

Time lags and the invasion debt in plant naturalisations

Richard Duncan¹

¹University of Canberra

February 11, 2021

Abstract

Ecological processes often exhibit time lags. For plant invasions, lags of decades to centuries between species' introduction and establishment in the wild (naturalisation) are common, leading to the idea of an invasion debt: accelerating rates of introduction result in an expanding pool of introduced species that will naturalise in the future. Here, I show how a concept from survival analysis, the hazard function, provides an intuitive way to understand and forecast time lags. For plant naturalisation, theoretical arguments predict that lags between introduction and naturalisation will have a unimodal distribution, and that increasing horticultural activity will cause the mean and variance of lag times to decline over time. These predictions were supported by data on introduction and naturalisation dates for plant species introduced to Britain. While increasing trade and horticultural activity can generate an invasion debt by accelerating introductions, the same processes could lower that debt by reducing lag times.

Introduction

Many ecological processes are characterised by time lags: a delay between initiation of a process and its outcome. The concept of an extinction debt, for example, arises from time delays between the start of processes that erode biodiversity and when species go extinct (Tilman *et al.* 1994; Kuussaari *et al.* 2009). Biodiversity recovery following conservation interventions can likewise be delayed due to temporal lags in species' responses to those actions (Watts *et al.* 2020). Time lags are a particular feature of biological invasions (Williamson 1996; Crooks 2005; Aikio *et al.* 2010). For example, delays of decades to centuries between introduction and establishment in the wild (naturalisation) have been documented for many invasive plant species (Kowarik 1995; Caley *et al.* 2008; Daehler 2009; van Klinken *et al.* 2015), leading to the concept of an invasion debt. Accelerating global trade and transport, leading to dramatic increases in the numbers of species introduced to regions outside their native range (Meyerson & Mooney 2007; Hulme 2009; Bradley *et al.* 2012; Seebens *et al.* 2015; Sikes *et al.* 2018), has resulted in an expanding pool of introduced species that will naturalise in the future, some proportion of which will become problematic invaders (Esslet *et al.* 2011; Rouget *et al.* 2016; Dehnen-Schmutz & Conroy 2018; Haeuser *et al.* 2018).

The examples above imply that the consequences of certain actions are yet to play out, with the trajectory of future ecological outcomes depending on the nature of time lags associated with those actions. While we can often identify factors likely to cause time lags in ecological processes (e.g., Geerts *et al.* 2013; Tenhumberg *et al.* 2018; Watts *et al.* 2020), we currently lack a framework for quantitatively analysing lag times and forecasting their consequences, such as the size of an invasion or extinction debt. My aim in this paper is to show how approaches from survival analysis can be used to understand time lags arising from stochastic processes, focusing on the lag between introduction and naturalisation, and the associated invasion debt in plant naturalisations. I show how the hazard function provides an intuitive way to understand how the risk of an outcome changes over time, with the time-varying shape of the hazard function determining the expected distribution of lag times. Consequently, we can use data on lag times to assess how risk has changed over time, and we can predict the shape of lag time distributions given expectations about changing risk and evaluate those predictions against data.

For plant naturalisations, I first outline in theory how the shape of the hazard function, and hence the risk of naturalisation, should increase for many species following introduction, and how the hazard function should steepen for more recently introduced species. These shifts in the shape of the hazard function imply predictable changes over time in the distribution of the lag times between introduction and naturalisation. I test these predictions using data from Britain, show that the data match the theory well, describe how the approach can be used to estimate the invasion debt, and consider the implications of these findings.

Materials and Methods

Theoretical distribution of lag times for plant naturalisation

Globally an estimated 170000 plant species are cultivated in gardens and parks and most of the world's approximately 13000 naturalised plant species are garden escapees (Reichard & White 2001; Niemiera & Holle 2009; Bradley *et al.* 2012; van Kleunen *et al.* 2018). In some regions, historical nursery and planting records provide estimates of the time lag between introduction and naturalisation, which vary widely among species ranging from years to centuries (Kowarik 1995; Caley *et al.* 2008; Daehler 2009; van Klinken *et al.* 2015). What underlies such extreme variation? And what form should lag time distributions take?

Lag time can be thought of as a random variable resulting from the stochastic nature of plant naturalisations (Mack 2000; Rejmánek *et al.* 2005). Consider that each plant grown in cultivation has some probability its propagules will establish a naturalised population per unit time. The probability a species will naturalise per time interval should depend on the number of plants of that species in cultivation, with more plants releasing more propagules into the environment, any one of which could establish a naturalised population if they encounter the right conditions (Duncan *et al.* 2014). In line with this, planting effort is often the strongest predictor of whether plant species have naturalised or not (Rejmánek *et al.* 2005; Dehnen-Schmutz *et al.* 2007; Pyšek *et al.* 2009; McGregoret *et al.* 2012; Maurel *et al.* 2016).

The time lag between introduction and naturalisation should depend on the probability of naturalisation per unit time, with a higher probability implying a shorter time lag. For a given species, however, naturalisation probability could change over time following introduction if the number of plants in cultivation changes. An increase in the number of individuals of a species grown in gardens and parks over time, for example, should cause the naturalisation risk to increase. Borrowing from survival analysis, the change in naturalisation risk over time can be quantified using a hazard function, with hazard defined as the instantaneous rate of naturalisation, a number proportional to the probability a species will naturalise in the next time step given it has not yet naturalised (Muenchow 1986; Tableman & Kim 2003).

