

# Modeling unobserved variables in dendrochronological age structures improves inferences about population dynamics

Ezequiel Aráoz and Juan Manuel Morales

**Abstract:** The loss of evidence poses a major challenge to historical ecology. For example, dendroecological studies aiming at relating tree establishment with past climate should consider the possibility that many plants might not survive to be recorded at the date of the study. A standard approach to deal with this data loss consists in fitting an exponential decay curve to the observed age structure and using the residuals of this fit as a proxy of tree establishment. Here, we show that hierarchical Bayesian analysis (HBA), where tree establishment is modeled as a latent variable, can outperform the standard approach. We illustrate the use of HBA with a simulation study in which the goal is to infer population dynamics from dendrochronological age structures. Both methods are also used to analyze empirical data from expanding *Alnus acuminata* Kunth forests in northwestern Argentina. The simulation study showed that the standard approach underestimated the association between rainfall and tree establishment. The HBA was unbiased and had narrower uncertainty around estimates. In the empirical study, the HBA detected effects of rainfall on tree establishment, which were deemed not significant by the standard approach. Besides these advantages, the flexibility of the HBA should allow for the analysis of more complex (and realistic) models.

**Key words:** Bayes, coverage, dendrochronology, hierarchical analysis, uncertainty, unobserved variable.

**Résumé :** La perte de preuves pose un sérieux défi à l'écologie historique. Par exemple, les études dendroécologiques visant à relier l'établissement des arbres au climat antérieur devraient envisager la possibilité que plusieurs individus aient pu disparaître avant qu'on puisse en tenir compte lors d'une étude. Une approche standard pour tenir compte de cette perte de données consiste à ajuster une courbe de décroissance exponentielle à la structure d'âge observée et d'utiliser les résidus de cet ajustement comme approximation de l'établissement des arbres. Dans cette étude, nous montrons que l'analyse bayésienne hiérarchique (ABH), dans laquelle l'établissement des arbres est modélisé comme variable latente, peut surpasser l'approche standard. Nous illustrons l'utilisation de l'ABH à l'aide d'une étude de simulation ayant pour but de déduire la dynamique de populations à partir de structures d'âge dendrochronologiques. Les deux méthodes sont aussi utilisées pour analyser des données empiriques provenant d'aulnaies en expansion dans le nord-ouest de l'Argentine. L'étude de simulation a montré que l'approche standard sous-estimait la relation entre la pluie et l'établissement des arbres. L'ABH était non biaisée et associée à une marge d'incertitude plus étroite autour des estimations. Dans l'étude empirique, l'ABH a détecté des effets de la pluie sur l'établissement des arbres qui ont été jugés non significatifs par l'approche standard. Outre ces avantages, la flexibilité de l'ABH devrait permettre l'analyse de modèles plus complexes et réalistes. [Traduit par la Rédaction]

**Mots-clés :** Bayes, couverture, dendrochronologie, analyse hiérarchique, incertitude, variable non observée.

## 1. Introduction

Historical ecology aims at reconstructing and depicting the environmental history of particular ecosystems to understand their interactions with ecological agents and other factors in the past. Thus, this discipline is informative for conservation, restoration, and forecasting potential consequences of climate and land use changes. The most important constraint that historical reconstructions must deal with is the loss of evidence through time due to the natural degradation of the record (Egan and Howell 2001). This loss poses a challenge for any study of historical ecology because it reduces the reliability of reconstructions as time before present increases, and therefore, it must be taken into account to interpret data from historical records (Swetnam et al. 1999).

Dendrochronology is one of the most widespread methods used in historical ecology (Schweingruber 1988). Dendrochronology is used to reconstruct different aspects of past environments such as climate (Briffa 2000), disturbance regime (Veblen et al. 1994), and human activity (Haneca et al. 2009). Sometimes this technique is

used to describe the age structure of woody plants (e.g., Villalba and Veblen 1997; Miller and Halpern 1998; Carilla and Grau 2010). Other studies aim at assessing the environmental factors that controlled tree establishment in the past to understand the processes leading forest dynamics, which has significant implications in the management and forecasting of forests affected by environmental changes.

In many parts of the world, forests are shifting their extent due to forest transition (forests expand in areas where land use intensity is reduced (Rudel et al. 2005)) or due to global warming (forests move upwards or toward higher latitudes (Harsch et al. 2009)). Under both scenarios, the dynamics of forests are likely to be affected by climate change, so accurate forecasting based on historical evidence will be crucial for the implementation of adequate management policies (Johnson et al. 1994). For example, *Alnus acuminata* Kunth montane forests of northwestern Argentina encroached over adjacent grasslands during the 20th century (Grau 1985). An increase in regional rainfall (Minetti and Vargas 1998) has been proposed as one of the causes of this expansion of forests, but to our knowledge, no formal analysis has been performed.

Received 20 February 2015. Accepted 18 August 2015.

E. Aráoz. Instituto de Ecología Regional, Universidad Nacional de Tucumán, CC 34, 4107 Yerba Buena, Tucumán, Argentina.

J.M. Morales. Laboratorio Ecotono, INBIOMA-CONICET, Universidad Nacional del Comahue, Quintral 1250, (8400) Bariloche, Argentina.

**Corresponding author:** Ezequiel Aráoz (e-mail: ezequielaraoz@gmail.com).

The analysis of tree age structures generated through dendro-chronological methods can be helpful to understand the climate control on tree establishment. However, the effect of tree mortality (and the consequent loss of information) in these assessments is usually overlooked or handled deficiently. Different approaches have been proposed to deal with data loss in static age structure analyses (e.g., Miller and Halpern 1998; Mast et al. 1999; Grau and Veblen 2000; Wang et al. 2006; Danby and Hik 2007; Carilla and Grau 2010). Some studies do not take mortality into account and consider that the observed frequency in each age class is a good indicator of tree establishment in the corresponding time interval (e.g., Miller and Halpern 1998). These studies use correlation analysis between age-class frequency and meteorological records of the corresponding period of establishment to make inferences about climate control on tree establishment (e.g., Miller and Halpern 1998; Wang et al. 2006). Other studies combine data of living trees with that of dead logs to reconstruct past tree establishment (e.g., Mast et al. 1999). This approach overcomes some of the problems arising from mortality, but it cannot be applied in environments in which dead logs decay rapidly or when their dating may be inaccurate (Danby and Hik 2007). Another approach was used by Grau and Veblen (2000), in which observed frequencies were grouped in 5-year age classes for 10 age structures of trees and a principal component analysis was performed that assigned the first component to mortality (which was supported by a strong correlation with calendar year) and used the second and third components to evaluate the relationships with environmental factors. Alternatively, mortality rates from an independent source such as permanent plots (e.g., Carilla and Grau 2010) can be used to estimate tree establishment from observed age structures. In this approach, mortality can be accurately measured, but it does not take into account potential differences between populations. Moreover, permanent plots usually do not include all age classes, and they have a limited time span. As these methods are based on ad hoc assumptions, their use is not universal. Yet, there exists a “standard method” that is based on simple statistical analysis. This method fits a negative exponential curve to the age structure to estimate the expected number of trees established at time  $t$ , assuming constant mortality rates. Sometimes, instead of a negative exponential, a power function is fitted to model age-dependent mortality (Hett and Loucks 1976; Batllori and Gutiérrez 2008). The residuals of the curve are anomalies that represent an indicator of tree establishment. Therefore, climate control on tree establishment is inferred by regressing these residuals against meteorological records (e.g., Daniels and Veblen 2004; Gamache and Payette 2005; Danby and Hik 2007). This method implies two steps: the estimation of mortality and the parameterization of the association between establishments and climate. The proxy of establishment is based on a single estimate of mortality, so the error of the estimate of mortality is carried over to the next step, reducing the statistical power of the method.

