Annals of Botany **00**: 1–17, 2026
https://doi.org/10.1093/aob/mcag175, available online at www.academic.oup.com/aob

ANNALS OF BOTANY
Nonprofit since 1887

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

ORIGINAL ARTICLE

Can long-term survival models improve inference in seed germination studies? Evidence from tropical forest species

Sebastião Martins Filho[1,*], Marciel Lelis Duarte[1,*], Lausanne Soraya de Almeida[2], Glauciana da Mata Ataíde[3] and Barbara França Dantas[4]

[1]Departamento de Estatística, Universidade Federal de Viçosa, Viçosa, Minas Gerais 36.570-900, Brazil, [2]Departamento de Engenharia Florestal, Universidade Federal de Viçosa, Viçosa, Minas Gerais 36.570-900, Brazil, [3]Departamento de Engenharia Florestal, Universidade Federal de São João Del Rei, Sete Lagoas, Minas Gerais 35701-970, Brazil and [4]Embrapa Semiárido, Petrolina, Pernambuco 56302-970, Brazil

[*]For correspondence. E-mail martinsfilho@ufv.br or marciel.duarte@ufv.br

Received: 29 May 2026 Accepted: 17 June 2026

• **Background and Aims** Seed germination analysis often ignores temporal dynamics and non-germinating fractions, limiting biological inference. Conventional methods assume all seeds germinate, which is unrealistic for heterogeneous forest seed lots. Long-term survival (cure fraction) models offer a framework to address this limitation. This study uses seed germination of native Brazilian forest species to demonstrate the applicability and advantages of long-term modelling.
• **Methods** Five independent datasets from native Brazilian forest species were analysed under varying environmental conditions, including temperature, osmotic stress, light and seed traits. Germination was modelled as a time-to-event process using parametric (log-normal, Weibull, log-logistic) and semiparametric (Cox) approaches with long-term survival models. Kaplan–Meier estimators were used for initial exploration and model comparison. Model selection was based on graphical fit and Cox–Snell residuals.
• **Key Results** All datasets showed a fraction of non-germinated seeds, with germination curves stabilizing below 100 %, which corroborates the use of cure models. Long-term models consistently outperformed traditional survival approaches, providing better fit and biological interpretability. Environmental factors differentially affected incidence and latency components: extreme temperatures and water stress significantly reduced the probability of germination, while having smaller effects on germination speed. The probability of germination varied widely between species and conditions, from 10 to 95 %. The models also revealed complex interactions between environmental and seed characteristics that influence germination dynamics.
• **Conclusions** This study demonstrates that long-term survival models provide a robust, biologically realistic statistical framework for analysing germination data from forest seeds. Partitioning the germination process into incidence (probability of germination) and latency (time to germination) components enabled key advances. First, it allowed precise quantification of the fractions of germinating and non-germinating seeds. Second, it provided more accurate and informative estimates than traditional survival methods. Finally, it enabled the isolation of the effects of environmental factors on germination capacity and speed.

Key words: Censored survival data, environmental stress, germination modelling, mixture cure models, seed trait, time-to-event analysis.

INTRODUCTION

In forest ecosystems, germination plays a fundamental role by directly influencing natural regeneration, population age structure and community resilience to environmental disturbances (Manso *et al.*, 2013; Hsu *et al.*, 2024). A comprehensive understanding of its temporal patterns is essential not only for the biology and ecology of species but also for practical applications such as ecological restoration, seedling production, *ex situ* conservation and forest management (Manso *et al.*, 2013). Consequently, germination is a key process for understanding species distribution, responses to environmental gradients and the impacts of climate change, providing support for conservation strategies (Carta *et al.*, 2022).

© The Author(s) 2026. Published by Oxford University Press on behalf of the Annals of Botany Company. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Despite its ecological relevance, seed germination analysis faces substantial methodological challenges. Several native tree species exhibit high variability in germinative response, often resulting from variations in dormancy, physiological quality and the environmental conditions to which they were exposed before, during and after harvest (Barak *et al.*, 2018). In such contexts, statistical strategies must rigorously dissociate the biological phenomenon from experimental bias while preserving the ecological realism of partial germination (Solarik *et al.*, 2016).

Traditionally, germination data analysis is based on summary indices, such as the final percentage and mean time (Mamani *et al.*, 2024), as well as measures of germination rate (McNair *et al.*, 2012; Romano and Stevanato, 2020) and synchronization (de Santana Santos *et al.*, 2026). Such methods reduce complex temporal processes to point statistics, thereby losing critical information about germination dynamics over time. Additionally, these metrics restrict formal statistical inference concerning treatment effects (Onofri *et al.*, 2022). Although non-linear models, such as the logistic, Weibull and Gompertz distributions, represent an advancement by describing the germinative trajectory, they lack mechanisms to handle censored data.

This limitation arises because, in the context of germination, censoring involves partial observation of the response, including seeds that rot or do not germinate by the end of the experimental evaluations. This information, often neglected, is fundamental and must be integrated into the modelling to avoid bias and obtain more precise conclusions (Onofri *et al.*, 2010; Mamani *et al.*, 2024). A statistical technique that incorporates such observations is survival analysis.

In this approach, germination is modelled as a time-to-event process (McNair *et al.*, 2012), allowing the use of non-parametric, semiparametric or parametric methods to test the influence of covariates (Harich *et al.*, 2016; Romano and Stevanato, 2020). Parametric and semiparametric models enable the estimation of parameters associated with germination speed and synchrony, while providing direct biological interpretations of these effects (Onofri *et al.*, 2010; Manhães Saint'Clair *et al.*, 2025). However, conventional survival modelling implicitly assumes that all individuals will experience the event over time. In practice, this assumption forces the asymptote of the germination curve towards 100 % (Onofri *et al.*, 2010), a premise often incompatible with the biology of native forest seeds, which rarely achieve such percentages. Seed lots constitute heterogeneous populations, composed of viable, dormant and empty seed fractions, many of which do not germinate within ecologically relevant time scales (Roesler *et al.*, 2023).

In this context, long-term survival models, also known as cure fraction models, offer a suitable statistical framework for analysing germination in the presence of such heterogeneity (Patilea and Van Keilegom, 2020). In long-term mixture models, the process is decomposed into two components: incidence, which estimates the proportion of seeds that will not germinate (the cure fraction) even if the monitoring period is long enough; and latency, which models the time to germination for the seeds that eventually do (Li and Taylor, 2002). This distinction between the two components allows distinguishing germination capacity (final percentage) from germination speed, respectively (Onofri *et al.*, 2011), aligning with fundamental ecophysiological concepts such as viability, dormancy and the effects of pre-harvest stress (Bernau *et al.*, 2020; Revilla *et al.*, 2024).

Despite being widely established in fields such as oncology and reliability engineering, long-term survival models remain under-explored in seed science, particularly for forest species (Roesler *et al.*, 2023). In this area, the literature lacks investigations that cover multiple species and experimental conditions, as well as comparative analyses across different probability distributions. Additionally, more explicit discussion of the biological interpretation of these model components is needed. Given these gaps, this study utilizes five germination datasets of native Brazilian forest species to demonstrate the applicability and advantages of long-term modelling. Specifically, we aim to (1) identify and quantify the fraction of non-germinated seeds (cure fraction); (2) evaluate the performance of different long-term models compared with traditional survival approaches; and (3) interpret the incidence and latency components from an ecophysiological perspective, providing a robust quantitative framework for analysing germination.

MATERIALS AND METHODS

Five independent datasets on seed germination of four native Brazilian forest species, selected for their ecological and silvicultural relevance, were analysed. The analysis employed various statistical techniques, using simple and complex experiments conducted under different conditions, aiming to demonstrate the applicability and robustness of long-term survival models.

Databases

Example 1: Germination of Anadenanthera colubrina var. cebil seeds under different temperatures. The first *A. colubrina* dataset was sourced from Gomes *et al.* (2023). A completely randomized design was employed, using four replicates of 25 seeds each. The seeds were sown on two sheets of Germitest® paper towels moistened with distilled water and incubated at six constant temperatures (5, 15, 20, 25, 35 and 45 °C) for 216 h (9 d). The process was conducted in a BOD (biochemical oxygen demand) germination chamber with a 12-h photoperiod. Germination was recorded once the radicle reached ≥1 mm in length.

Example 2: Germination of Anadenanthera colubrina var. cebil seeds under different osmotic potentials induced by polyethylene glycol 6000. The second *A. colubrina* dataset was also sourced from Gomes *et al.* (2023). A completely randomized design was employed, with four replicates of 25 seeds each. Six osmotic potentials (0, −0.2, −0.4, −0.6, −0.7 and −0.8 MPa) were induced using polyethylene glycol 6000 (PEG 6000) solutions. Seeds were sown on two sheets of Germitest® paper towel, moistened with either distilled water or PEG 6000 solutions at a ratio of 2.5 times the dry paper weight. The resulting paper rolls were maintained for 216 h (9 d) in a BOD germination chamber at 25 °C

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

and a 12-h photoperiod. Germination was defined by a radicle protrusion of ≥1 mm.

Example 3: Germination of Astronium urundeuva seeds under different osmotic potentials induced by NaCl solutions. The third dataset, focusing on *A. urundeuva*, was sourced from Oliveira *et al.* (2019). A completely randomized design was employed, using three replicates of 50 seeds each. Seven osmotic potentials (0, −0.072, −0.144, −0.216, −0.288, −0.360 and −0.432 MPa) were induced using sodium chloride (NaCl) solutions. Seeds were sown on two sheets of Germitest® paper towel, moistened with either distilled water or NaCl solutions at a ratio of 2.5 times the dry paper weight. The resulting paper rolls were maintained for 384 h (16 d) in a BOD germination chamber at 25 °C and a 12-h photoperiod. Germination was recorded based on a radicle protrusion of ≥1 mm.

Example 4: Germination of *Jacaranda caroba* seeds under different light and temperature conditions. The fourth dataset, concerning *J. caroba* seeds, was sourced from the study by Ordóñez-Parra *et al.* (2023). A completely randomized design was employed using a factorial arrangement, combining two alternating temperature regimes (30/20 °C and 25/15 °C) and two light conditions (12-h photoperiod and continuous darkness). Each of the four replicates consisted of a 9-cm Petri dish containing 25 seeds and two sheets of filter paper moistened with a 2 % nystatin solution as an antifungal agent. Petri dishes assigned to the dark treatment were wrapped in double-layered aluminium foil, and germination counts were conducted under a green safety light. All dishes were maintained in a BOD germination chamber and monitored every 24 h for 720 h (30 d). Radicle protrusion served as the criterion for germination.