The time-varying shape of the hazard function determines the shape of the lag time distribution (and vice versa, see Appendix S1). If the hazard for a group of species remains constant or declines over time following introduction, their lag time distribution will be monotonically declining (Fig 1A). For species that become popular garden plants, however, we expect the hazard to increase over time following introduction as they become widely planted (Mack 2000). If the hazard increases over time following introduction, lag times will have a unimodal distribution (Fig 1B-D).

Changing lag time distributions

For popular species, planting effort should increase over time following introduction. In addition, we expect total planting effort across all species to increase over time due to increasing horticultural activity driven by an ever-growing human population and associated demand for garden and amenity plantings (Lawson 1996; Dehnen-Schmutz *et al.* 2010; Drew *et al.* 2010; Humair *et al.* 2015). Because of this, relative to a species first introduced in the year 1800, we would expect a similar species introduced in 2000 to have an initially faster rate of planting following introduction. Consequently, an increase in horticultural activity over time should result in steeper hazard functions for more recently introduced species relative to earlier introductions. A steepening of the hazard function over time could also occur through land-use changes that, over the long-term, result in more modified habitats providing greater opportunities for plants to establish in the wild (Hobbs & Huenneke 1996; Essl *et al.* 2011; González-Moreno *et al.* 2017). Hence, both increasing

horticultural activity and greater habitat modification should increase the naturalisation risk for more recent introductions, hastening the rate at which species escape into the wild and establish.

I used a Weibull distribution and its associated hazard function to explore how changes in the shape and steepness of the hazard function alter the shape of the lag time distribution. The Weibull distribution is commonly used to model time-to-event data because it has a flexible hazard function that can take a variety of shapes (Carroll 2003; Tableman & Kim 2003; Fig. 1 and Appendix S1). If the hazard is constant over time following introduction, the Weibull simplifies to a negative exponential distribution (Fig 1A). If the hazard increases over time following introduction, the Weibull can model that increase as downward accelerating, linear or upward accelerating (Fig 1B-D). In Appendix S1 I show that, regardless of its shape, as the slope of the hazard function steepens (the different coloured lines in Fig. 1 panels), both the mean and variance of the lag time distribution decline. This outcome is evident in Figure 1B-D, with lag time distributions shifting toward zero and becoming more peaked as the hazard functions steepen. Consequently, if the hazard function for introduced species has steepened over time, we expect species introduced in the more distant past to have, on average, a longer lag time with greater among-species variation in their lag times, while more recent introductions should have, on average, shorter lag times with less among-species variation. I next test these predictions using data on the lag times of naturalised plants introduced to Britain.

Data collection

I obtained a list of all non-native flowering plants recorded as naturalised in Britain from Stace & Crawley (2015), which is the most up-to-date published list currently available. I excluded taxa of uncertain origin and taxa listed as casual or survivors, retaining only taxa classed as naturalised, which Stace & Crawley (2015) defined as self-reproducing in the wild in Britain after 1987 (i.e., the list excludes species recorded as naturalised before 1987 that no longer appear to be self-reproducing in the wild). I obtained a list of the naturalised non-flowering plants (gymnosperms and ferns) from Stace (2010) and Hill *et al.* (2005). Both lists included hybrids, cultivars, and subspecies. For convenience, I term each taxon a species. For each species in the combined list, I searched for information on the year the species was first recorded as grown in cultivation in Britain (which I term year of introduction), and the year it was first recorded as growing outside cultivation in Britain (which I term year of naturalisation), using Hill *et al.* (2005), Stace & Crawley (2015) and The Online Atlas of the British and Irish Flora (<https://www.brc.ac.uk/plantatlas/> accessed in December 2020). In Hill *et al.* (2005), the year of introduction was sometimes given as a range when it could not be determined precisely. In those cases, I used the earliest year of introduction, and I used the earliest recorded year of introduction and naturalisation when sources differed.

For the analyses in this paper, I included only naturalised species with data for both the year of first introduction and naturalisation, and I excluded species introduced before 1500 AD because their introduction dates were less certain. The final list comprised 708 naturalised species in 99 plant families. The year in which species were recorded as first introduced ranged from 1500 to 1960, while the year in which species were recorded as first naturalised ranged from 1597 to 2000. I calculated the lag time (in years) as the year of first recorded naturalisation minus the year of first recorded introduction.

I classified each species into one of three life-form categories: tree/shrub (213 species), perennial herb (398 species) or annual/biennial herb (97 species). I did this because we expect that, on average, lag times should differ among these life-form categories, with longer-lived, slower-growing species taking longer to reach reproductive maturity and hence taking longer to produce propagules that could establish wild populations relative to faster-growing, shorter-lived species (Wangen & Webster 2006; Caley *et al.* 2008). The life-form classifications were determined primarily from information in The Online Atlas of the British and Irish Flora (<https://www.brc.ac.uk/plantatlas/>).

Specifying a model for lag time distributions

I modelled lag times as drawn from both a Weibull and normal distribution. The normal distribution closely approximates the Weibull distribution for certain parameter values (e.g., Fig 1D) but differs in being a symmetric distribution with an upward accelerating hazard function that steepens more rapidly than the

Weibull hazard function, thus allowing for a wider range of shapes for the lag time distribution than modelled by the Weibull alone.