In this study, we propose the use of hierarchical Bayesian analysis (HBA; Clark 2005) to explicitly deal with the different processes involved in population dynamics, using a latent variable to model the unobservable establishments in the past. To compare the proposed and the standard methods, we simulate forest dynamics with constant mortality and establishment controlled by climate. We estimated the parameters of the emergent data with HBA and with the standard method to assess their performance. We also used both methods to analyze empirical data from *A. acuminata* dominated forests from a subtropical montane system.

## 2. Material and methods

### 2.1. A model for tree recruitment

In forest–grassland interfaces, the seedling stage is considered to be the bottleneck for tree establishment, so that environmental conditions during this stage constrain recruitment (Sankaran

et al. 2004). For simplicity, we assumed that water availability in year “ $t$ ” is the only restriction to tree establishment ( $N_t$ ) and that the association between rainfall ( $R_t$ ) and tree establishment is linear (e.g., Daniels and Veblen 2004). The number of seedlings established at time  $t$  is modeled as a latent (unobserved) variable, because usually, it cannot be observed directly as establishments occurred in the past. The stochasticity in the number of establishments was modeled with a Poisson distribution, with intensity ( $\lambda_t$ ) being a linear function of rainfall

$$(1) \quad \begin{aligned} N_t &\sim \text{Poisson}(\lambda_t) \\ \lambda_t &= \max(a + bR_t, 0) \end{aligned}$$

where  $a$  is the expected number of establishments when rainfall is zero, and  $b$  is the increase in this expectation when there is a one unit raise in annual rainfall. Clearly, other formulations for this relationship are possible; a log-linear model could be used to naturally constrain  $\lambda$  to be positive or the relationship between establishments and rainfall could be modeled with some nonlinear function. We used a linear model to make the comparison with the standard approach straightforward, and eq. 1 was truncated at zero to avoid negative  $\lambda$  values.

Due to mortality, only some (if any) of the trees ( $s_c$ ) that were established at time  $t$  would be present at the time of observations ( $t_{\text{obs}}$ ), which we modeled with a Binomial distribution

$$(2) \quad \begin{aligned} O_c &\sim \text{Binomial}(N_t, s_c) \\ s_c &= \exp(-d(t_{\text{obs}} - t)) \end{aligned}$$

Here,  $O_c$  is the observed number of trees belonging to age class  $c$ . Age structure is a snapshot of a population at the sampling time ( $t_{\text{obs}}$ ). Yet, inferences are referred to the establishment time  $t$ . The number of trees in each age class ( $O_c$ ) is a fraction of those that recruited ( $N_t$ ). The individual-level probability of surviving  $t_{\text{obs}} - t$  years was modeled with an exponential function that decreases with constant mortality rate  $d$ . Assuming constant mortality makes it easier to compare the proposed approach with the standard method (Hett and Loucks 1976).

### 2.2. Simulated data

We used simulated age structures from tree populations in which mortality rates and the relationship with rainfall were known to assess the performance of HBA and the standard method in recovering these parameters. The estimates from both methods were compared with the “true values” to assess their power, including bias, precision, and coverage (Bolker 2008). Bias, the mean difference between the estimate and the parameter, was statistically assessed through a  $t$  test. The precision of the estimate was assessed through the confidence and credible interval widths, which are indicators of the uncertainty of the estimates. Coverage was calculated as the proportion of analyses in which the true parameter value is included within the estimated confidence or credible intervals.

Test data were generated by simulating rainfall, establishments, and mortality over a period of 1000 years, which was long enough to deplete the oldest age classes under any of the considered mortality rates. For every year, rainfall was sampled from a normal distribution with a mean of 150 cm and standard deviation of 10 cm, which constitute an index of rainfall variability and were determined arbitrarily.

$$R_t \sim \text{Normal}(150, 10) \quad \text{for } t = 1, \dots, T$$

Establishments followed eq. 1, where parameter  $b$  took values of  $-3$ ,  $3$ , and  $10$ , and parameter  $a$  was set to 1000. Surviving trees at age class  $c$  were simulated using eq. 2, and  $d$  took the values of

0.001, 0.005, 0.01, and 0.05 year<sup>-1</sup>. Age structures were generated from a sample of 250 trees, which is within the limits of many dendrochronological age structure reconstructions (e.g., Daniels and Veblen 2003). At least 100 age structures for every combination of parameters  $b$  and  $d$  were simulated; however, about 12% were discarded because the Markov chain Monte Carlo of the HBA failed to converge. The standard methods' estimates of the discarded analysis were farther from parameters than the mean of the remaining estimates.

### 2.3. Comparison between methods

In the standard method, a negative exponential curve is fitted to the age structure (Hett and Loucks 1976). Point estimates of mortality ( $d$ ) and their 95% confidence intervals were recorded directly; however, the estimates of the rainfall coefficients ( $b$ ) and their 95% confidence intervals were scaled to the whole population by multiplying them by the ratio between sample size and population size.

