity – Walker, 2005) also strongly affect model performance (Rivadeneira et al., 2009; Solow et al., 2011; Bradshaw et al., 2012a; Lee et al., 2014). Various classical frequentist methods have been tested and validated as a function of both the number of records and sampling intensity (Rivadeneira et al., 2009; Fisher and Blomberg, 2012), highlighting performance problems specifically when sampling probabilities decrease through time (Rivadeneira et al., 2009). Newly emerging Bayesian methods can, if used appropriately, reduce such performance issues and improve species classification (endangered or about to go extinct; Alroy, 2014), but the effect of inherent dating error and their variation over time on model performance have barely been assessed (Bradshaw et al., 2012a). As dating errors typically increase as sampling reaches deeper back in time (such as in palaeontological time series; Walker, 2005), providing rigorous measures of the biases generated by dating errors on $\hat{\theta}$ is therefore essential.

Here we explore how the characteristics of time series of dated records, such as the number of occurrences, time gaps between records, and uncertainties in measured dates, act and interact to constrain different frequentist models used commonly to infer θ . More specifically, we provide both quantitative and qualitative criteria for: (i) maximizing the inferential capability of eight classical methods used to generate confidence intervals for θ ; and (ii) provide a general guideline for selecting the most appropriate method to infer θ from a given time series of dated records. We first describe these eight frequentist methods focussing on their conceptual assumptions with respect to five summary variables characterizing the types of time series usually available (henceforth,

Download English Version:

<https://daneshyari.com/en/article/6446398>

Download Persian Version:

<https://daneshyari.com/article/6446398>

[Daneshyari.com](https://daneshyari.com)