I specified truncated distributions for both the normal and Weibull because the data were censored: for species introduced in year Y_i , we can only observe the lag times of species that have naturalised up to the present year Y_p . Other species introduced in year Y_i could naturalise in the future, meaning the observed distribution of lag times is truncated at an upper limit $Y_p - Y_i$. I aimed to estimate the full distribution of lag times by modelling the truncated portion of the distribution using the observed lag times. The number of species likely to naturalise in the future (the invasion debt) can then be estimated from that portion of the full distribution that extends beyond the present (Fig. 2 and Appendix S3). I set the present year Y_p to the year 2000, which was the most recent year of recorded naturalisation for species in the dataset. The most recent year of first introduction was 1960, meaning I estimated the invasion debt associated with plant species introduced to Britain between 1500 and 1960.

I compared the fit of ten models to the data (Table 1) to assess how well the data matched theoretical predictions (Fig. 1 and Appendix S1) and to identify a best-fitting model to infer the full distribution of lag times. Models 1-5 modelled lag times as drawn from a truncated normal distribution, with the five models differing in how the mean and standard deviation were specified: as either constant or changing over time (as predicted by theory), and whether lag times differed among life-forms (tree/shrub, perennial herb and annual/biennial herb; Table 1). Models 6-10 had the same specifications for the mean and standard deviation, but modelled lag times as drawn from a truncated Weibull distribution.

Model fitting

I fitted the models to the lag time data in a Bayesian framework in R version 4.0.3 (R Core Team 2017) using a Metropolis-Hastings algorithm (modified from the R package MHadaptive; Chivers 2012) to estimate the posterior distributions of the parameters. Being a Bayesian model, I specified a prior distribution for all parameters using relatively uninformative priors: a normal distribution with mean zero and standard deviation 10. I ran the models in an adaptive phase for 20000 iterations, using the final 2000 iterations to determine the parameters for the proposal distribution and initial parameter starting values. I then ran three chains for 10000 iterations using the proposal distribution and initial parameter values from the adaptive phase, and checked the resulting parameters for convergence using the Gelman-Rubin statistic (Gelman & Rubin 1992), which was less than 1.1 for all parameters indicating adequate convergence. I identified the best performing model using the approximate leave-one-out cross-validation (LOO) criteria (Vehtari *et al.* 2017), which estimates the predictive accuracy of each model. LOO is considered an improvement on other information-criterion based model selection measures such as AIC, WAIC and DIC that are widely used to compare model performance (see Vehtari *et al.* 2017 for details). I used Pareto smoothed importance sampling to estimate LOO (PSIS-LOO) and compared models by calculating the difference in expected predictive accuracy (Δ PSIS-LOO) between each model and the best-performing model on the deviance scale using the *loo* package in R (Vehtari *et al.* 2019), with smaller PSIS-LOO values implying a model with better predictive accuracy.

Assessing model fit and estimating the invasion debt

I used the best-performing model based on the LOO criteria to model the full lag time distribution and used simulation to assess how well the modelled lag time distribution matched the data. To do this, I simulated a lag time for each species given its actual year of introduction by drawing a random value from the lag time distribution specified by the best-performing model. I carried out 10000 simulated draws for each species, binned the simulated lag times into 20 year intervals, calculated the mean and 95% quantiles for the number of species in each bin, and compared these simulated outcomes with the actual distribution of lag times in 20 year bins.

For a species introduced in year Y_i , the modelled lag time distribution for that year allows us to calculate the probability that the species will naturalise in year Y_t , where $Y_t > Y_i$. The total number of species expected to naturalise in year Y_t can be calculated as the sum of the probabilities that each previously

introduced species will naturalise in year Y_t . We can therefore use this approach to estimate the number of species expected to naturalise in each year beyond the present and sum those estimates to calculate the invasion debt (Fig. 2). The difficulty is that we do not know the total number of species introduced each year that are going to naturalise because we only observe species that have naturalised. In Appendix S3, I show how the fitted lag time distribution coupled with the introduction dates can be used to estimate the total number of species introduced each year that will naturalise, and hence estimate the invasion debt. For each life-form group, I used the methods in Appendix S3 to model the average number of species introduced per year between 1500 and 1960 that have or will naturalise in the foreseeable future, and thus calculate the invasion debt.

Simulation check and uncertainty in introduction and naturalisation dates

I used simulation to check the methods described above worked as expected, using known parameter values to generate artificial data similar to the real data, fitting the lag time model to the artificial data and estimating the invasion debt, and assessing how well the methods recovered the true parameter values using repeated simulations. Full details are in Appendix S4.

The dates of first introduction and naturalisation in this study are taken from historical records and are subject to uncertainty, notably because the true dates might not have been documented or the records have not been uncovered, resulting in first record dates being more recent than true dates. While dates can be updated as new records are discovered, any set of dates will be subject to uncertainty due to the difficulty of determining precisely when past events occurred. The question is whether this uncertainty could affect the results reported in this study. To evaluate this, I used artificial data simulation as described above but added measurement error to the simulated dates. I then assessed how well the methods recovered the true parameter values in the presence of dating errors using repeated simulations. Full details are in Appendix S5.

Results

Figure 2 plots lag time as a function of introduction year for the 708 species naturalised in Britain for which data were available, along with marginal histograms for lag time and year of introduction. An apparent spike in the number of early introductions is evident in the decades 1580-1600 and 1620-1640. These spikes occur because many early introductions were first recorded in two compilations of garden plants published during those periods. While these compilations provide the first record of cultivation for many introduced species, the true dates of introduction for cultivation could have been earlier but were undocumented. Simulations using artificial data revealed that errors in the true dates of introduction and naturalisation should not unduly affect the lag time estimates reported in this study, but that dating errors could result in underestimates of the invasion debt (Appendix S5).