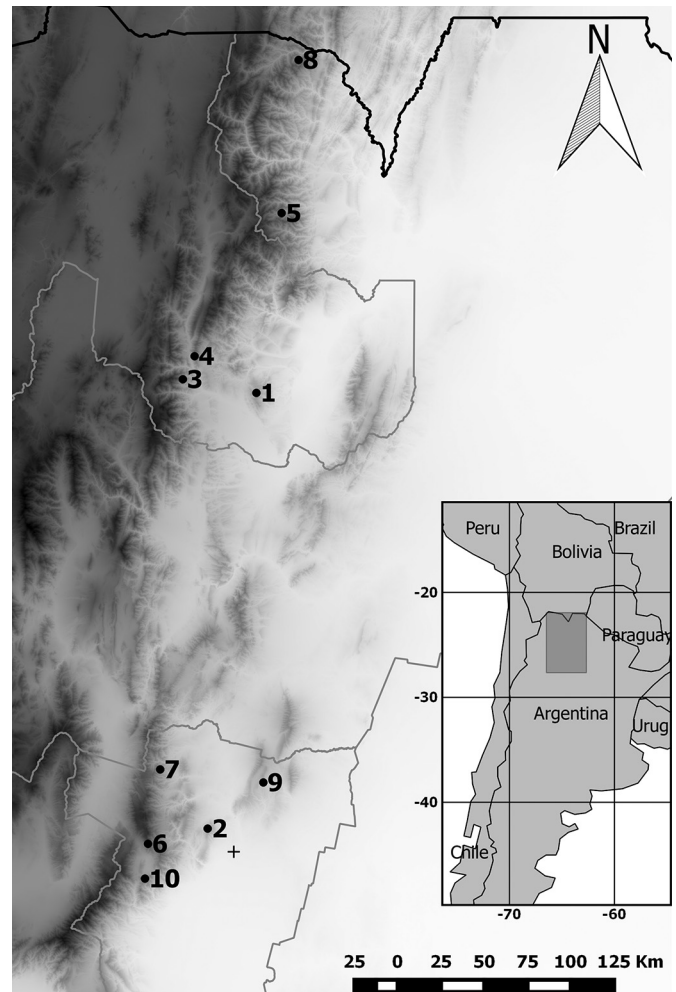
We also used a HBA that estimates all parameters simultaneously. The HBA model was implemented in WinBUGS, a software designed to perform Bayesian inference via Markov Chain Monte Carlo simulations (Spiegelhalter 2003). Flat, uninformative priors for all parameters were used so that data dominated the results (see bug model in the Supplementary material<sup>1</sup>). For all parameters, three Markov chains were started at random locations and updated 400 000 times; the convergence of the chains was evaluated through the "potential scale reduction factor"  $\hat{R}$  (Gelman and Rubin 1992). The joint posterior distributions of the parameters resulted from combining the independent chains after discarding the first 200 000 steps and keeping every third observation to reduce autocorrelation. The mean of the combined chains was used as a point estimate for parameters, and the 95% highest posterior density interval was used to estimate credible intervals. As with the standard method, the estimate of  $b$  and its credible interval were scaled to the whole population.

### 2.4. An example with *A. acuminata* field data

We compared the standard and proposed approaches analyzing dendrochronological data sets from *A. acuminata* forests from northwestern Argentina. *Alnus acuminata* is a pioneer species distributed in mountain ranges from Mexico to northwestern Argentina. In northwestern Argentina, the tree line is dominated by this species, so it is expected that the establishment is controlled to some extent by climate (Holtmeier and Broll 2005). In xeric environments, tree establishment might be enhanced in wet years (e.g., Daniels and Veblen 2004). In contrast, in environments with higher water availability, temperature might be the main constraint to tree establishment; rainy periods are cloudy and cool, and thus, establishment could be associated with dryer (and warmer) years. For example, in montane systems of northwestern Argentina, a negative association between rainfall and temperature has been observed. In this system, tree growth of different species is controlled by rainfall or temperature alternatively depending on the water availability of the site (Villalba et al. 1998). Accurately assessing the effect of rainfall on tree establishment would have implications in forest forecasting and management.

From 1997 to 2007, the age structures of *A. acuminata* forests from 10 watersheds of northwestern Argentina were constructed using dendrochronology (Fig. 1). The selected watersheds were characterized by a mosaic of patches covered by grasslands, shrublands, or *A. acuminata* dominated forests. In each watershed, between 150 and 700 boles of *A. acuminata* were sampled with increment borers to date tree establishment. Grau and Veblen (2000) estimated that this species grows at least 18 cm in the first year. Thus, to obtain the innermost ring, trees were cored within

**Fig. 1.** Map of the position of the 10 watersheds analyzed. The cross indicates the position of the meteorological station of San Miguel de Tucumán. 1, El Centinela; 2, Taficillo; 3, Yala; 4, Volcán; 5, San Andrés; 6, La Quebradita; 7, Hualinchay; 8, Cerro Bravo; 9, Medina; 10, Tafi Sur.