Example 5: Germination of Mabea fistulifera seeds with different seed coat colours, elaiosome conditions, and substrate types. The fifth dataset, focusing on *M. fistulifera* seeds, was sourced from Santos *et al.* (2025). The experiment followed a completely randomized design with a $3 \times 2 \times 2$ factorial arrangement and four replicates of 25 seeds each. The treatments consisted of combinations of seed coat colour (red, orange and brown), elaiosome condition (present or absent) and substrate type (paper rolls and top-of-paper). For the paper rolls (PR) method, seeds were placed between two sheets of Germitest® paper, covered with a third sheet, rolled and placed in plastic bags. For the top-of-paper (TOP) method, seeds were placed on two sheets of Germitest® paper inside plastic Petri dishes. The substrate was sterilized and moistened with distilled water at a ratio of 2.5 times its dry weight. All samples were incubated in a BOD germination chamber at 25 °C for 576 h (24 d). Radicle protrusion served as the criterion for germination.

Statistical methods

Kaplan–Meier estimator. Initially, the data were evaluated using the Kaplan–Meier non-parametric estimator (Kaplan and Meier, 1958) to identify the presence of long-term events and to compare with the parametric and semiparametric methods employed. In this approach, the estimator quantifies the failure probability distribution within a population. The survival function is defined as:

$$S(t) = P(T > t) \tag{1}$$

where $S(t)$ is the probability of a seed not germinating by a specific time T, i.e. the probability of an observation surviving beyond time t.

Consequently, the cumulative distribution function is defined as:

$$F(t) = 1 - S(t) \tag{2}$$

where $F(t)$ is the probability that germination occurs by time t, and the germination curves were plotted based on this function.

The Kaplan–Meier estimator, used to calculate non-parametric estimates of the survival function, is defined as:

$$\hat{S}(t) = \prod_{j:t_j<t} \left(1 - \frac{d_j}{n_j}\right) \tag{3}$$

where $t_j < t = t_1 < t_2 \ldots < t_k$ are the k distinct and ordered germination times; d_j is the number of germinations at t_j, $j = 1, \ldots, k$; and n_j is the number of individuals at risk at t_j, i.e. the individuals that have neither germinated nor been censored up to the instant immediately prior to t_j.

Long-term parametric and semi-parametric survival models. The standard mixture model proposed by Boag (1949) and Berkson and Gage (1952) can be used to model seed germination. It allows simultaneous estimation of the probability of germination (incidence) and the time to germination (latency). Considering U as a Bernoulli random variable, where $U = 1$ if the individual is susceptible to the event, and $U = 0$ otherwise, such that $P(U = 1) = \pi$ and $P(U = 0) = 1 - \pi$, the formulation of this long-term mixture model is as follows:

$$S_{pop}(t|\boldsymbol{x}, z) = 1 - \pi(z) + \pi(z)\, S_U(t|\boldsymbol{x}) \tag{4}$$

where $S_{pop}(t)$ corresponds to the cumulative proportion of seeds that have not germinated after time t. The function $S_U(t)$ describes the probability of a seed not germinating, conditional on a vector of covariates $\boldsymbol{x} = (x_1, \ldots, x_p)'$, while π is the probability of a seed germinating given a vector of covariates $z = (z_1, \ldots, z_q)'$, which may include the same covariates present in $\boldsymbol{x}$. When $\pi = 1$, eqn (4) reduces to a standard survival model.

The function $S_U(t)$ can take the form of parametric or semi-parametric distributions. Among parametric approaches, the exponential, Weibull, log-normal and log-logistic distributions are commonly used to model survival data.

$$S_U(t) = \begin{cases} \exp\left[-\exp\left(\log t - \mu\right)\right], & \text{Exponential;} \\ \exp\left[-\exp\left(\dfrac{\log t - \mu}{\sigma}\right)\right], & \text{Weibull;} \\ 1 - \Phi\left(\dfrac{\log t - \mu}{\sigma}\right), & \text{Log-normal;} \\ \left[1 + \exp\left(\dfrac{\log t - \mu}{\sigma}\right)\right]^{-1}, & \text{Log-logistic;} \end{cases} \tag{5}$$

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

The inclusion of covariates is usually performed through the parameterization $\mu = \boldsymbol{\beta}'\boldsymbol{x}$, where $\boldsymbol{\beta}' = (\beta_0, \ \ldots, \ \beta_p)$ represents the vector of unknown regression parameters. In this context, the vector $\boldsymbol{x}$ acts multiplicatively on the scale parameter μ, accelerating or decelerating the seed germination time. The shape parameter σ represents germination synchrony. A lower value indicates that seeds germinate more synchronously (within a shorter time interval), which characterizes a seed lot of higher uniformity.

In semiparametric models, such as the Cox proportional hazards model (Cox, 1972), the function $S_U(t)$ is expressed as:

$$S_U(t|\boldsymbol{x}) = \exp\left(-\int_0^{t_i} \lambda_0(u)\exp(\boldsymbol{\beta}'\boldsymbol{x})du\right) = S_0(t)^{\exp(\boldsymbol{\beta}'\boldsymbol{x})} \quad (6)$$

where $\lambda_0(u)$ and $S_0(t)$ correspond, respectively, to the baseline hazard and survival functions, which are non-parametrically specified. More specifically, $S_0(t) = exp[-\Lambda_0(t)]$ where $\Lambda_0(t) = \int_0^{t_i} \lambda_0(u)\, du$ is the baseline cumulative hazard function.

The fundamental assumption of the Cox proportional hazards model is that the hazard ratios for the covariates remain constant over time, implying that the hazard curves do not intersect. Since this assumption was violated in the present study, we employed robust standard errors. This approach adjusts the variance–covariance matrix of the estimated coefficients, ensuring their validity even under the violation of proportionality (Lin and Wei, 1989).

Finally, the effect of the vector z on the probability $\pi(z)$ can be modelled through a binomial regression with a logit link function.

$$\pi(z) = \frac{\exp(\boldsymbol{\gamma}'z)}{1+\exp(\boldsymbol{\gamma}'z)} \quad (7)$$

where $\boldsymbol{\gamma}' = (\gamma_0, \ \ldots, \ \gamma_q)$ is the vector of unknown parameters, and the logit link function ensures that each probability $\pi(\mathbf{z})$ remains restricted to the interval [0, 1]. Other link functions, such as probit, complementary log-log (cloglog) and identity, may also be used.

The maximum likelihood (ML) method was employed to estimate the parameters of the long-term mixture model. These estimates were obtained by numerically maximizing the likelihood function:

$$\ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = log\prod_{i=1}^{n}\left[\pi(z_i)f_U(t_i|\boldsymbol{x}_i)\right]^{\delta_i}\left[1-\pi(z_i) + \pi(z_i)S_U(t_i|\boldsymbol{x}_i)\right]^{1-\delta_i} \quad (8)$$

where $\delta_i = 1$ if the event occurred and $\delta_i = 0$ otherwise, and f_U is the conditional probability density function of T.

In the parametric mixture model, maximum likelihood estimators (MLEs) are obtained using the Newton–Raphson method. However, in the semiparametric mixture model, the presence of a non-parametric component renders direct estimation of the function unfeasible. In this case, the EM algorithm provides an efficient alternative for separately estimating $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $S_0(t)$, through two independent components, ℓ_I and ℓ_S:

$$\ell_I(\boldsymbol{\gamma}|\mathbf{u}) = \log\prod_{i=1}^{n}\pi(z_i)^{u_i}\ [1-\pi(z_i)]^{1-u_i} \quad (9)$$

and

$$\ell_S(\boldsymbol{\beta}, \Lambda_0(t)|\mathbf{u}) = \log\prod_{i=1}^{n}\lambda_U(t|\mathbf{x_i})^{\delta_i u_i}\ S_U(t|\mathbf{x_i})^{u_i} \quad (10)$$

Statistical analysis, including parameter estimation and their respective standard errors, were calculated using the *mixcure* 2.0 package (Peng, 2020) in R software version 4.6.0 (R Core Team, 2026).

Germination curve determination. The obtained estimates were used to determine the survival function $S_{pop}(t)$. Based on the cumulative distribution function, defined by $F_{pop}(t) = 1 - S_{pop}$, the probabilities of seed germination over time were calculated. The germination curves were then plotted based on $F_{pop}(t)$.

Model selection and evaluation. The selection of the model that best describes the germination process was based on graphical inspection, comparing the model-fitted curve with the non-parametric method (Peng and Yu, 2021). Additionally, modified Cox–Snell residuals (Peng and Taylor, 2017) were analysed to verify the adequacy of the selected models.

RESULTS

In all datasets studied, the preliminary non-parametric Kaplan–Meier analysis revealed the presence of individuals immune to the event of interest, i.e. seeds that failed to germinate (Fig. 1). It was observed that the cumulative distribution function (germination curve) increases and reaches a plateau, but does not reach the value of 1 as $t \to \infty$. Additionally, late germination was observed, particularly under less favourable conditions.

Example 1

Based on a comparative evaluation of parametric models against Kaplan–Meier estimates, the log-normal model with long-term survivors was selected to describe *Anadenanthera colubrina* seed germination across different temperatures (Supplementary Data Fig. S1). This approach provided a superior fit to the traditional log-normal model, which overestimates the final germination fraction (Fig. 2). The goodness of fit was further confirmed by residual analysis (Supplementary Data Fig. S2).

The latency component showed that higher temperatures (15–45 °C) yielded negative coefficients, indicating that germination speed increases and, consequently, the time to event is reduced (Table 1). The intercept β_0 corresponds to germination at 5 °C, which was used as the reference level in the model. The shape parameter σ was 0.3940, characterizing a seed lot with greater uniformity.

The incidence component showed that the germination proportion π (calculated using parameters according to eqn (7)) varied significantly with temperature (Table 1). At 5 °C it was only 6.7 %, indicating a dormant/unviable seed fraction of 93.3 %. This percentage increased substantially

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

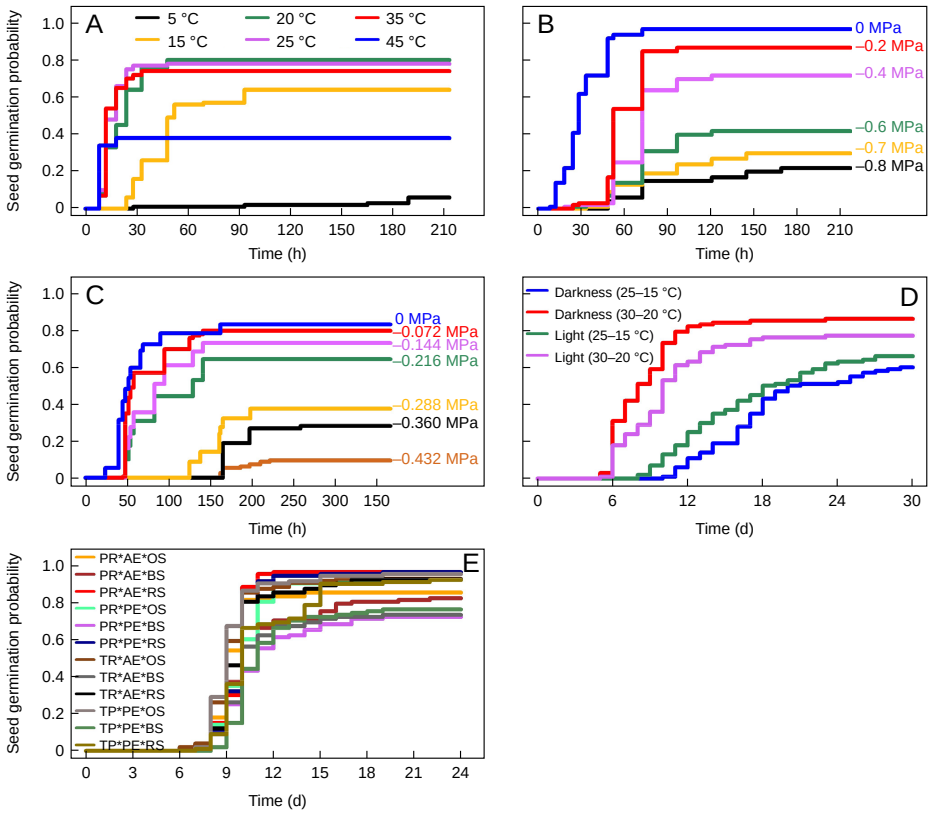

FIG. 1. Time-to-event analysis of seed germination estimated by Kaplan–Meier. Curves represent cumulative germination probabilities showing long-term survivors under different covariates. (A) Seed germination of *Anadenanthera colubrina* at different temperatures and (B) at different osmotic potentials induced by PEG 6000. (C) Seed germination of *Astronium urundeuva* at different osmotic potentials induced by NaCl. (D) Seed germination of *Jacaranda caroba* under different light and temperature conditions. (E) Seed germination of *Mabea fistulifera* across different substrates, seed colours and the presence or absence of elaiosomes.