Apart from the two early spikes, the number of introductions tended to increase over time, peaking in the early 1800s before declining. This decline is potentially due to the lag between introduction and naturalisation, which could result in a greater proportion of more recent introductions having not yet naturalised. The lag time distribution peaks at just under 100 years with a long tail. The mean and median lag times between introduction and naturalisation were 145 and 122 years, respectively.

Table 1 compares the performance of the ten candidate models. For all models, the estimated effective number of parameters was close to the actual number of parameters, indicating none of the models were badly mis-specified, and plots of the Pareto k diagnostics indicated most models were well behaved with few outlying observations (Appendix S2). Model 4, which specified a truncated normal distribution of lag times with mean and standard deviation changing over time, and mean lag time differing by plant life-form, was the best-performing of the ten candidate models (Table 1). Model 5, which specified that the standard deviation also differed by life-form, was the second-best performing and close in predictive accuracy to model 4. Nevertheless, there seemed no reason to favour the more complex model 5 given that model 4 performed slightly better. Apart from model 5, model 4 performed substantially better than other models: the difference in PSIS-LOO between model 4 and model 10 (the next best-performing after model 5) was

nearly twice the standard error of their difference (roughly equivalent to a 95% confidence interval) suggesting that, allowing for model uncertainty, we could be reasonably confident that model 4 performed better. I therefore used model 4 for subsequent inferences about the distribution of lag times. Simulations using artificial data revealed that fitting a truncated distribution to censored lag time data accurately recovered the true parameter values (Appendix S4), and that parameter estimates were not unduly affected by dating errors (Appendix S5).

The parameter estimates for model 4 revealed two features of the lag time distribution. First, as predicted, both the mean and standard deviation of lag times declined over time (Fig 3B) causing the lag time distribution to shift towards zero and become narrower and more peaked for more recent introductions (Fig. 4). The lag time distribution for trees/shrubs introduced in 1500, for example, had a mean of 571 years and a standard deviation of 169 years, while the distribution for trees/shrubs introduced in 1960 had a mean of 70 years and a standard deviation of 26 years, equating to an overall decline in mean lag time of around 100 years per century. The hazard functions associated with model 4 were all upward accelerating and steepened appreciably for more recent introductions (Fig 3D-E).

Second, trees/shrubs had a longer mean lag time than perennial herbs, which in turn had a longer mean lag time than biennial/annual species (Fig. 3C). For species first introduced in 1500, the predicted mean lag times for trees/shrubs, perennial herbs and biennial/annual species were 571, 455 and 377 years, respectively, although these differences narrowed over time: for species introduced in 1960, the predicted mean lag times were 70, 56 and 46 years, respectively.

The distribution of lag times predicted by model 4 fitted the observed distribution of lag times well, both overall (Fig 4A) and when naturalised species were split by their century of introduction (Fig 4B-F). In almost all cases, the actual number of species in the 20 year lag time bins were within the 95% quantiles obtained from the model simulations.

The estimated number of species introduced to Britain per year that have or will naturalise has increased over time for trees/shrubs and perennial herbs (curved lines in Fig 5A, C) but remained relatively constant for annual/biennial species (Fig 5E). The number of species predicted to naturalise every 20 years, based on introductions prior to 1960 and the fitted lag time model, matched the observed data well (Fig 5B, D, F). For each life-form, the area under the predicted distribution that extended beyond the year 2000 (shaded red in Fig 5) estimated the size of the invasion debt. These areas indicated that, in addition to the 708 species already naturalised, a further 84 trees/shrubs (95% credible interval 63-110), 65 perennial herbs (48-86) and 9 biennial/annual species (5-14) are expected to naturalise within the next 150 years; a total of 158 species (116-210). These figures could be underestimates because artificial data simulations revealed that dating errors can downwardly bias estimates of invasion debt (Appendix S5).

Discussion

Time lags are a feature of many ecological processes. Here, I have shown how hazard functions underpin the distribution of lag times arising from stochastic processes, and consequently how data on lag times can be used to examine changes in the risk of an outcome over time. I applied this framework to understanding lag times between plant introduction and naturalisation and the associated invasion debt, but the approach has wider applicability. It could, for example, be used to estimate the size of an extinction debt from the pattern of extinctions over time (Elphick *et al.* 2010).

For plant naturalisations in Britain, the key finding is that lag times between introduction and naturalisation have a unimodal distribution with mean and variance that have declined substantially over the last 500 years, implying that, on average: a) for naturalised species, hazards increase over time following introduction; and b) hazard functions are steeper for more recently introduced species.

Changes over time in the shape of the lag time distribution have consequences for the invasion debt, implying that recently introduced species with the potential to naturalise will realise that potential faster than species introduced in the past. The total number of species introduced to Britain for cultivation has likely increased

substantially over the last century in line with global gardening trends (Lawson 1996; Drew *et al.* 2010; Humair *et al.* 2015; Dehnen-Schmutz & Conroy 2018). The analyses presented here suggest that the invasion debt associated with introductions prior to 1960 has largely been realised: just over 80% (708/866) of species purposefully introduced to Britain before 1960 that were capable of naturalising are estimated to have naturalised by the year 2000 (Fig 5).

Of the species yet to naturalise, most are predicted to be longer-lived trees or shrubs. This likely reflects the fact that trees/shrubs have, on average, a longer lag-phase than shorter-lived perennial herbaceous and biennial/annual species (Fig 3C; see also Caley *et al.* 2008). But it also results from a shifting trend in introductions, with the numbers of newly introduced trees/shrubs capable of naturalising increasing at a faster rate over time than perennial herbs and annual/biennial species (compare the curved lines in Fig 5 A, C, E).