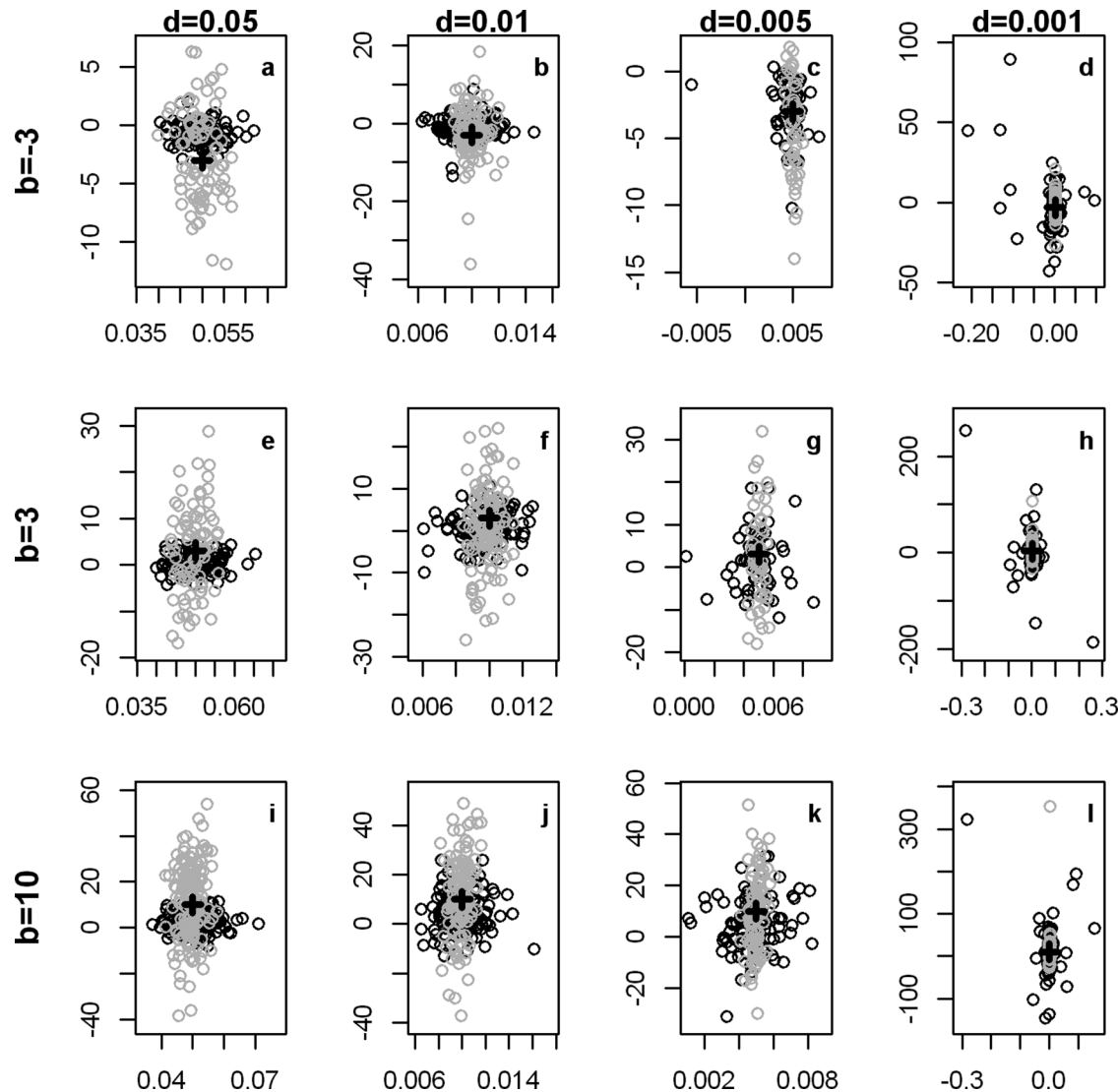
18 cm of the soil from the downhill side of the tree to reduce problems in the manipulation of the borer handle. After counting backwards from the outermost to the innermost ring and correcting for nonattained centers (Duncan 1989), we reconstructed the age structures of the forests of each watershed.

Rainfall records are scarce in the study area and they are restricted to lowland areas, whereas *A. acuminata* dominates montane forests. In this exercise, we used a single rainfall series from San Miguel de Tucumán (26°48'S, 65°12'W), which is one of the longest (it begins in 1884) and most reliable records in the region. Tree establishments occur during rainy springtime (from October to December), and seedlings must harden before the beginning of winter. We rescaled rainfall records between October of the establishment year ( $Y_0$ ) and March of the following year ( $Y_1$ ) to have a mean of zero and a standard deviation of one and used them as an indicator of the climate conditions. Although we acknowledge that temperature might affect the water balance of the environment, we consider that extrapolation from lowlands to highlands could increase the noise of the proxy. Analyses were restricted to

<sup>1</sup>Supplementary data are available with the article through the journal Web site at <http://nrcresearchpress.com/doi/suppl/10.1139/cjfr-2015-0073>.



**Fig. 2.** Point estimates of the simulated tree population parameters. Point estimates of rainfall coefficient  $b$  and mortality rate  $d$  under the standard (black circles) and the proposed approach (hierarchical Bayesian analysis (HBA), gray circles). Crosses represent the combination of parameters used in simulations. Figure parts: (a), (e), and (i),  $d = 0.05$ ; (b), (f), and (j),  $d = 0.01$ ; (c), (g), and (k),  $d = 0.005$ ; (d), (h), and (l),  $d = 0.001$ .



the 20th century when more than 99% of the sampled trees were established.

### 3. Results

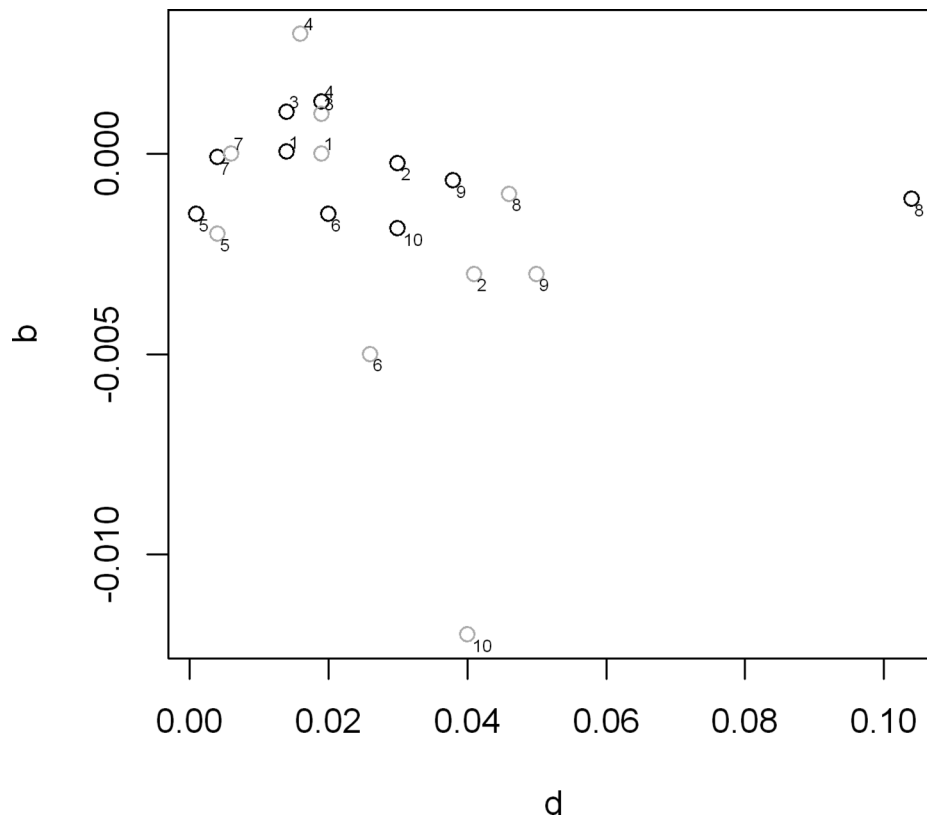
Under high simulated mortality ( $d \geq 0.01$ ), both the standard method and HBA recovered similar estimates (Figs. 2a, 2b, 2e, 2f, 2i, and 2j; Supplementary Table S1<sup>1</sup>). With low mortality ( $d \leq 0.005$ ), the standard approach estimated  $d$  with some bias, even estimating negative mortalities when mortality was set to 0.001 (Figs. 2c, 2d, 2g, 2h, 2k, and 2l; Supplementary Table S1<sup>1</sup>). The  $d$  estimates of HBA were always more precise than those of the standard method; the standard deviation of the mortality estimate obtained with the standard method was about twice that of the mortality estimate obtained through HBA when  $d \geq 0.01$  (Figs. 2a, 2b, 2e, 2f, 2i, and 2j; Supplementary Table S2<sup>1</sup>). The difference increased to four times with low mortality ( $d = 0.005$ ; Figs. 2c, 2g, and 2k), and this difference was two orders of magnitude higher when parameter  $d$  was 0.001 (Figs. 2d, 2h, and 2l).