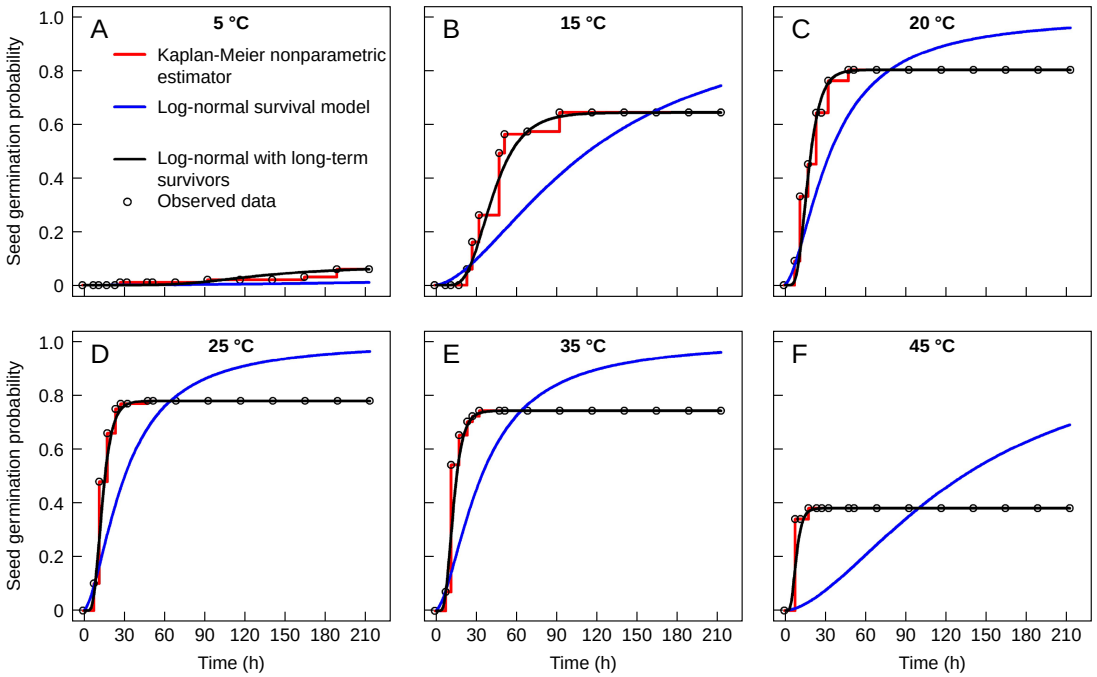

FIG. 2. Time-to-event analysis of *Anadenanthera colubrina* seed germination at different temperatures. Curves represent cumulative germination probabilities estimated by Kaplan–Meier and parametric log-normal models (with and without long-term survivors).

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

TABLE 1. Estimates, standard error, 95 % confidence interval and *P*-value of the log-normal model with long-term survivors for germination of *Anadenanthera colubrina* seeds under different temperatures.

Parameter	Estimate	Standard error	95 % confidence interval		*P*-value
			Lower	Upper	
$\beta_{0(\text{intercept})}$	4.8608	0.1914	4.4856	5.2359	< 0.0001
$\beta_{1(15\ ^\circ C)}$	−1.1105	0.1977	−1.4979	−0.7231	< 0.0001
$\beta_{2(20\ ^\circ C)}$	−1.9750	0.1964	−2.3599	−1.5900	< 0.0001
$\beta_{3(25\ ^\circ C)}$	−2.2148	0.1966	−2.6000	−1.8296	< 0.0001
$\beta_{4(35\ ^\circ C)}$	−2.2569	0.1968	−2.6426	−1.8711	< 0.0001
$\beta_{5(45\ ^\circ C)}$	−2.6959	0.2018	−3.0915	−2.3004	< 0.0001
σ	0.3940	–	–	–	–
$\gamma_{0(\text{intercept})}$	−2.6351	0.2793	−3.1825	−2.0877	0.0001
$\gamma_{1(15\ ^\circ C)}$	3.2106	0.4839	2.2636	4.1607	< 0.0001
$\gamma_{2(20\ ^\circ C)}$	4.0214	0.5033	3.0366	5.0095	< 0.0001
$\gamma_{3(25\ ^\circ C)}$	3.9008	0.4991	2.9242	4.8806	< 0.0001
$\gamma_{4(35\ ^\circ C)}$	3.6811	0.4927	2.7170	4.6484	< 0.0001
$\gamma_{5(45\ ^\circ C)}$	2.1456	0.4830	1.2006	3.0938	< 0.0001
$\pi_{(5\ ^\circ C)}$	0.0669	–	–	–	–
$\pi_{(15\ ^\circ C)}$	0.6404	–	–	–	–
$\pi_{(20\ ^\circ C)}$	0.8000	–	–	–	–
$\pi_{(25\ ^\circ C)}$	0.7800	–	–	–	–
$\pi_{(35\ ^\circ C)}$	0.7400	–	–	–	–
$\pi_{(45\ ^\circ C)}$	0.3800	–	–	–	–

$\boldsymbol{\beta}$, Parameter vector associated with the latency component (including the shape parameter $\boldsymbol{\sigma}$); $\boldsymbol{\gamma}$, parameter vector associated with the incidence component; $\boldsymbol{\pi}$, probability of seed germination.

to 80 and 78 % at 20 and 25 °C, respectively. Meanwhile, at 45 °C π decreased to 38 %, reflecting the deleterious effects of supra-optimal temperatures.

Example 2

In accordance with the selection criterion adopted in the preceding example, the Weibull model with long-term survivors was selected to describe the seed germination of *Anadenanthera colubrina* under different osmotic potentials induced by PEG 6000 (Supplementary Data Fig. S3). This approach provided a better fit than the traditional Weibull model, which tends to underestimate germination speed and overestimate the final germination percentage (Fig. 3). Cox–Snell residuals also indicate goodness of fit (Supplementary Data Fig. S4).

The intercept β_0 represents the reference osmotic potential ($\Psi = -0.8$ MPa), with the parameter $\sigma = 0.3235$ (Table 2). The magnitude of water stress is inversely proportional to osmotic potential; thus, reductions in Ψ intensify the stress on seeds. In the latency submodel the coefficients (β) were negative and significant for most treatments, indicating a difference relative to the reference, except for $\Psi = -0.7$ MPa. Specifically, for the lowest coefficients (β_5 and β_4) seeds germinated more quickly (Fig. 3), demonstrating that water stress reduces germination speed.

Considering the incidence submodel, the germination probability decreased progressively with increasing water stress (Table 2, Fig. 3). At a potential of 0.0 MPa (pure water), the germination percentage was 97 %, decreasing steadily to 22 % at $\Psi = -0.8$ MPa.

Example 3

The Cox semiparametric model with long-term survivors was used to analyse the germination of *Astronium urundeuva* under different osmotic potentials induced by NaCl. As shown in Fig. 4, including the long-term component improved the model goodness of fit. The same pattern can be observed in the residual analysis (Supplementary Data Fig. S5).

Unlike parametric models, which model the survival function, the semiparametric Cox model directly estimates the hazard (λ). Consequently, higher coefficients in the latency submodel increased germination speed (Table 3, Fig. 4).

The incidence component demonstrated that the germination probability increased as water stress decreased (Table 3). At a potential of 0.00 MPa, π was 82.7 %, progressively decreasing to 9.3 % at −0.432 MPa.

Example 4

Following an evaluation based on the comparison of the fitted curves from the parametric models with the Kaplan–Meier estimates, the log-normal model with long-term survivors was selected to describe the germination of *Jacaranda caroba* subjected to different light and temperature conditions (Supplementary Data Fig. S6). This modelling framework provided a superior fit to the conventional log-normal model, successfully capturing the fraction of non-germinated seeds across distinct light and temperature conditions (Fig. 5). The goodness of fit is supported by the residual analysis (Supplementary Data Fig. S7).

In the latency submodel, the β_0 intercept corresponds to the reference condition (darkness and 25–15 °C). The σ parameter was 0.3110, and the photoperiod × temperature interaction was significant (Table 4). The analysis of this interaction revealed that, regardless of the light regime, the alternating temperature of 30–20 °C provided higher germination speeds (Fig. 5A, B).

Regarding the effect of photoperiod, it was found that at 25–15 °C the presence of light reduced the germination time (Fig. 5C). Conversely, under the temperature of 30–20 °C the absence of light promoted an increase in the germination speed (Fig. 5D). These results show that the influence of light on the germination of this species is modulated by temperature.

Similarly, the incidence submodel showed that germination probability also varied with the factors studied (Table 4). Under dark conditions, the germination percentage was 62.2 % at 25–15 °C, increasing to 86 % at 30–20 °C. Under light, the observed values were 66.7 and 77 % for the temperatures of 25–15 and 30–20 °C, respectively.