The decline in the mean and variance of the lag time distribution (Fig. 4) and the associated steepening of hazard functions (Fig. 3D-F) imply that the probability of a species naturalising per time interval has increased over time in Britain. I have argued that a reduction in lag times is most likely driven by increased horticultural activity associated with accelerating human population growth. For species that become widely planted, and hence more likely to naturalise, an increase over time in the rate at which species are propagated, distributed and planted will accelerate the rate at which cultivated plants can supply propagules, increasing the probability that species capable of naturalising will do so in a given time period. Changes in other factors that increase naturalisation probability could also contribute to or account for changes in the lag time distribution. In particular, an increase over time in the extent and severity of habitat disturbance could result in more opportunities for plants to naturalise in increasingly modified environments (Hobbs & Huenneke 1996; Essl *et al.* 2011; González-Moreno *et al.* 2017). Changing climate could also alter naturalisation probability (Dullinger *et al.* 2017; Haeuser *et al.* 2018) although anthropogenic climate change, whose impacts have been most pronounced in the last 100 years, is unlikely to account for the sustained decline in the mean and variance of the lag time distribution in Britain over the last 500 years (Fig. 4). Finally, lag times could shorten through a systematic change in recording effort if, for example, efforts to document newly naturalised species have increased over time. The simulations assessing the effects of dating errors reported in Appendix S5 did not consider a systematic shift in recording effort over time, but such a shift would have to be large to generate the observed outcomes.

The overall mean lag-phase in this study (145 years) was similar to the mean lag times calculated for temperate woody species in two previous studies (Kowarik 1995; Caley *et al.* 2008), but longer than recorded for introduced grasses in Australia (van Klinken *et al.* 2015) and trees in Hawaii (Daehler 2009). Here, and in Caley *et al.* (2008), the mean lag time between introduction and naturalisation differed by plant life-form, with longer-lived species having longer lag times. Nevertheless, the difference in mean lag time between trees/shrubs and biennial/annual species (around 190 years for species introduced in 1500) was substantially less than the decline over time for trees/shrubs (mean lag time declined by about 500 years from 1500 to 1960). Hence, lag times are not a fixed property of species or locations but vary markedly in response to factors affecting naturalisation probability that can change over time, such as planting effort and habitat modification. While other plant traits have the potential to influence lag times, the extent to which lag time distributions have changed over time in Britain suggests that temporally varying factors may be as or more important than species traits in determining the length of the lag phase among species capable of naturalising (see also Mack 2000).

Britain, like other European countries, may be unique in having a long history of plant introductions and records that allow lag times to be reconstructed (see also Kowarik 1995). It is unclear whether the marked shift in lag time distributions over almost five centuries in Britain would be replicated in other regions with a shorter history of sustained plant introductions, such as followed European colonisation of the Americas, Australia, southern Africa and New Zealand. I predict the same processes and outcomes will be operating, albeit with a compressed timeframe in regions with a shorter introduction history. Consequently, the global invasion debt associated with non-native species already introduced for horticulture may be substantially less

than expected based on previously reported lag times (Kowarik 1995; Caley *et al.* 2008; Essl *et al.* 2011; van Klinken *et al.* 2015; but see Daehler 2009). This has implications for preventing future plant naturalisations. It suggests that, while an existing pool of cultivated species in a region will contain species capable of naturalising in the future, that risk may be lower than expected, and lower than the risk posed by future introductions of new species, which are capable of naturalising more rapidly than in the past. Models that aim to predict future invasion trajectories need to account for the changing nature of lag time distributions (Leung *et al.* 2012; Essl *et al.* 2019).

ACKNOWLEDGEMENT

I am very grateful to Tim Blackburn, Lizzie Wandrag and three reviewers for insightful and very helpful comments on previous versions of the manuscript.

References

- Aikio, S., Duncan, R.P. & Hulme, P.E. (2010). Lag-phases in alien plant invasions: separating the facts from the artefacts. *Oikos* , 119, 370–378.
- Bradley, B.A., Blumenthal, D.M., Early, R., Grosholz, E.D., Lawler, J.J., Miller, L.P., *et al.* (2012). Global change, global trade, and the next wave of plant invasions. *Frontiers in Ecology and the Environment* , 10, 20–28.
- Caley, P., Groves, R.H. & Barker, R. (2008). Estimating the invasion success of introduced plants. *Diversity and Distributions* , 14, 196–203.
- Carroll, K.J. (2003). On the use and utility of the Weibull model in the analysis of survival data. *Controlled Clinical Trials* , 24, 682–701.
- Chivers, C. (2012). *MHadaptive: general Markov Chain Monte Carlo for Bayesian inference using adaptive Metropolis-Hastings sampling* . <https://cran.r-project.org/web/packages/MHadaptive/index.html>.
- Crooks, J.A. (2005). Lag times and exotic species: The ecology and management of biological invasions in slow-motion. *Ecoscience* , 12, 316–329.
- Daehler, C.C. (2009). Short lag times for invasive tropical plants: evidence from experimental plantings in Hawai'i. *PLOS ONE* , 4, e4462.
- Dehnen-Schmutz, K. & Conroy, J. (2018). Working with gardeners to identify potential invasive ornamental garden plants: testing a citizen science approach. *Biol Invasions* , 20, 3069–3077.
- Dehnen-Schmutz, K., Holdenrieder, O., Jeger, M.J. & Pautasso, M. (2010). Structural change in the international horticultural industry: Some implications for plant health. *Scientia Horticulturae* , 125, 1–15.
- Dehnen-Schmutz, K., Touza, J., Perrings, C. & Williamson, M. (2007). The horticultural trade and ornamental plant invasions in Britain. *Conservation Biology* , 21, 224–231.
- Drew, J., Anderson, N. & Andow, D. (2010). Conundrums of a complex vector for invasive species control: a detailed examination of the horticultural industry. *Biol Invasions* , 12, 2837–2851.
- Dullinger, I., Wessely, J., Bossdorf, O., Dawson, W., Essl, F., Gattringer, A., *et al.* (2017). Climate change will increase the naturalization risk from garden plants in Europe. *Global Ecology and Biogeography* , 26, 43–53.
- Duncan, R.P., Blackburn, T.M., Rossinelli, S. & Bacher, S. (2014). Quantifying invasion risk: the relationship between establishment probability and founding population size. *Methods Ecol Evol* , 5, 1255–1263.
- Elphick, C.S., Roberts, D.L. & Michael Reed, J. (2010). Estimated dates of recent extinctions for North American and Hawaiian birds. *Biological Conservation* , 143, 617–624.