The 95% confidence intervals of the estimated mortality from the standard method and the 95% credible intervals of the HBA included the true  $d$  parameter value 92.5% and 95% of the time,

respectively, despite the fact that the confidence intervals of the first approach were wider than the credible intervals of the HBA (Supplementary Tables S3 and S4<sup>1</sup>). Confidence intervals were two orders of magnitude broader than the corresponding credible intervals under very low mortality rates ( $d = 0.001$ ), and this difference decreased to between 3 and 1.3 times when mortality was higher ( $d \geq 0.005$ ).

On average, the rainfall coefficient ( $b$ ) was more precisely estimated by HBA than by the standard approach, implying that HBA estimated this coefficient with no bias in contrast with the standard method. The difference in the performance of both methods increased with the magnitude of the parameter, and in general, the standard approach underestimated its magnitude (Figs. 2a, 2b, 2e, 2f, 2i, and 2j; Supplementary Table S5<sup>1</sup>). Bias in  $b$  estimates obtained with the standard approach was positively associated with mortality rates (Fig. 2; Supplementary Table S6<sup>1</sup>), being significant ( $p < 0.001$ ) whenever mortality was high ( $d \geq 0.01$ ; Appendix A; note that the R script is also provided as a file in the Supplementary material<sup>1</sup>). Under very low mortality rates ( $d = 0.001$ ), the standard approach estimated  $b$  almost as well as HBA, which correctly estimated it under all combinations of parameters.

**Fig. 3.** Point estimates of real tree population parameters. Point estimates of rainfall coefficient  $b$  and mortality rate  $d$  under the standard (black circles) and proposed approach (hierarchical Bayesian analysis (HBA), gray circles) for *A. acuminata* populations of 10 watersheds from northwestern Argentina. 1, El Centinela; 2, Taficillo; 3, Yala; 4, Volcán; 5, San Andrés; 6, La Quebradita; 7, Hualinchay; 8, Cerro Bravo; 9, Medina; 10, Tafi Sur.



The parameter for the effect of rainfall on tree establishment ( $b$ ) was contained within the 95% confidence and credible intervals of estimates (i.e., coverage) 78% of the time with the standard approach and 93.5% of the time with HBA (Supplementary Table S7<sup>1</sup>). This coverage was above 90% with the standard method when mortality was low or very low ( $d \leq 0.005$ ), it decreased to around 85% under high mortality ( $d = 0.01$ ), and with very high mortality ( $d = 0.05$ ), coverage decreased again to just 48%. Meanwhile, the coverage of parameter  $b$  with HBA remained above 90% under all combinations of parameters. The width of the 95% credible intervals of the rainfall coefficient estimated through HBA was, on average, similar to the 95% confidence interval of the standard approach (Supplementary Table S8<sup>1</sup>). The analyses of coverage show that the standard approach overestimates the precision of  $b$  parameter, whereas HBA is more realistic.

The analysis of *A. acuminata* populations with both approaches produced different results. The mortality estimate was, in general, higher with HBA than with the standard approach. However, in Cerro Bravo, the standard method estimated a high mortality, which was above 0.1, far from the rest of the populations that were between 0.001 and 0.04. Mortalities estimated with the proposed method ranged from 0.004 to 0.05 (Fig. 3). In the five watersheds with highest mortality rates, the 95% credible interval of HBA and 95% confidence interval of standard method did not include the point estimate of the other method (Supplementary Table S9<sup>1</sup>).

Although both methods estimated negative rainfall coefficients ( $b$ ) for six out of 10 sheds, some important differences were observed (Fig. 3). None of the  $b$  estimates of the standard approach were significant, as all of the 95% confidence intervals included zero. In contrast, HBA identified two populations in which the 95% credible interval of the estimate was completely below zero.

This method also identified a population (Volcán, one of the most xeric environments) with a positive  $b$  estimate that did not include zero in its 95% credible interval (Supplementary Table S9<sup>1</sup>).

#### 4. Discussion

We used a simple model to highlight the advantages of modeling unobserved variables instead of using a single surrogate of past events to estimate population parameters in historical ecology. A simulation experiment showed that the estimates of mortality and rainfall coefficients of forest dynamics attained through HBA were more precise and accurate than those obtained through the standard approach. Although in some scenarios the average estimated mortality ( $d$ ) was similar with both approaches, estimates were more accurate with HBA under a broad set of combinations of parameters. This apparent contradiction is explained by the high dispersion of  $d$  estimates among standard analyses in comparison with HBA, suggesting that the former is unstable. Moreover, having similar coverage, the average 95% confidence interval of the standard method was broader than the corresponding credible interval of the HBA, showing that this estimate was more precise (Supplementary Tables S3 and S4<sup>1</sup>).

Historical reconstructions are expensive; replicates are not always available, and usually inferences are based on single data sets, so analyses should take full advantage of reliable methods, especially when they are used to inform environmental policies. The estimation of mortality in *A. acuminata* populations illustrates this point. In five out of 10 watersheds, neither the 95% credible interval of HBA nor the 95% confidence interval of the standard method included the point estimate of the other method (Supplementary Table S9<sup>1</sup>). The high dispersion of mortality estimates

obtained through the standard method in the simulation analysis (Fig. 2; Supplementary Table S4<sup>1</sup>) suggests that HBA is more reliable. Thus, the high mortality rate estimated for Cerro Bravo by the standard method could be an effect of this instability. Historical ecology has been proposed as the best tool to depict environmental baselines for landscape management (Egan and Howell 2001), so errors in the estimates of sensitive parameters such as mortality (Pinero et al. 1984) could mislead the interpretation of the processes that shape landscapes. In contrast with the standard method, which relies on a single estimate of mortality to infer past events progressively compounding errors, HBA models the establishment process and mortality simultaneously. This is useful to accurately characterize the age structure of the population and to take into account different sources of uncertainty when estimating environmental effects on tree establishment, which are overlooked by the standard approach. For example, in the standard method, the rainfall coefficient  $b$  is estimated given a point estimate of mortality  $d$ , so all the uncertainty is assigned to  $b$ . In the simulation experiment, the  $b$  estimate of the standard method was biased, and its confidence intervals were inaccurate (Supplementary Tables S5 and S6<sup>1</sup>). In contrast, HBA estimated realistic, although wider, credible intervals. Differences were also observed when empirical data were analyzed; mortality estimates for the studied forests ranged between 0.001 and 0.1 with the standard method and between 0.004 and 0.05 with the HBA approach. The HBA detected important effects of rainfall on *A. acuminata* tree establishments in three out of 10 watersheds, which were deemed not significant by the standard approach. These differences may have implications in the understanding and forecasting of forest dynamics.