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

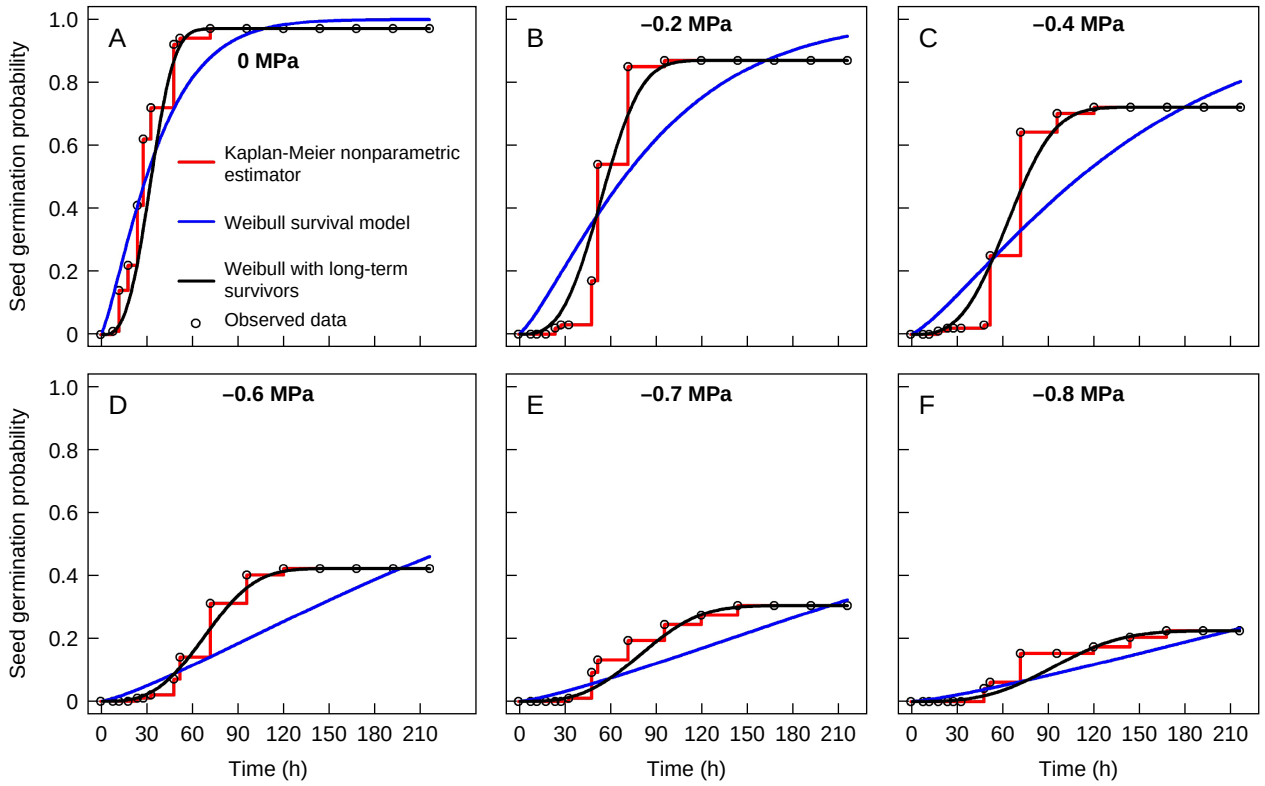

FIG. 3. Time-to-event analysis of *Anadenanthera colubrina* seed germination at different osmotic potentials induced by PEG. Curves represent cumulative germination probabilities estimated by Kaplan–Meier and parametric Weibull models (with and without long-term survivors).

TABLE 2. Estimates, standard error, 95 % confidence interval and *P*-value of the Weibull model with long-term survivors for germination of *Anadenanthera colubrina* seeds under different osmotic potentials induced by PEG.

Parameter	Estimate	Standard error	95 % confidence interval		*P*-value
			Lower	Upper	
$\beta_{0(\text{intercept})}$	4.6677	0.0809	4.5091	4.8263	0.0000
$\beta_{1(\Psi=-0.7\ \text{MPa})}$	−0.1555	0.1242	−0.3989	0.0879	0.2104
$\beta_{2(\Psi=-0.6\ \text{MPa})}$	−0.3175	0.0862	−0.4865	−0.1485	0.0002
$\beta_{3(\Psi=-0.4\ \text{MPa})}$	−0.3925	0.0842	−0.5575	−0.2275	< 0.0001
$\beta_{4(\Psi=-0.2\ \text{MPa})}$	−0.5482	0.0889	−0.7224	−0.3740	< 0.0001
$\beta_{5(\Psi=0.0\ \text{MPa})}$	−1.0598	0.1316	−1.3177	−0.8019	< 0.0001
σ	0.3235	–	–	–	–
$\gamma_{0(\text{intercept})}$	−1.2655	0.2104	−1.6779	−0.8531	< 0.0001
$\gamma_{1(\Psi=-0.7\ \text{MPa})}$	0.4182	0.3298	−0.2282	1.0646	0.2048
$\gamma_{2(\Psi=-0.6\ \text{MPa})}$	0.9427	0.2016	0.5476	1.3378	< 0.0001
$\gamma_{3(\Psi=-0.4\ \text{MPa})}$	2.2100	0.3932	1.43	2.9807	< 0.0001
$\gamma_{4(\Psi=-0.2\ \text{MPa})}$	3.1665	0.4144	2.3543	3.9787	< 0.0001
$\gamma_{5(\Psi=0.0\ \text{MPa})}$	4.7416	0.4471	3.8653	5.6179	0.0000
$\pi_{(\Psi=-0.8\ \text{MPa})}$	0.2200	–	–	–	–
$\pi_{(\Psi=-0.7\ \text{MPa})}$	0.3000	–	–	–	–
$\pi_{(\Psi=-0.6\ \text{MPa})}$	0.4200	–	–	–	–
$\pi_{(\Psi=-0.4\ \text{MPa})}$	0.7200	–	–	–	–
$\pi_{(\Psi=-0.2\ \text{MPa})}$	0.8700	–	–	–	–
$\pi_{(\Psi=0.0\ \text{MPa})}$	0.9700	–	–	–	–

$\boldsymbol{\beta}$: Parameter vector associated with the latency component (including the shape parameter σ); $\boldsymbol{\gamma}$: parameter vector associated with the incidence component; $\boldsymbol{\pi}$: probability of seed germination.

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

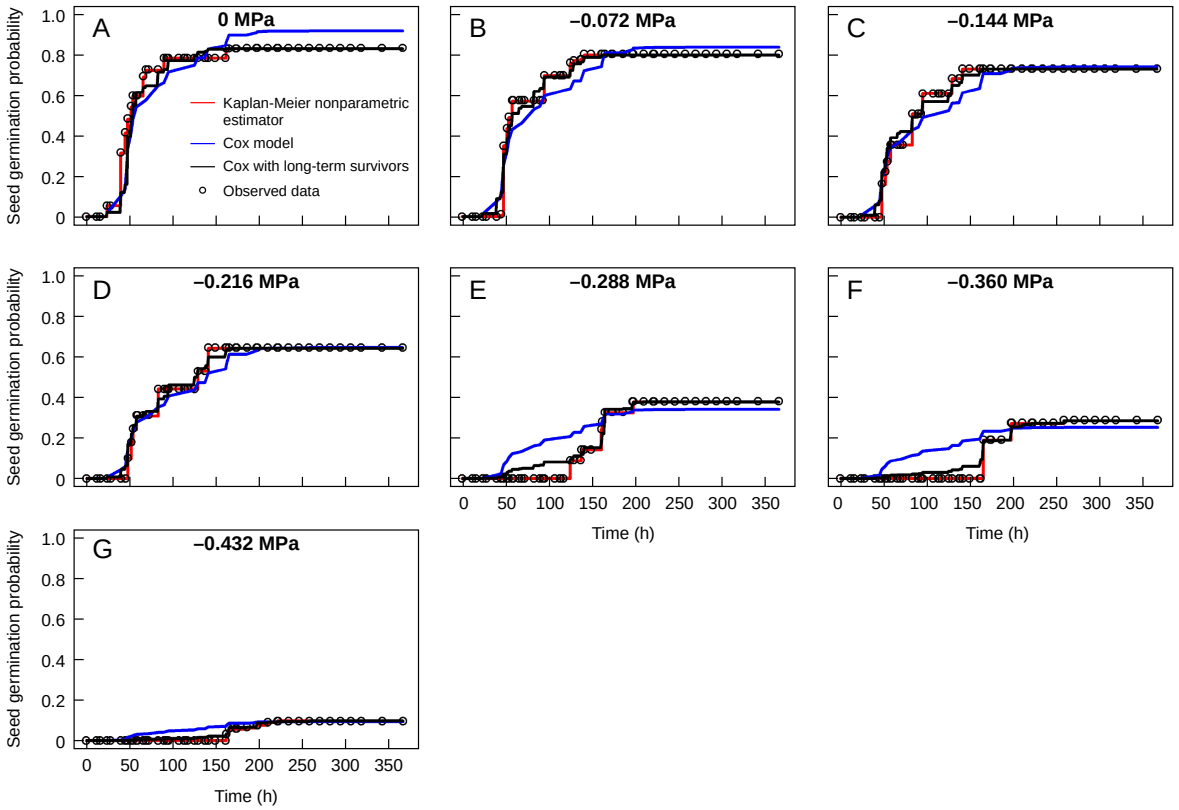

FIG. 4. Time-to-event analysis of *Astronium urundeuva* seed germination at different osmotic potentials induced by NaCl solutions. Curves represent cumulative germination probabilities estimated by Kaplan–Meier and Cox proportional hazards models (with and without long-term survivors).

Example 5

According to the adopted selection criterion, the log-logistic model with long-term survivors was selected to analyse the germination of *Mabea fistulifera* across different combinations of seed coat colour, elaiosome presence and substrate type (Supplementary Data Fig. S8). As illustrated in Figs 6–8, incorporating the long-term component enhanced the goodness of fit for all combinations. This pattern is corroborated by the residual analysis (Supplementary Data Fig. S9).

In the latency submodel, the intercept β_0 represents the reference condition (paper towel, absence of elaiosome and orange colour). The parameter σ was 0.0844 and the TP × PE × BS and TP × PE × RS three-way interactions were significant, indicating complex effects on germination speed (Table 5).

The incidence component showed that the germination probability varied widely across combinations of the studied factors (Table 5). The maximum probability was 96.2 % for PR × AE × RS, while the minimum probability was 72 % for PR × PE × BS.

DISCUSSION

Methodological implications and effectiveness of long-term survival models

The results of this study demonstrate that long-term survival models offer several advantages over traditional approaches. These models naturally accommodate censored observations, such as seeds that did not germinate by the end of the experiment. In traditional analyses, in which final germination percentage and mean germination time are evaluated, these seeds are often excluded, leading to bias and loss of information (Onofri *et al.*, 2011; Romano and Stevanato, 2020; Mamani *et al.*, 2024; Manhães Saint'Clair *et al.*, 2025).

In addition to failing to properly handle censored data, traditional statistical approaches, such as ANOVA and comparison of means, focus on inferences about the final percentage and the germination speed index (GSI) at the end of the experiment. Generally, these results are associated with cumulative germination curves, based on observed data rather than predicted, to describe the temporal behaviour of germination. However, both approaches are applied separately for each covariate, limiting their ability to capture combined effects. In this context, the long-term survivors approach presents the advantage of generating cumulative germination curves based on predictions that incorporate multiple covariates simultaneously. Consequently, it enables germination predictions based on combinations of different covariate values, such as temperature and photoperiod. Moreover, each parameter of these models has a clear biological interpretation: σ, associated with germination uniformity; β, related to germination speed; and π, corresponding to the probability of germination.