- Essl, F., Dullinger, S., Rabitsch, W., Hulme, P.E., Hulber, K., Jarošík, V., *et al.* (2011). Socioeconomic legacy yields an invasion debt. *PNAS* , 108, 203–207.
- Essl, F., Lenzner, B., Courchamp, F., Dullinger, S., Jeschke, J.M., Kühn, I., *et al.* (2019). Introducing AlienScenarios: a project to develop scenarios and models of biological invasions for the 21 st century. *NeoBiota* , 45, 1–17.
- Geerts, S., Moodley, D., Gaertner, M., Roux, J.J.L., McGeoch, M.A., Muofhe, C., *et al.* (2013). The absence of fire can cause a lag phase: The invasion dynamics of *Banksia ericifolia* (Proteaceae). *Austral Ecology* , 38, 931–941.
- Gelman, A. & Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* , 7, 457–472.
- González-Moreno, P., Pino, J., Cózar, A., García-de-Lomas, J. & Vilà, M. (2017). The effects of landscape history and time-lags on plant invasion in Mediterranean coastal habitats. *Biol Invasions* , 19, 549–561.
- Haeuser, E., Dawson, W., Thuiller, W., Dullinger, S., Block, S., Bossdorf, O., *et al.* (2018). European ornamental garden flora as an invasion debt under climate change. *Journal of Applied Ecology* , 55, 2386–2395.
- Hill, M., Baker, R., Broad, G., Chandler, P., Copp, G., Ellis, J., *et al.* (2005). *Audit of non-native species in England* . English Nature Research Reports. Centre for Ecology & Hydrology.
- Hobbs, R.J. & Huenneke, L.F. (1996). Disturbance, diversity, and invasion: implications for conservation. In: *Ecosystem Management: Selected Readings* (eds. Samson, F.B. & Knopf, F.L.). Springer, New York, NY, pp. 164–180.
- Hulme, P.E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* , 46, 10–18.
- Humair, F., Humair, L., Kuhn, F. & Kueffer, C. (2015). E-commerce trade in invasive plants. *Conservation Biology* , 29, 1658–1665.
- Justus, C.G., Hargraves, W.R., Mikhail, A. & Graber, D. (1978). Methods for estimating wind speed frequency distributions. *J. Appl. Meteor.* , 17, 350–353.
- van Kleunen, M., Essl, F., Pergl, J., Brundu, G., Carboni, M., Dullinger, S., *et al.* (2018). The changing role of ornamental horticulture in alien plant invasions. *Biological Reviews* , 93, 1421–1437.
- van Klinken, R.D., Panetta, F.D., Coutts, S. & Simon, B.K. (2015). Learning from the past to predict the future: an historical analysis of grass invasions in northern Australia. *Biol Invasions* , 17, 565–579.
- Kowarik, I. (1995). Time lags in biological invasions with regard to the success and failure of alien species. In: *Plant invasions: general aspects and special problems* (eds. Pyšek, P., Prach, K., Rejmánek, M. & Wade, M.). SPB Academic Publishing, Amsterdam, pp. 15–38.
- Kuussaari, M., Bommarco, R., Heikkinen, R.K., Helm, A., Krauss, J., Lindborg, R., *et al.* (2009). Extinction debt: a challenge for biodiversity conservation. *Trends in Ecology & Evolution* , 24, 564–571.
- Lawson, R.H. (1996). Economic importance and trends in ornamental horticulture. *Acta horticultrae* , 432, 226–237.
- Leung, B., Roura-Pascual, N., Bacher, S., Heikkilä, J., Brotons, L., Burgman, M.A., *et al.* (2012). TEASing apart alien species risk assessments: a framework for best practices. *Ecology Letters* , 15, 1475–1493.
- Mack, R.N. (2000). Cultivation fosters plant naturalization by reducing environmental stochasticity. *Biological Invasions* , 2, 111–122.
- Maurel, N., Hanspach, J., Kuhn, I., Pyšek, P. & Kleunen, M. van. (2016). Introduction bias affects relationships between the characteristics of ornamental alien plants and their naturalization success. *Global*

Ecology and Biogeography , 25, 1500–1509.

McGregor, K.F., Watt, M.S., Hulme, P.E. & Duncan, R.P. (2012). What determines pine naturalization: species traits, climate suitability or forestry use? *Diversity Distrib.* , 18, 1013–1023.

Meyerson, L.A. & Mooney, H.A. (2007). Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment* , 5, 199–208.