The shift in tree lines in response to environmental change is a worldwide issue whose mechanisms have not been totally elucidated (Trant and Hermanutz 2014). Accelerating global changes will continue to affect environmental patterns, so an accurate knowledge of the mechanisms that control ecosystems will be critical to understand, predict, and manage them. Climate warming has been proposed as the main cause of tree-line advance (Harsch et al. 2009), but other causes such as interaction with rainfall were also reported (Daniels and Veblen 2004; Wang et al. 2006). The HBA identified associations between rainfall and *A. acuminata* establishments, which were undetected by the standard approach (Supplementary Table S9<sup>1</sup>). This is a plausible association as tree establishment on grasslands depends on environmental constraints (Sankaran et al. 2004). Differences in the sign of the associations could be explained by variations in rainfall regimes; global warming pushes forests upwards only when water is available (Holtmeier and Broll 2005). Rainy years are cooler, so in environments where water is available, a negative association between rainfall and establishments is expected. In contrast, in more xeric environments, establishment is mainly constrained by water availability. The simulation experiment evidenced that the  $b$  estimate of the standard method was biased mainly under high mortality rates ( $d \geq 0.01$ ; Fig. 2; Supplementary Table S6<sup>1</sup>), which are within the ranges observed in *A. acuminata* populations, suggesting that estimated differences between both methods are not spurious.

Although we used a simplified model to facilitate the comparison of the performance of both methods, we acknowledge that forest dynamics is not so simple. Aráoz and Grau (2010) observed that, at the scale of stands, tree establishment was facilitated by fire events, which in turn were controlled by climate. These complex interactions are beyond the scope of the present study, which addresses forest dynamics at the watershed scale. However, further studies addressing these interactions at multiple scales are necessary to understand and forecast forest dynamics. The HBA framework is specially designed to deal with complex problems in which different hierarchical structures (e.g., in space or time), hidden variables, processes, and their uncertainties can be con-

sidered to estimate the parameters of the model (Clark 2005). Bayesian methods provide adequate tools to combine different sources of information, which is useful in historical ecology where information is scarce. One of the main advantages of Bayesian analysis is the straightforward formulation of the models and the possibility of estimating unobservable processes such as those involved in historical ecology (e.g., Boreux et al. 2009).

Here, tree populations were analyzed separately to compare the performance of both methods, but the empirical data could have been analyzed within a hierarchical framework, in which the regional climatic effect could be separated from the local watershed effect. The implementation of this framework would be straightforward by adding a watershed index and hyperparameters in the HBA, but it would be quite difficult in the standard approach (Gelman and Hill 2007). In this study, we illustrated the use of HBA to infer tree population dynamics from dendrochronological data, but the logic of this approach could be applied to infer processes from data obtained through other techniques of historical ecology, which are generally affected by loss of evidence or by other sources of uncertainty. For example, this approach could enhance reconstructions of disturbance regimes in which the evidence is erased by vegetation succession and by new disturbance events. HBA could also shed some light on palynological studies in which the contribution of each species to the record and its uncertainty could be modeled instead of using a single proxy.

We presented a Bayesian approach to model hard-to-get historic ecological data, which outperformed the current, best available method in recovering known parameters. An empirical data set illustrated the power of HBA, which detected associations between rainfall and tree establishment in expanding subtropical montane forests, which were not detected through the standard method. The use of latent variables to model unobserved processes in combination with an enhanced flexibility allowed by Bayesian analysis will constitute an essential tool to depict historic environments more accurately, taking advantage of different sources of information.

## Acknowledgements

We thank Ricardo Grau for providing dendrochronological age structures from three watersheds and for suggestions that improved the manuscript. Valuable suggestions and comments of two anonymous reviewers were also useful to improve a previous version of the manuscript. E. Aráoz was awarded with a postdoctoral fellowship by Fundación Bunge y Born. This work was supported by research projects from Agencia Nacional de Promoción científica y Técnica (PICT 00123) and CONICET (PIP 0968).

## References

- Aráoz, E., and Grau, H.R. 2010. Fire-mediated forest encroachment in response to climatic and land-use change in subtropical Andean treelines. *Ecosystems*, 13(7): 992–1005. doi:10.1007/s10021-010-9369-7.
- Batlloori, E., and Gutiérrez, E. 2008. Regional tree line dynamics in response to global change in the Pyrenees. *J. Ecol.* 96(6): 1275–1288. doi:10.1111/j.1365-2745.2008.01429.x.
- Bolker, B.M. 2008. *Ecological models and data* in R. Princeton University Press.
- Boreux, J.-J., Naveau, P., Guin, O., Perreault, L., and Bernier, J. 2009. Extracting a common high frequency signal from northern Quebec black spruce tree-rings with a Bayesian hierarchical model. *Clim. Past*, 5(4): 607–613. doi:10.5194/cp-5-607-2009.
- Briffa, K.R. 2000. Annual climate variability in the Holocene: interpreting the message of ancient trees. *Quat. Sci. Rev.* 19(1): 87–105. doi:10.1016/S0277-3791(99)00056-6.
- Carilla, J., and Grau, H.R. 2010. 150 years of tree establishment, land use and climate change in montane grasslands, northwest Argentina. *Biotropica*, 42(1): 49–58. doi:10.1111/j.1744-7429.2009.00655.x.
- Clark, J.S. 2005. Why environmental scientists are becoming Bayesians. *Ecol. Lett.* 8(1): 2–14. doi:10.1111/j.1461-0248.2004.00702.x.
- Danby, R.K., and Hik, D.S. 2007. Variability, contingency and rapid change in recent subarctic alpine tree line dynamics. *J. Ecol.* 95(2): 352–363. doi:10.1111/j.1365-2745.2006.01200.x.
- Daniels, L.D., and Veblen, T.T. 2003. Regional and local effects of disturbance