Furthermore, long-term mixture models provide separate and interpretable estimates of the incidence and latency components, each with its own confidence intervals and hypothesis tests. This feature enables formal statistical inference on

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

TABLE 3. Estimates, standard error, 95 % confidence interval and *P*-value of the Cox proportional hazards model with long-term survivors for germination of *Astronium urundeuva* seeds under different osmotic potentials induced by NaCl solutions.

Parameter	Estimate	Standard error	95 % confidence interval		*P*-value
			Lower	Upper	
$\beta_{0(\Psi=-0.432\ MPa)}$	–	–	–	–	–
$\beta_{1(\Psi=-0.360\ MPa)}$	−0.0354	0.2220	−0.4705	0.3997	0.8734
$\beta_{2(\Psi=-0.288\ MPa)}$	0.7477	0.2605	0.2371	1.2583	0.0041
$\beta_{3(\Psi=-0.216\ MPa)}$	2.4057	0.2378	1.9396	2.8718	< 0.0001
$\beta_{4(\Psi=-0.144\ MPa)}$	2.5801	0.2369	2.1158	3.0444	< 0.0001
$\beta_{5(\Psi=-0.072\ MPa)}$	2.8630	0.2410	2.3906	3.3354	< 0.0001
$\beta_{6(\Psi=0.000\ MPa)}$	3.1408	0.2883	2.5757	3.7059	< 0.0001
$\gamma_{0(intercept)}$	−2.2736	0.2998	−2.8613	−1.6859	< 0.0001
$\gamma_{1(\Psi=-0.360\ MPa)}$	1.3291	0.2371	0.8644	1.7939	< 0.0001
$\gamma_{2(\Psi=-0.288\ MPa)}$	1.7557	0.3956	0.9803	2.5310	< 0.0001
$\gamma_{3(\Psi=-0.216\ MPa)}$	2.8490	0.2943	2.2.22	3.4257	< 0.0001
$\gamma_{4(\Psi=-0.144\ MPa)}$	3.2514	0.3585	2.54	3.9540	< 0.0001
$\gamma_{5(\Psi=-0.072\ MPa)}$	3.6187	0.3046	3.0217	4.2158	< 0.0001
$\gamma_{6(\Psi=0.000\ MPa)}$	3.8358	0.3460	3.1577	4.5139	< 0.0001
$\pi_{(\Psi=-0.432\ MPa)}$	0.0933	–		–	–
$\pi_{(\Psi=-0.360\ MPa)}$	0.2800	–	–	–	–
$\pi_{(\Psi=-0.288\ MPa)}$	0.3733	–	–	–	–
$\pi_{(\Psi=-0.216\ MPa)}$	0.6400	–	–	–	–
$\pi_{(\Psi=-0.144\ MPa)}$	0.7267	–	–	–	–
$\pi_{(\Psi=-0.072\ MPa)}$	0.7933	–	–	–	–
$\pi_{(\Psi=0.00\ MPa)}$	0.8267	–	–	–	–

$\boldsymbol{\beta}$: Parameter vector associated with the latency component (including the shape parameter $\boldsymbol{\sigma}$); $\boldsymbol{\gamma}$: parameter vector associated with the incidence component; $\boldsymbol{\pi}$: probability of seed germination.

treatment effects for each component of the germination process, which is not possible when using summary indices such as final germination percentage or mean germination time (Onofri *et al.*, 2011; Mamani *et al.*, 2024).

Again, the structure of mixture models allows the incorporation of different parametric distributions or a semiparametric approach for the latency component, providing flexibility to capture different shapes of germination curves. In the present study, different distributions were selected for distinct species and experimental conditions, reflecting real biological differences in germination dynamics. Overall, long-term models showed better fit than traditional survival models. This is evidenced by agreement with the Kaplan–Meier estimator, observed data and modified Cox–Snell residuals.

Traditional survival models assume that the distribution function $F(t)$, representing the germination curve, is proper in the sense that it converges to 1 (100 % germination) as $t \to \infty$ (Berkson and Gage, 1952; Peng and Yu, 2021). However, this behaviour does not always reflect the biological reality of seed assays. In contrast, in long-term models the inclusion of the incidence component makes $F(t)$ an improper function, characterized by growth that stabilizes at a level below 1 (Berkson and Gage, 1952; Peng and Yu, 2021). Thus, these models allow capturing the fraction of seeds that do not germinate, providing a more accurate estimate of the seed lot's germination capacity.

Biological interpretation of the long-term survivors in native forest seeds

The results of this study demonstrated that long-term survival models effectively capture the heterogeneity of forest seed lots. This approach offers a biological interpretation superior to traditional models. In this context, the non-germinated fraction $(1 - \pi)$ reflects a complex mixture of physiological states. This includes non-viable seeds due to mechanical damage, predation or senescence, as well as those exhibiting secondary dormancy induced by stress. Finally, it also encompasses seeds that, while viable, lack sufficient reserves or structural integrity to complete germination (Onofri *et al.*, 2011; Manso *et al.*, 2013).

The disentangling of germination capacity (incidence) from speed (latency) is fundamental in these models to understanding responses to environmental factors (Onofri *et al.*, 2011). This premise, discussed by Manso *et al.* (2013) for *Pinus pinea*, maintains that long-term survival models can capture natural germination variability with greater precision than traditional approaches.

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

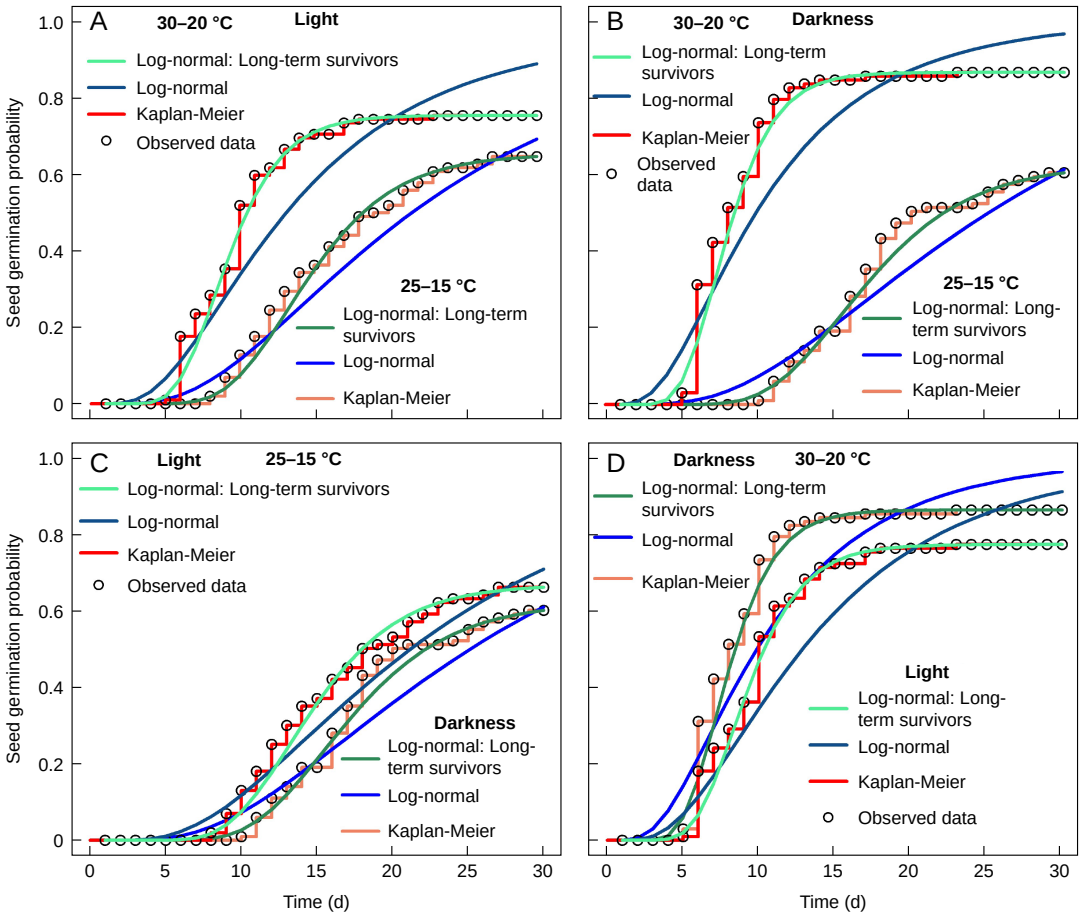

FIG. 5. Time-to-event analysis of *Jacaranda caroba* seed germination under different alternating temperature and light conditions. Curves represent cumulative germination probabilities estimated by Kaplan–Meier and parametric log-normal models (with and without long-term survivors).

TABLE 4. Estimates, standard error, 95 % confidence interval and *P*-value of the log-normal model with long-term survivors for the germination of *Jacaranda caroba* seeds under different light and temperature conditions.

Parameter	Estimate	Standard error	95 % confidence interval		*P*-value
			Lower	Upper	
$\beta_{0(\text{intercept})}$	2.8358	0.0388	2.7598	2.9118	0.0000
$\beta_{1(\text{photoperiod})}$	−0.1557	0.0587	−0.2708	−0.0406	0.0080
$\beta_{2(\text{temperature})}$	−0.7667	0.0506	−0.8659	−0.6675	0.0000
$\beta_{3(\text{photoperiod} \times \text{temperature})}$	0.3048	0.0723	0.1631	0.4465	< 0.0001
σ	0.3110	–	–	–	–
$\gamma_{0(\text{intercept})}$	0.4959	0.2214	0.0620	0.9298	0.0251
$\gamma_{1(\text{photoperiod})}$	0.1979	0.3108	−0.4113	0.8071	0.5242
$\gamma_{2(\text{temperature})}$	1.3195	0.3964	0.5426	2.0964	0.0009
$\gamma_{3(\text{photoperiod} \times \text{temperature})}$	−0.8047	0.5029	−1.7904	0.1810	0.1096
$\pi_{(\text{darkness and 25–15 °C})}$	0.6215	–		–	–
$\pi_{(\text{darkness and 30–20 °C})}$	0.8600	–	–	–	–
$\pi_{(\text{light and 25–15 °C})}$	0.6668	–	–	–	–
$\pi_{(\text{light and 30–20 °C})}$	0.7700	–	–	–	–

$\boldsymbol{\beta}$: Parameter vector associated with the latency component (including the shape parameter σ); $\boldsymbol{\gamma}$: parameter vector associated with the incidence component; $\boldsymbol{\pi}$: probability of seed germination.