Muenchow, G. (1986). Ecological use of failure time analysis. *Ecology* , 67, 246–250.

Niemiera, A.X. & Holle, B.V. (2009). Invasive Plant Species and the Ornamental Horticulture Industry. In: *Management of Invasive Weeds* , Invading Nature – Springer Series In Invasion Ecology (ed. Inderjit). Springer Netherlands, Dordrecht, pp. 167–187.

Pyšek, P., Krivánek, M. & Jarošík, V. (2009). Planting intensity, residence time, and species traits determine invasion success of alien woody species. *Ecology* , 90, 2734–2744.

R Core Team. (2017). *R: A language and environment for statistical computing* . <https://www.R-project.org/>.

Reichard, S.H. & White, P. (2001). Horticulture as a pathway of invasive plant introductions in the United States. *BioScience* , 51, 103–113.

Rejmánek, M., Richardson, D.M., Higgins, S., Pitcairn, M. & Grotkopp, E. (2005). Ecology of invasive plants: state of the art. In: *Invasive alien species: searching for solutions* (eds. Mooney, H., Mack, R., McNeely, J., Neville, L., Schei, P. & Waage, J.). Island Press, Washington, D.C., pp. 104–161.

Rouget, M., Robertson, M.P., Wilson, J.R.U., Hui, C., Essl, F., Renteria, J.L., *et al.* (2016). Invasion debt – quantifying future biological invasions. *Diversity and Distributions* , 22, 445–456.

Seebens, H., Essl, F., Dawson, W., Fuentes, N., Moser, D., Pergl, J., *et al.* (2015). Global trade will accelerate plant invasions in emerging economies under climate change. *Global Change Biology* , 21, 4128–4140.

Sikes, B.A., Bufford, J.L., Hulme, P.E., Cooper, J.A., Johnston, P.R. & Duncan, R.P. (2018). Import volumes and biosecurity interventions shape the arrival rate of fungal pathogens. *PLoS Biol* , 16, e2006025.

Stace, C. (2010). *New Flora of the British Isles* . Cambridge University Press.

Stace, C.A. & Crawley, M.J. (2015). *Alien Plants* . William Collins.

Tableman, M. & Kim, J.S. (2003). *Survival analysis using S: analysis of time-to-event data* . Chapman and Hall/CRC.

Tenhumberg, B., Crone, E.E., Ramula, S. & Tyre, A.J. (2018). Time-lagged effects of weather on plant demography: drought and *Astragalus scaphoides*. *Ecology* , 99, 915–925.

Tilman, D., May, R.M., Lehman, C.L. & Nowak, M.A. (1994). Habitat destruction and the extinction debt. *Nature* , 371, 65–66.

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y. & Gelman, A. (2019). *loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models* . <https://mc-stan.org/loo>.

Vehtari, A., Gelman, A. & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput* , 27, 1413–1432.

Wangen, S.R. & Webster, C.R. (2006). Potential for multiple lag phases during biotic invasions: reconstructing an invasion of the exotic tree *Acer platanoides* . *Journal of Applied Ecology* , 43, 258–268.

Watts, K., Whytock, R.C., Park, K.J., Fuentes-Montemayor, E., Macgregor, N.A., Duffield, S., *et al.* (2020). Ecological time lags and the journey towards conservation success. *Nature Ecology & Evolution* , 4, 304–311.

Williamson, M. (1996). *Biological Invasions* . Springer Science & Business Media.

Table 1. Ten candidate models to describe lag time distributions for naturalised plants in Britain. Models 1-5 specified a truncated normal distribution with $T(\cdot)$ the upper limit for truncation. Y_p is the present (set to the year 2000), Y_i is the year of introduction of the i th species and P_i and A_i are variables specifying whether the i th species was perennial ($P_i = 1$) or not ($P_i = 0$), and annual ($A_i = 1$) or not ($A_i = 0$). The β are the parameters estimated for each model. Log-transformations of the mean [e.g., $\log(\mu_{Y_i})$] and standard deviation [e.g., $\log(\sigma_{Y_i})$] ensured these remained positive. Models 6-10 specified a truncated Weibull distribution parameterised in terms of a shape (k) and scale (λ) parameter, with k and λ related to the mean and standard deviation as shown (Justus *et al.* 1978). Δ PSIS-LOO is the difference in PSIS-LOO (Pareto smoothed importance sampling leave-one-out cross-validation) between each model and the best fitting model, which had Δ PSIS-LOO = 0 (shown in bold).

Lag time distribution	Model	Mean $\log(\mu_{Y_i}) =$	Standard deviation $\log(\sigma_{Y_i}) =$	Δ PSIS-LOO	Standard error of Δ PSIS-LOO
Normal (μ_{Y_i}, σ_{Y_i}) $T(Y_p - Y_i)$	1	β_1	β_2	-215.0	26.4
	2	$\beta_1 + \beta_2 Y_i$	β_3	-161.0	22.8
	3	$\beta_1 + \beta_2 Y_i$	$\beta_3 + \beta_4 Y_i$	-52.3	14.0
	4	$\beta_1 + \beta_2 P_i + \beta_3 A_i + \beta_4 Y_i$	$\beta_5 + \beta_6 Y_i$	0	0
	5	$\beta_1 + \beta_2 P_i + \beta_3 A_i + \beta_4 Y_i$	$\beta_5 + \beta_6 P_i + \beta_7 A_i + \beta_8 Y_i$	-1.5	1.9
Weibull (k_{Y_i}, λ_{Y_i}) $T(Y_p - Y_i)$		β_1	β_2	-116.2	21.4
$k_{Y_i} = \left(\frac{\sigma_{Y_i}}{\mu_{Y_i}}\right)^{-1.086}$ $\lambda_{Y_i} = \frac{\mu_{Y_i}}{\Gamma(1 + \frac{1}{k})}$					
	7	$\beta_1 + \beta_2 Y_i$	β_3	-90.9	22.4
	8	$\beta_1 + \beta_2 Y_i$	$\beta_3 + \beta_4 Y_i$	-79.5	19.3
	9	$\beta_1 + \beta_2 P_i + \beta_3 A_i + \beta_4 Y_i$	$\beta_5 + \beta_6 Y_i$	-72.8	19.6
	10	$\beta_1 + \beta_2 P_i + \beta_3 A_i + \beta_4 Y_i$	$\beta_5 + \beta_6 P_i + \beta_7 A_i + \beta_8 Y_i$	-21.6	12.7