- and climate on altitudinal treelines in northern Patagonia. *J. Veg. Sci.* **14**(5): 733–742. doi:10.1111/j.1654-1103.2003.tb02205.x.
- Daniels, L.D., and Veblen, T.T. 2004. Spatiotemporal influences of climate on altitudinal treeline in northern Patagonia. *Ecology*, **85**(5): 1284–1296. doi:10.1890/03-0092.
- Duncan, R. 1989. An evaluation of errors in tree age estimates based on increment cores in kahikatea (*Dacrydium dacrydioides*). *N.Z. Nat. Sci.* **16**(1): 31–37.
- Egan, D., and Howell, E.A. 2001. The historical ecology handbook: a restorationists guide to reference ecosystems. Island Press.
- Gamache, I., and Payette, S. 2005. Latitudinal response of subarctic tree lines to recent climate change in eastern Canada. *J. Biogeogr.* **32**(5): 849–862. doi:10.1111/j.1365-2699.2004.01182.x.
- Gelman, A., and Hill, J. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York.
- Gelman, A., and Rubin, D.B. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**(4): 457–472. doi:10.1214/ss/1177011136.
- Grau, A. 1985. La expansión del aliso del cerro (*Alnus acuminata* HBK subsp. *acuminata*) en el noroeste de Argentina. *Lilloa*, **36**(2): 237–247.
- Grau, H.R., and Veblen, T.T. 2000. Rainfall variability, fire and vegetation dynamics in neotropical montane ecosystems in northwestern Argentina. *J. Biogeogr.* **27**(5): 1107–1121. doi:10.1046/j.1365-2699.2000.00488.x.
- Haneca, K., Čufar, K., and Beeckman, H. 2009. Oaks, tree-rings and wooden cultural heritage: a review of the main characteristics and applications of oak dendrochronology in Europe. *J. Archeol. Sci.* **36**(1): 1–11. doi:10.1016/j.jas.2008.07.005.
- Harsch, M.A., Hulme, P.E., McGlone, M.S., and Duncan, R.P. 2009. Are treelines advancing? A global metaanalysis of treeline response to climate warming. *Ecol. Lett.* **12**(10): 1040–1049. doi:10.1111/j.1461-0248.2009.01355.x.
- Hett, J.M., and Loucks, O.L. 1976. Age structure models of balsam fir and eastern hemlock. *J. Ecol.* **64**(3): 1029–1044. doi:10.2307/2258822.
- Holtmeier, F.-K., and Broll, G. 2005. Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental change at landscape and local scales. *Glob. Ecol. Biogeogr.* **14**(5): 395–410. doi:10.1111/j.1466-822X.2005.00168.x.
- Johnson, E., Miyanishi, K., and Kleb, H. 1994. The hazards of interpretation of static age structures as shown by stand reconstructions in a *Pinus contorta* – *Picea engelmannii* forest. *J. Ecol.* **82**(4): 923–931. doi:10.2307/2261455.
- Mast, J.N., Fule, P.Z., Moore, M.M., Covington, W.W., and Waltz, A.E. 1999. Restoration of presettlement age structure of an Arizona ponderosa pine forest. *Ecol. Appl.* **9**(1): 228–239. doi:10.1890/1051-0761(1999)009[0228:ROPASO]2.0.CO;2.
- Miller, E.A., and Halpern, C.B. 1998. Effects of environment and grazing disturbance on tree establishment in meadows of the central Cascade Range, Oregon, U.S.A. *J. Veg. Sci.* **9**(2): 265–282. doi:10.2307/3237126.
- Minetti, J.L., and Vargas, W.M. 1998. Trends and jumps in the annual precipitation in South America, south of the 15° S. *Atmósfera*, **11**(4): 205–221.
- Pinero, D., Martínez-Ramos, M., and Sarukhan, J. 1984. A population model of *Astrocaryum mexicanum* and a sensitivity analysis of its finite rate of increase. *J. Ecol.* **72**(3): 977–991. doi:10.2307/2259545.
- Rudel, T.K., Coomes, O.T., Moran, E., Achard, F., Angelsen, A., Xu, J., and Lambin, E. 2005. Forest transitions: towards a global understanding of land use change. *Glob. Environ. Change*, **15**(1): 23–31. doi:10.1016/j.gloenvcha.2004.11.001.
- Sankaran, M., Ratnam, J., and Hanan, N.P. 2004. Tree–grass coexistence in savannas revisited — insights from an examination of assumptions and mechanisms invoked in existing models. *Ecol. Lett.* **7**(6): 480–490. doi:10.1111/j.1461-0248.2004.00596.x.
- Schweingruber, F.H. 1988. Tree rings-basics and applications of dendrochronology. D. Reidel Publishing Company.
- Spiegelhalter, D. 2003. WinBUGS user manual version 1.4. Available from <http://www.mrc-bsu.cam.ac.uk/wp-content/uploads/manual14.pdf> [accessed December 2014].
- Swetnam, T.W., Allen, C.D., and Betancourt, J.L. 1999. Applied historical ecology: using the past to manage for the future. *Ecol. Appl.* **9**(4): 1189–1206. doi:10.1890/1051-0761(1999)009[1189:AHEUTP]2.0.CO;2.
- Trant, A.J., and Hermanutz, L. 2014. Advancing towards novel tree lines? A multispecies approach to recent tree line dynamics in subarctic alpine Labrador, northern Canada. *J. Biogeogr.* **41**(6): 1115–1125. doi:10.1111/jbi.12287.
- Veblen, T.T., Hadley, K.S., Nel, E.M., Kitzberger, T., Reid, M., and Villalba, R. 1994. Disturbance regime and disturbance interactions in a Rocky Mountain sub-alpine forest. *J. Ecol.* **82**(1): 125–135. doi:10.2307/2261392.
- Villalba, R., and Veblen, T.T. 1997. Regional patterns of tree population age structures in northern Patagonia: climatic and disturbance influences. *J. Ecol.* **85**(2): 113–124. doi:10.2307/2960643.
- Villalba, R., Grau, H.R., Boninsegna, J.A., Jacoby, G.C., and Ripalta, A. 1998. Tree-ring evidence for long-term precipitation changes in subtropical South America. *J. Climatol.* **18**(13): 1463–1478. doi:10.1002/(SICI)1097-0088(19981115)18:13<1463::AID-JOC324>3.0.CO;2-A.
- Wang, T., Zhang, Q.B., and Ma, K. 2006. Treeline dynamics in relation to climatic variability in the central Tianshan Mountains, northwestern China. *Glob. Ecol. Biogeogr.* **15**(4): 406–415. doi:10.1111/j.1466-822X.2006.00233.x.