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

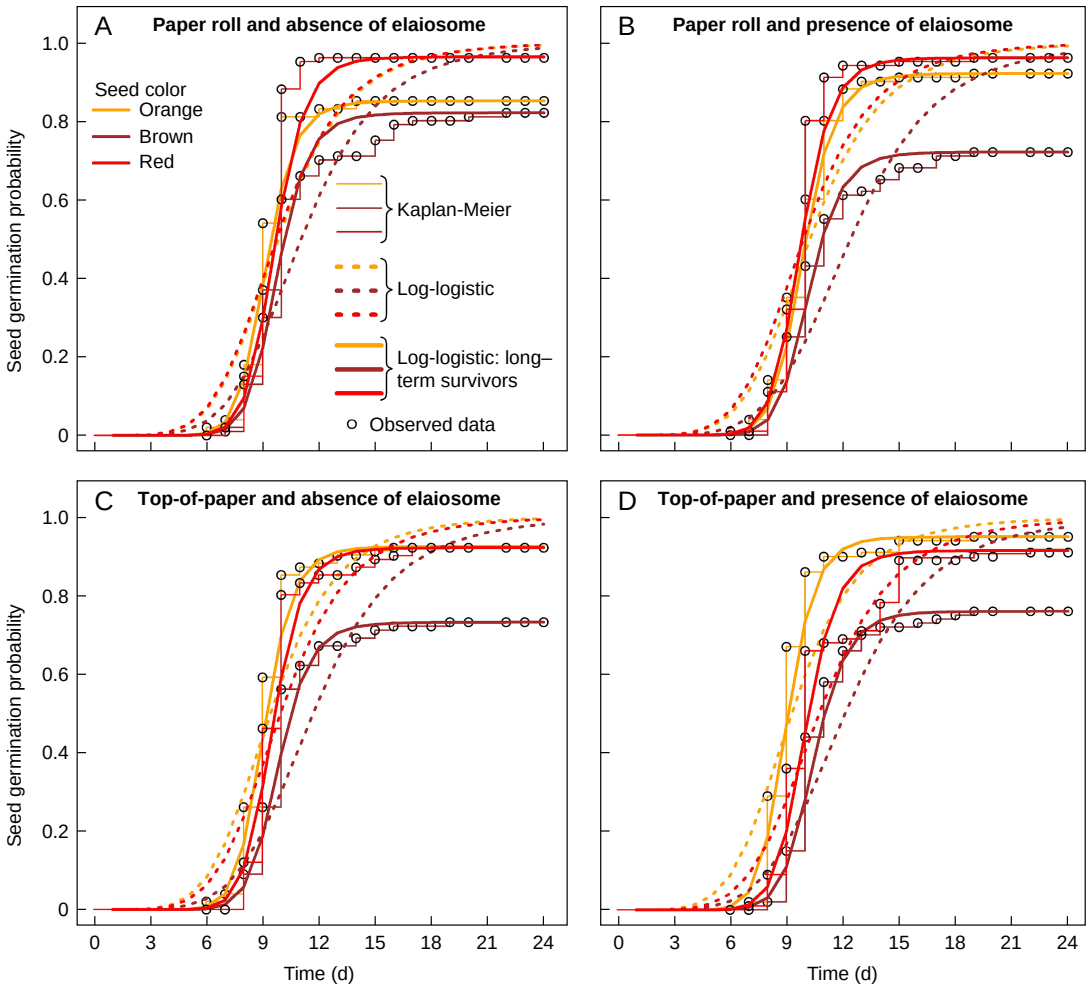

FIG. 6. Time-to-event analysis of *Mabea fistulifera* seed germination with different seed coat colours, as a function of substrate type and the presence or absence of elaiosomes. Curves represent cumulative germination probabilities estimated by Kaplan–Meier and parametric log-logistic models (with and without long-term survivors).

In the present study, the germination percentage varied widely among species and experimental conditions. For instance, in *Anadenanthera colubrina* at 5 °C germination was 6.7 %, reflecting the inability of most seeds to germinate under suboptimal temperatures. In contrast, at 20–25 °C the germination percentage was 80–78 %, indicating conditions close to the physiological optimum (Fig. 2). Under severe osmotic stress (−0.8 MPa), the percentage of germinated seeds was 22 %, suggesting that water stress not only delays germination but also prevents it in a large proportion of seeds (Fig. 3). These patterns support the hypothesis that the cure fraction reflects critical physiological limitations, in which seeds fail to take up water for growth, mobilize reserves and activate metabolism (Bewley *et al.*, 2013).

Effects of environmental factors on incidence and latency components

One of the contributions of this study lies in the finding that environmental factors affect the incidence (π) and latency components of seed germination in distinct ways. In *Anadenanthera colubrina* temperature affected both the probability and the speed of germination, but in different ways. Suboptimal (5 °C) and supra-optimal (45 °C) temperatures drastically reduced π, whereas intermediate temperatures (15–35 °C) had relatively minor effects on incidence but significantly accelerated germination speed (Fig. 2). This pattern suggests that temperature functions primarily as a viability filter under thermal extremes and as a metabolic rate modulator under favourable conditions (Meeussen *et al.*, 2022; Mamani *et al.*, 2024).

Osmotic stress in *Anadenanthera colubrina* and saline stress in *Astronium urundeuva* showed similar patterns, with more pronounced effects on the incidence component than on the latency component (Figs 3 and 4). In *Anadenanthera colubrina* the germination percentage decreased from 97 % (0.0 MPa) to 22 % (−0.8 MPa), a 75 % reduction, whereas the β coefficients for latency declined moderately (Table 2). Similarly, in *Astronium urundeuva* the percentage declined from 82.7 (0.0 MPa) to 9.3 (−0.432 MPa), representing a 73.4 % reduction (Table 3). These results indicate that water and salt stress primarily reduce the fraction of seeds able to germinate, possibly due to toxic effects, osmotic imbalance or inhibition of imbibition

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

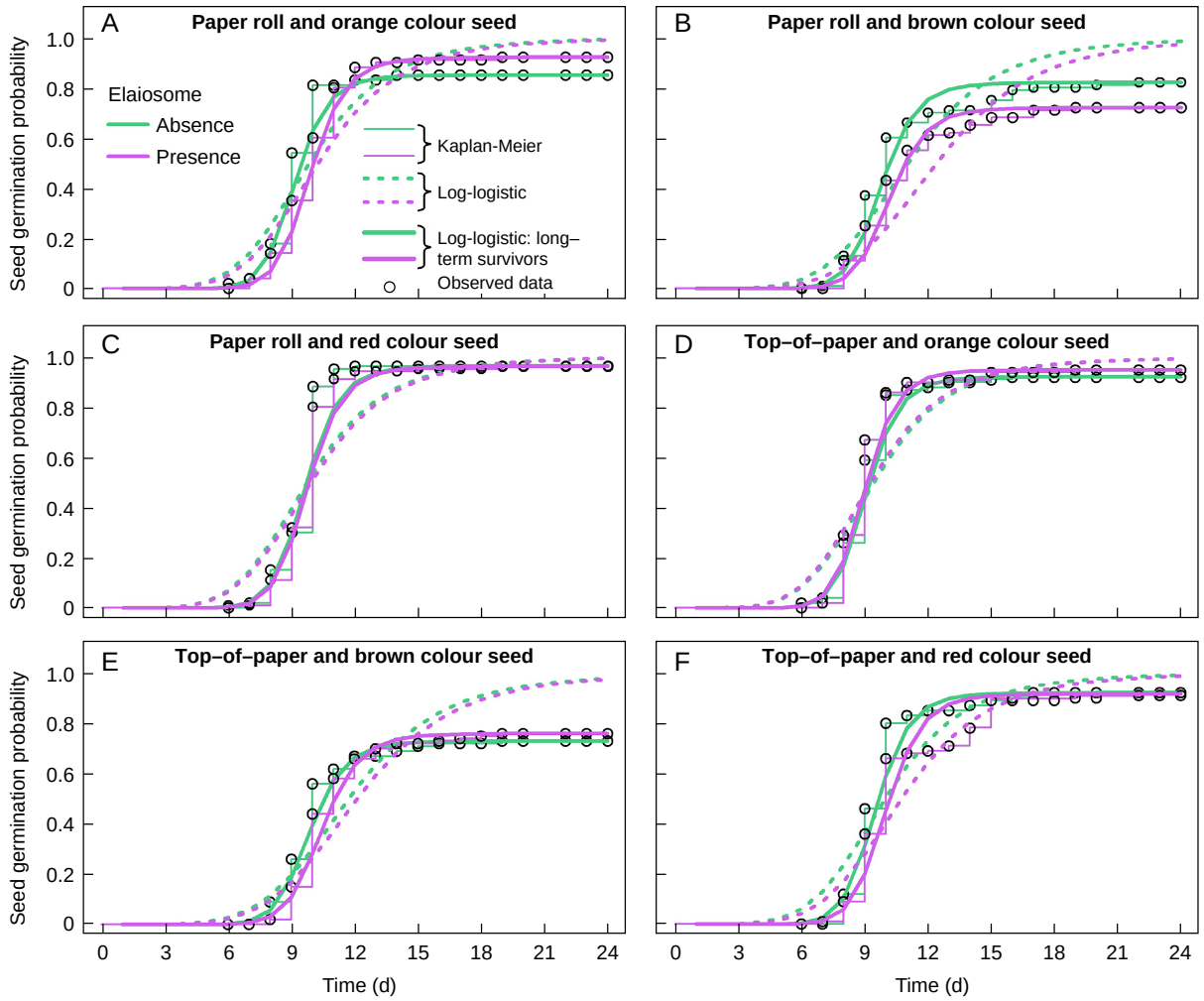

FIG. 7. Time-to-event analysis of *Mabea fistulifera* seed germination in the presence or absence of elaiosomes, as a function of substrate type and seed coat colours. Curves represent cumulative germination probabilities estimated by Kaplan–Meier and parametric log-logistic models (with and without long-term survivors).

(Bewley *et al.*, 2013; Daibes and Cardoso, 2020). However, those that germinate maintain relatively similar speed.

These results have significant implications for understanding the responses of forest species to climate change. Models that assume complete germination and focus solely on germination speed may drastically underestimate the impacts of environmental stresses, particularly droughts and heatwaves. These events can reduce germination percentage without necessarily affecting the speed. Carme *et al.* (2026) report that climate warming affects germination time in populations of *Quercus suber*, but they did not quantify its effects on the germination percentage. Additionally, Dantas *et al.* (2020) reported that temperature and osmotic potential affect germination percentage and germination rate in dry forest species, although rainfall reduction was identified as the main factor restricting the temporal window suitable for germination.

Ecological effects of the incidence and latency components

In *Jacaranda caroba*, germination percentage ranged from 62 to 86 % depending on light and temperature conditions (Fig. 5, Table 4). These factors influenced the incidence component. In contrast, the latency component was predominantly influenced by temperatures of 30–20 °C (Fig. 5A, B). In ecological terms, the sensitivity of these components to temperature may represent an adaptive mechanism that synchronizes germination with resource availability and reduces competition. This behaviour is consistent with the pioneer character of this species, which colonizes forest gaps and edges. In these environments, higher temperatures and greater light availability are reliable indicators of favourable conditions for seedling establishment (Lacerda and Figueiredo, 2009; Mamani *et al.*, 2024).