FIGURE CAPTIONS

Figure 1 . Hazard functions (inset boxes) and the associated probability density distributions (main panels) where lag times follow a Weibull distribution (see Appendix S1). Parameter k determines the shape of the hazard function: **A** if $k = 1$, the hazard is constant over time and the lag time distribution is negative exponential; **B** if $1 < k < 2$ the hazard function is downward accelerating and the lag time distribution shifts to unimodal for all $k > 1$; **C** if $k = 2$ the hazard function is linear; **D** if $k > 2$ the hazard function is upward accelerating. Each panel shows three hazard functions and associated lag time distributions given the shape parameter k , with distributions having a mean lag time of approximately 100 (red lines), 200 (blue lines) or 300 (black lines) years. In each of **B-D**, a steeper hazard function results in the lag time distribution having a smaller mean and variance (see Appendix S1 for details).

Figure 2 . Lag time (the lag between introduction and naturalisation in years) plotted against year of first introduction for 708 plant species naturalised in Britain, with marginal histograms for lag time and year of first introduction (shaded grey). The lag times for each introduction year are truncated at the sloping line because we can only observe species that have naturalised up to the present (set to the year 2000). The two inset histograms (shaded blue) show the distribution of lag times for species introduced in the periods 1620-1640 and 1890-1910 (shown by the blue vertical lines). These lag time distributions are truncated at the dotted lines on the histograms. The black curves are normal distributions fitted to the truncated lag

time data with the area under the curve beyond the year of truncation (shaded red) estimating the invasion debt (the number of species introduced in each period that are yet to naturalise based on the truncated lag time distribution for species that have naturalised).

Figure 3 . A-C Parameter estimates for model 4 (the best-performing model, see Table 1). Filled circles are the means of the posterior distributions and vertical lines show the 95% credible intervals. **A** shows the estimated mean and standard deviation of the expected lag times for trees and shrubs introduced in 1500 AD; **B** shows the rate at which the mean and standard deviation of the lag time distribution changed per year after 1500 AD; **C** shows the mean lag times of perennial herbs and annual/biennial species relative to trees/shrubs, which were set as the reference class. All parameter estimates are on the \log_e scale. **D-F** Hazard functions by life-form for model 4 for the years 1650, 1750, 1850 and 1950, showing the increase in the hazard (a number proportional to the probability a species will naturalise in the next small time step given it has not yet naturalised) over time following introduction.

Figure 4 . The observed distribution of lag times (grey histograms) for: **A** all species and **B-F** species introduced in different centuries (starting in 1500 through to 1900). For each century, the observed lag times are truncated at a maximum value shown as dashed lines on each histogram. The black lines show the mean predicted distribution of lag times based on the actual dates of introduction and 10000 simulations of each species lag time derived from model 4 (see text), with the red shading showing the 95% quantiles from the simulations.

Figure 5 . Histograms with 20 year bins for year of introduction and year of naturalisation for 708 naturalised plant species in Britain grouped by life-form (**A, D** = trees/shrubs; **B, E** = perennial herbs; **C, F** = annual/biennial species). The most recent year of introduction for naturalised species in the data (1960) is shown as a dashed line in **A-C** , and the most recent year of recorded naturalisation (2000) is shown as a dashed line in **D-F** . The black curves in **D-F** show the expected number of species naturalising in each 20 year interval based on the actual dates of introduction (**A-C**) and the expected lag times estimated using model 4. For each life-form, the area under the curve beyond the year 2000 (shaded red in **D-F**) estimates the invasion debt for that life-form (the number of species introduced before 1960 that are yet to naturalise based on the truncated lag time distribution of the species that have naturalised). The black curves in **A, C** and **E** show the trend in the number of species introduced in each 20 year period that will naturalise in the foreseeable future (see Appendix S3 for details).

Figure 1 .

Hosted file

image1.emf available at <https://authorea.com/users/387007/articles/508449-time-lags-and-the-invasion-debt-in-plant-naturalisations>

Figure 2 .

Hosted file

image2.emf available at <https://authorea.com/users/387007/articles/508449-time-lags-and-the-invasion-debt-in-plant-naturalisations>

Figure 3 .

Hosted file

image3.emf available at <https://authorea.com/users/387007/articles/508449-time-lags-and-the-invasion-debt-in-plant-naturalisations>

Figure 4 .

Hosted file

image4.emf available at <https://authorea.com/users/387007/articles/508449-time-lags-and-the-invasion-debt-in-plant-naturalisations>

Figure 5 .

Hosted file

image5.emf available at <https://authorea.com/users/387007/articles/508449-time-lags-and-the-invasion-debt-in-plant-naturalisations>