## Appendix A

Note that this R script is provided as a file in the Supplementary material<sup>1</sup>.

# In this script we simulate population dynamics and sampling process. Afterwards population parameters are estimated from the emergent age structure through standard and HBA methods.

# We only show the case for  $d = 0.05$ . Analogous simulations were performed for other mortality rates.

library(emdbook); library(coda); library(R2WinBUGS)#Required packages are loaded

samplesize<-250 #Sample size is defined

b<-c(-3,3,10); d<- 0.05 #We set known parameters b (rainfall coefficient)and d (mortality rate)

n.simulations<-100 #We set the number of simulations for each b estimate

timelag<-999:0 # Time elapsed from establishment to sampling time

# We prepare the empty output vectors

n.sim<-n.simulations\*length(b)#n.sim is the total number of simulations taking into account every b parameter.

id<-numeric(n.sim); popsize<-numeric(n.sim); s.sample=numeric(n.sim)

# True parameters are to be recorded

d.true=numeric(n.sim); b.true<-numeric(n.sim)

# Estimates of the standard method

d.est1=numeric(n.sim); sd.d.est1<-numeric(n.sim)

b.est1=numeric(n.sim); sd.b.est1=numeric(n.sim)

a.est1<-numeric(n.sim); sd.a.est1<-numeric(n.sim)

# Estimates of HBA

d.est2=numeric(n.sim); sd.d.est2<-numeric(n.sim)# Point estimate of d through HBA

int1d<-numeric(n.sim); int2d<-numeric(n.sim)# Confidence interval of d estimate

d.hat<-numeric(n.sim)#r.hat for d estimate

b.est2=numeric(n.sim); sd.b.est2=numeric(n.sim) # Point estimate of b through HBA

int1b<-numeric(n.sim);int2b<-numeric(n.sim) # Confidence interval of b estimate

b.hat<-numeric(n.sim) #r.hat for b estimate

a.est2<-numeric(n.sim); sd.a.est2<-numeric(n.sim) # Point estimate of a through HBA

int1a<-numeric(n.sim); int2a<-numeric(n.sim) # Confidence interval of a estimate

a.hat<-numeric(n.sim) #r.hat for a estimate

for (j in 1:(length(b))) { #This loop is to replicate simulations for every b parameter

for (k in 1:n.simulations) {#This loop is to perform the defined simulations

i<-k+(j-1)\*n.simulations # i is the id of the simulation.

```

# Simulation of population dynamics.
Rainfall <-rnorm(n=1000, mean=150, sd=10) # Rainfall~Normal(150,10) for t = 1,...,T.
Establishments<-rpois(n=1000, lambda=1000+Rainfall*b[j])# Establishments~Lambda;(Lambda=a + b * rainfall)
Survival<-exp(-d*timelag)#Mortality is constant through age classes and through time.
Survivors<-rbinom(n=1000,size=Establishments,prob=Survival)#Survivors ~ Binomial (Establishmentt, exp(-d x (tobs-test))),
Population<-rep(timelag,Survivors)# We identify every individual by its establishment year
Sample<-sample(Population,size=samplesize)# We simulate a dendrochronological sampling of size=samplesize (250)
Agestructure<-as.vector(table(factor(Sample,levels=999:0)))# Individuals are grouped by age classes to get age structure.
# Empty output vectors are filled with results
popsize[i]<-length(Population)#We record the size of the population to rescale the estimated b parameter
id<-i; s.sample<-samplesize; d.true<-d; b.true<-b[j]
# In the standard approach some criterion must be used to "cut" the observed time series since oldest age classes (generally with zero
frequency) have negative residuals.
# We cut the age structure
# We define limits for HBA to speed up simulations
limit1<- max(which(cumsum (Agestructure)==min (cumsum(Agestructure))))+1
timelag1<-timelag[limit1:(length(timelag))]
agestructure1<-Agestructure[limit1:(length(timelag))]
rainfall1<-Rainfall[limit1:(length(timelag))]
# Parameters are estimated through standard method.
fit1<-try(nls(agestructure1~a*exp(d*timelag1), start=list(a=1,d=0.0051)),F)
proxy<-try(residuals(fit1),F)
residuals_model<-try(lm (proxy~rainfall1),F)
# parameters are estimated through HBA
nobs = length(agestructure1)
datasets <- list ("agestructure1", "rainfall1", "nobs", "timelag1")
inits <- function(){
list (a=runif(1,0,10),b=runif(1,0,10), d=runif(1,0.001,0.1))
}

parameters <- c("a", "b", "d")

simulation = try(bugs(data=datasets, inits, parameters, #
model.file = "modelobug.bug", n.thin=3,n.chains=3,#here the name of the bug model is provided.
bugs.directory="C:/Program Files/WinBUGS14",
n.iter=400000,debug=FALSE), F)

# Estimates of the standard method are recorded
d.est1[i]<-try(-(coef(fit1)[2]),F)
sd.d.est1<-try(summary(fit1)$coef[4],F)
b.est1<-try(coef(residuals_model)[2],F)
sd.b.est1<- try(coef(summary(residuals_model))[4],F)
a.est1<-try(coef(residuals_model)[1],F)
sd.a.est1<-try(coef(summary(residuals_model))[3],F)
# Estimates of the HBA are recorded
d.est2[i]<-try(simulation$mean$d,F)
sd.d.est2<-try(simulation$sd$d,F)
d.hat<-try(simulation$summary[31],F)
int1d<-try(simulation$summary[11],F)
int2d<-try(simulation$summary[27],F)
b.est2[i]<-try(simulation$mean$b,F)
sd.b.est2<-try(simulation$sd$b,F)
b.hat<-try(simulation$summary[30],F)
int1b<-try(simulation$summary[10],F)
int2b<-try(simulation$summary[26],F)
a.est2[i]<-try(simulation$mean$a,F)
sd.a.est2<- try(simulation$sd$a,F)
a.hat<-try(simulation$summary[29],F)
int1a<-try(simulation$summary[9],F)
int2a<-try(simulation$summary[25],F)
output<-cbind(id,popsize,s.sample,d.true,b.true,
d.est1,sd.d.est1,b.est1,sd.b.est1,a.est1,sd.a.est1,
d.est2,sd.d.est2,d.hat,int1d,int2d,b.est2,sd.b.est2,b.hat,int1b,int2b,a.est2,sd.a.est2,a.hat,int1a,int2a)
write(t(output),"path.txt",ncolumns=ncol(output))
}
}

```