In addition to responses associated with environmental conditions, the non-germinating fraction may also reflect variation in intrinsic seed traits within the same seed lot. In *Mabea fistulifera*, variation in germination percentage associated with seed coat colour (72–96 %) suggests heterogeneity in the degree of maturation or physiological quality within the same seed lot (Table 5). Seeds with a red seed coat showed a higher probability of germination when the elaiosome was absent, possibly because its removal

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

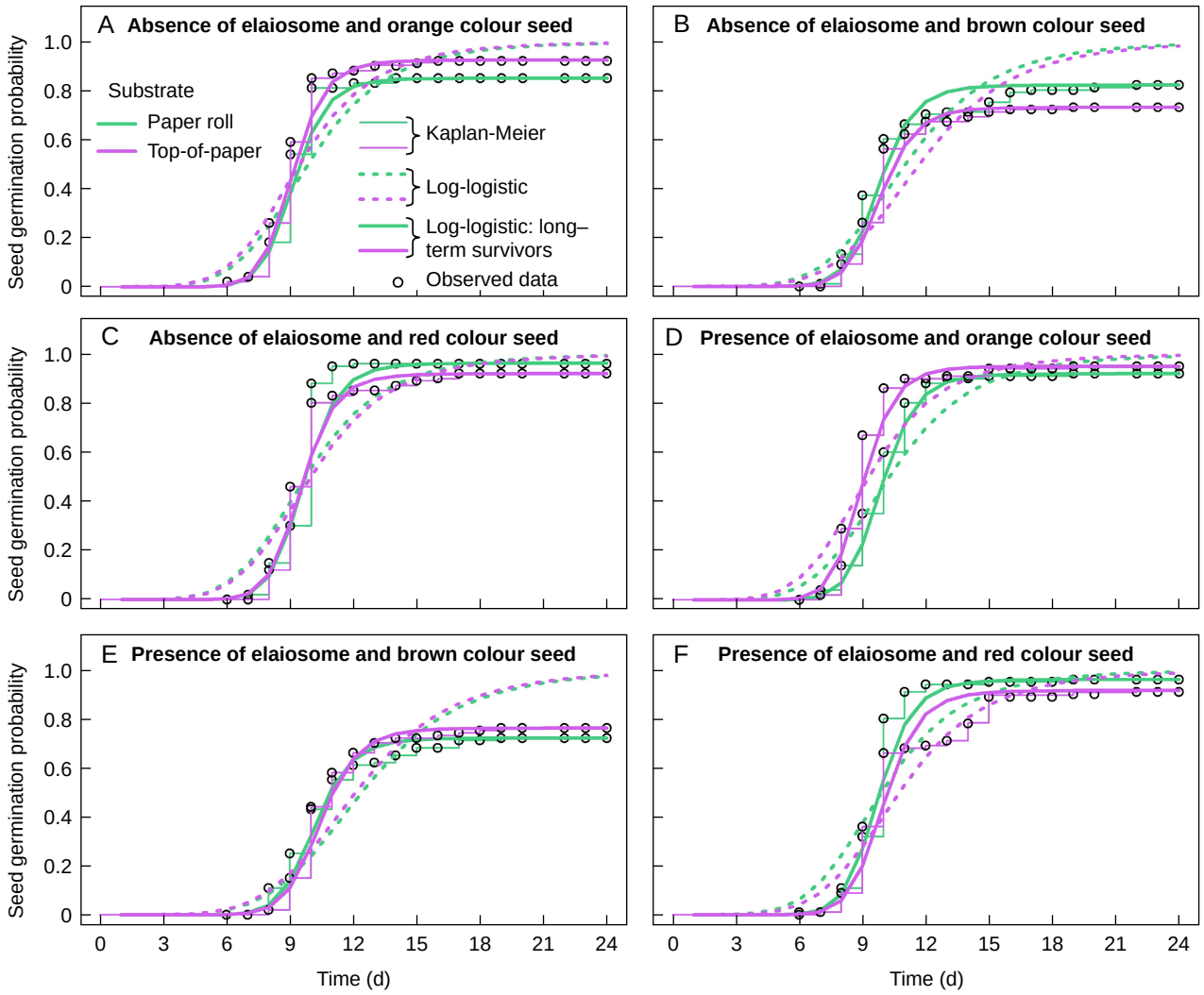

FIG. 8. Time-to-event analysis of *Mabea fistulifera* seed germination with different substrate types, as a function of seed coat colour and the presence or absence of elaiosomes. Curves represent cumulative germination probabilities estimated by Kaplan–Meier and parametric log-logistic models (with and without long-term survivors).

eliminates physical or chemical barriers to imbibition (Santos *et al.*, 2025). This complex interaction between seed traits and post-harvest treatments underscores the importance of accounting for within-lot heterogeneity in germination studies.

Practical applications for restoration and conservation

The results of this study have direct practical applications for ecological restoration, seedling production and *ex situ* conservation of native forest species. In this context, quantifying the cure fraction under different conditions enables optimization of germination protocols in nurseries. For *Anadenanthera colubrina*, for example, temperatures of 20–25 °C maximize both the probability (78–80 %) and the speed of germination, whereas temperatures of 5 or 45 °C should be avoided (Fig. 2). Similarly, for *Astronium urundeuva*, high sensitivity to salt stress (Fig. 4) indicates that this species is not suitable for sowing in saline substrates but can be used in areas with low salinity.

In this scenario, models with a cure fraction allow more realistic predictions of success in restoration projects. Establishment of forest seedlings is often limited by low germination percentage and high mortality (Cole *et al.*, 2011; Piaia *et al.*, 2023). Traditional survival models, which assume complete germination, may overestimate the number of seeds required. In contrast, long-term models provide more accurate estimates of the seeding density needed to achieve target seedling densities, accounting for both germination capacity and speed (Onofri *et al.*, 2011).

Long-term models can also support *ex situ* conservation strategies in germplasm banks. Quantifying the germinating fraction under controlled conditions provides a more informative seed lot quality metric than final germination percentage. This approach offers a more precise and biologically meaningful way to characterize seed germination, enabling improved decision-making for seed collection, storage and propagation strategies critical to *ex situ* conservation programmes (El-Kassaby *et al.*, 2008).

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

TABLE 5. Estimates, standard error, 95 % confidence interval and *P*-value of the log-logistic model with long-term survivors for germination of *Mabea fistulifera* seeds for different seed coat colours, elaiosome conditions and substrate types.

Parameter	Estimate	Standard error	95 % confidence interval		*P*-value
			Lower	Upper	
$\beta_{0(\text{intercept})}$	2.2144	0.0114	2.1920	2.2369	0.0000
$\beta_{1(\text{TP})}$	−0.0065	0.0165	−0.0390	0.0259	0.6929
$\beta_{2(\text{PE})}$	0.0775	0.0174	0.0434	0.1116	0.0000
$\beta_{3(\text{BS})}$	0.0658	0.0222	0.0223	0.1093	0.0030
$\beta_{4(\text{RS})}$	0.0514	0.0154		0.0815	0.0008
$\beta_{5(\text{TP×PE})}$	−0.0864	0.0263	−0.1380	−0.0349	0.0010
$\beta_{6(\text{TP×BS})}$	0.0146	0.0305	−0.0452	0.0745	0.6317
$\beta_{7(\text{TP×RS})}$	−0.0054	0.0227	−0.0499	0.0392	0.8137
$\beta_{8(\text{PE×BS})}$	−0.0376	0.0375	−0.1111	0.0359	0.3160
$\beta_{9(\text{PE×RS})}$	−0.0663	0.0245	−0.1143	−0.0183	0.0068
$\beta_{10(\text{TP×PE×BS})}$	0.1060	0.0471	0.0137	0.1984	0.0244
$\beta_{11(\text{TP×PE×RS})}$	0.1265	0.0393	0.0495	0.2034	0.0013
σ	0.0844	–	–	–	–
$\gamma_{0(\text{intercept})}$	1.7348	0.3328	1.0826	2.3870	0.0000
$\gamma_{1(\text{TP})}$	0.7495	0.5094	−0.2.89	1.7478	0.1412
$\gamma_{2(\text{PE})}$	0.7079	0.5453	−0.3610	1.7767	0.1943
$\gamma_{3(\text{BS})}$	−0.2183	0.4292	−1.0597	0.6230	0.6110
$\gamma_{4(\text{RS})}$	1.5060	0.8971	−0.2523	3.2642	0.0932
$\gamma_{5(\text{TP×PE})}$	−0.2476	0.9446	−2.0990	1.6039	0.7933
$\gamma_{6(\text{TP×BS})}$	−1.2713	0.6243	−2.4949	−0.0477	0.0417
$\gamma_{7(\text{TP×RS})}$	−1.5477	1.1206	−3.7440	0.6486	0.1672
$\gamma_{8(\text{PE×BS})}$	−1.2798	0.6768	−2.60	0.0468	0.0586
$\gamma_{9(\text{PE×RS})}$	−0.7700	1.8349	−4.3664	2.8263	0.6747
$\gamma_{10(\text{TP×PE×BS})}$	0.9776	1.0457	−1.07	3.0271	0.3498
$\gamma_{11(\text{TP×PE×RS})}$	0.2483	2.1139	−3.8949	4.3914	0.9065
$\pi_{(\text{PR×AE×OS})}$	0.8500	–	–	–	–
$\pi_{(\text{TP×AE×OS})}$	0.9300	–	–	–	–
$\pi_{(\text{PR×AE×BS})}$	0.8200	–	–	–	–
$\pi_{(\text{TP×AE×BS})}$	0.7300	–	–	–	–
$\pi_{(\text{PR×AE×RS})}$	0.9623	–	–	–	–
$\pi_{(\text{TP×AE×RS})}$	0.9200	–	–	–	–
$\pi_{(\text{PR×PE×OS})}$	0.9200	–	–	–	–
$\pi_{(\text{TP×PE×OS})}$	0.9500	–	–	–	–
$\pi_{(\text{PR×PE×BS})}$	0.7200	–	–	–	–
$\pi_{(\text{TP×PE×BS})}$	0.7600	–	–		–
$\pi_{(\text{PR×PE×RS})}$	0.9600	–	–	–	–
$\pi_{(\text{TP×PE×RS})}$	0.9154	–	–		–

$\boldsymbol{\beta}$: Parameter vector associated with the latency component (including the shape parameter σ); $\boldsymbol{\gamma}$:. parameter vector associated with the incidence component; π: probability of seed germination.
PR, paper roll; TP, top of paper; AE, absence of elaiosome; PE, presence of elaiosome; OS, orange colour seed; BS, brown colour seed; RS, red colour seed.

Methodological challenges and perspectives for future studies

The biological interpretation of the cure fraction remains ambiguous, as it is not possible to distinguish among non-viable seeds, dormant seeds and those that simply did not germinate under the specific experimental conditions without additional viability tests (tetrazolium, X-ray, etc.) and dormancy-breaking treatments (Onofri *et al.*, 2011). Thus, future studies should combine long-term survival models

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

with viability tests and dormancy-breaking treatments to decompose the cure fraction into its different biological components.

In addition, this study did not explore the effects of time-dependent covariates, such as temperature fluctuations throughout the experiment, nor did it examine spatial correlation structures or random effects (frailty), for example, among seeds from the same mother tree or population. Future extensions of cure fraction models should incorporate these complexities, enabling more realistic analyses of germination data (Manso *et al.*, 2013; Romano and Stevanato, 2020). With these advances, long-term survival models have the potential to become established tools for analysing germination data in forest ecology and seed science.

Conclusions

This study demonstrates that long-term survival models provide a robust, biologically realistic statistical framework for analysing germination data from forest seeds. Partitioning the germination process into incidence (probability of germination) and latency (time to germination) components enabled key advances. First, it allowed precise quantification of the fractions of germinating and non-germinating seeds. Second, it provided more accurate and informative estimates than traditional survival methods. Finally, it enabled the isolation of the effects of environmental factors on germination capacity and speed.

The results indicate that stressful conditions, such as extreme temperatures and water limitation, predominantly affect the incidence component. Under these scenarios, there is a marked reduction in the fraction of seeds capable of germinating. In contrast, these factors have smaller effects on latency.

This distinction is crucial for predicting the responses of forest species to climate change. In addition, the findings support optimization of seedling production protocols and ecological restoration planning. Given these results, in the presence of a non-germinating seed fraction, long-term survival models are recommended, especially for native forest species.

SUPPLEMENTARY DATA

Supplementary data are available at *Annals of Botany* online and consist of the following. Figure S1: time-to-event analysis of *Anadenanthera colubrina* seed germination under different temperatures. Curves represent cumulative germination probabilities estimated by Kaplan–Meier (dashed lines) and parametric models: (A) exponential, (B) Weibull, (C) log-normal and (D) log-logistic. Log-normal was the selected model. Figure S2: Cox–Snell residuals for the time-to-event analysis of *Anadenanthera colubrina* seed germination under different temperatures, using (A) a log-normal model and (B) a log-normal model with long-term survivors. Figure S3: time-to-event analysis of *Anadenanthera colubrina* seed germination at different osmotic potentials induced by PEG 6000. Curves represent cumulative germination probabilities estimated by Kaplan–Meier (dashed lines) and parametric models: (A) Exponential, (B) Weibull, (C) log-normal and (D) log-logistic. Weibull was the selected model. Figure S4: Cox–Snell residuals for the time-to-event analysis of *Anadenanthera colubrina* seed germination at different osmotic potentials induced by PEG, using (A) a Weibull model and (B) a Weibull model with long-term survivors. Figure S5: Cox–Snell residuals for the time-to-event analysis of *Astronium urundeuva* seed germination at different osmotic potentials induced by NaCl solutions, using (A) a Cox model and (B) a Cox model with long-term survivors. Figure S6: time-to-event analysis of *Jacaranda caroba* seed germination at different alternating temperature and light conditions. Curves represent cumulative germination probabilities estimated by Kaplan–Meier (dashed lines) and parametric models: (A) exponential, (B) Weibull, (C) log-normal and (D) log-logistic. Log-normal was the selected model. Figure S7: Cox–Snell residuals for the time-to-event analysis of *Jacaranda caroba* seed germination at different alternating temperature and light conditions, using (A) a log-normal model and (B) a log-normal model with long-term survivors. Figure S8: time-to-event analysis of *Mabea fistulifera* seed germination with different seed coat colours, substrate type and the presence or absence of elaiosomes. Curves represent cumulative germination probabilities estimated by Kaplan–Meier (dashed lines) and parametric models: (A) exponential, (B) Weibull, (C) log-normal and (D) log-logistic. PR, paper roll; TP, top-of-paper; AE, absence of elaiosome; PE, presence of elaiosome; OS, orange colour seed; BS, brown colour seed; RS, red colour seed. Log-logistic was the selected model. Figure S9: Cox–Snell residuals for the time-to-event analysis of *Mabea fistulifera* seed germination with different seed coat colours, substrate types and the presence or absence of elaiosomes, using (A) a Log-logistic model and (B) a Log-logistic model with long-term survivors.

FUNDING

The authors express their sincere thanks to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Finance Code 001), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Grant 318412/2025-0) and the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for their financial support.

ACKNOWLEDGEMENTS

The authors are grateful to the reviewers and editors for their valuable suggestions.

DATA AVAILABILITY

The data are available at https://doi.org/10.5281/zenodo.20853968

AUTHOR CONTRIBUTIONS

S.M.F., M.L.D.: conceptualization, methodology, formal analysis, data curation, visualization, investigation,

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

resources, writing – original draft, writing – review and editing. L.S.A., G.M.A., B.F.D.: writing – review and editing. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CODE AVAILABILITY

The code are available at https://doi.org/10.5281/zenodo.20853968

AI ASSISTANCE

During the preparation of this work, the authors utilized Gemini and Grammarly solely for linguistic editing, grammar checking and minor stylistic improvements to enhance the readability of the manuscript. These AI tools were used under the direct oversight of the authors, who carefully reviewed and edited the output. Consequently, the authors take full responsibility for the final content, integrity and conclusions of the published article.

REFERENCES

Barak RS, **Lichtenberger TM**, **Wellman-Houde A**, **Kramer AT**, **Larkin DJ. 2018**. Cracking the case: seed traits and phylogeny predict time to germination in prairie restoration species. *Ecology and Evolution* **8**: 5551–5562. doi:10.1002/ece3.4083

Berkson J, **Gage RP. 1952**. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47**: 501–515. doi:10.2307/2281318

Bernau VM, **Jardón Barbolla L**, **McHale LK**, **Mercer KL. 2020**. Germination response of diverse wild and landrace Chile peppers (*Capsicum* spp.) under drought stress simulated with polyethylene glycol. *PLoS One* **15**: e0236001. doi:10.1371/journal.pone.0236001

Bewley JD, **Bradford KJ**, **Hilhorst HWM**, **Nonogaki H. 2013**. *Seeds: physiology of development, germination and dormancy*, 3rd edn. New York: Springer. doi:10.1007/978-1-4614-4693-4

Boag JW. 1949. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society: Series B, Statistical Methodology* **11**: 15–44. doi:10.1111/j.2517-6161.1949.tb00020.x

Carme M, **Vicente E**, **Costa ESF**, ***et al.* 2026**. Seed origin determines cork oak germination: the warmer the higher, faster and more synchronized. *New Forests* **57**: 20. doi:10.1007/s11056-026-10163-8

Carta A, **Fernández-Pascual E**, **Gioria M**, ***et al.* 2022**. Climate shapes the seed germination niche of temperate flowering plants: a meta-analysis of European seed conservation data. *Annals of Botany* **129**: 775–786. doi:10.1093/aob/mcac037

Cole RJ, **Holl KD**, **Keene CL**, **Zahawi RA. 2011**. Direct seeding of late-successional trees to restore tropical montane forest. *Forest Ecology and Management* **261**: 1590–1597. doi:10.1016/j.foreco.2010.06.038

Cox DR. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B, Statistical Methodology* **34**: 187–202. doi:10.1111/j.2517-6161.1972.tb00899.x

Daibes LF, **Cardoso VJM. 2020**. Effect of reduced water potential on seed germination of a forest tree: a hydrotime approach. *Journal of Seed Science* **42**: e202042003. doi:10.1590/2317-1545v42224519

Dantas BF, **Moura MSB**, **Pelacani CR**, ***et al.* 2020**. Rainfall, not soil temperature, will limit the seed germination of dry forest species with climate change. *Oecologia* **192**: 529–541. doi:10.1007/s00442-019-04575-x

de Santana Santos B, **Santos CS**, **Meiado MV. 2026**. Tolerance to 1 year of drought: the role of seed hydration memory in the germination and early seedling development of *Xiquexique gounellei* subsp. *gounellei* under prolonged water deficit. *Plant Species Biology* **41**: e70043. doi:10.1111/1442-1984.70043

El-Kassaby YA, **Moss I**, **Kolotelo D**, **Stoehr M. 2008**. Seed germination: mathematical representation and parameters extraction. *Forest Science* **54**: 220–227. doi:10.1093/forestscience/54.2.220

Gomes SEV, **Gomes RA**, **Dantas BF. 2023**. Climate and seed size of a dry forest species: influence on seed production, physiological quality, and tolerance to abiotic stresses. *Journal of Seed Science* **45**: e202345013. doi:10.1590/2317-1545v45264166

Harich FK, **Treydte AC**, **Ogutu JO**, ***et al.* 2016**. Seed dispersal potential of Asian elephants. *Acta Oecologica* **77**: 144–151. doi:10.1016/j.actao.2016.10.005

Hsu H-W, **Stuke M**, **Bakker JD**, **Kim S-H. 2024**. A time-to-event analysis for temperature dependence of seed germination in four conifers: ecological niche and environmental gradients. *Forest Ecology and Management* **562**: 121972. doi:10.1016/j.foreco.2024.121972

Kaplan EL, **Meier P. 1958**. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53**: 457–481. doi:10.1080/01621459.1958.10501452

Lacerda DMA, **Figueiredo PSD. 2009**. Restauração de matas ciliares do rio Mearim no município de Barra do Corda-MA: seleção de espécies e comparação de metodologias de reflorestamento. *Acta Amazonica* **39**: 295–303. doi:10.1590/S0044-59672009000200008

Li C-S, **Taylor JMG. 2002**. A semi-parametric accelerated failure time cure model. *Statistics in Medicine* **21**: 3235–3247. doi:10.1002/sim.1260

Lin DY, **Wei LJ. 1989**. The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* **84**: 1074–1078. doi:10.1080/01621459.1989.10478874

Mamani GQ, **Duarte ML**, **Almeida LSD**, **Martins Filho S. 2024**. Non-parametric survival analysis in seed germination of forest species. *Journal of Seed Science* **46**: e202446036. doi:10.1590/2317-1545v46288345

Manhães Saint'Clair V, **Lelis Duarte M**, **Dos Anjos da Silva M**, **Soraya de Almeida L**, **Martins Filho S. 2025**. Parametric time-to-event analysis to evaluate the germination of faba bean and lentil seeds. *Brazilian Journal of Biometrics* **43**, e-43763. doi:10.28951/bjb.v43i4.763

Manso R, **Fortin M**, **Calama R**, **Pardos M. 2013**. Modelling seed germination in forest tree species through survival analysis. The *Pinus pinea* L. case study. *Forest Ecology and Management* **289**: 515–524. doi:10.1016/j.foreco.2012.10.028

McNair JN, **Sunkara A**, **Frobish D. 2012**. How to analyse seed germination data using statistical time-to-event analysis: non-parametric and semi-parametric methods. *Seed Science Research* **22**: 77–95. doi:10.1017/S0960258511000547

Meeussen C, **De Pauw K**, **Sanczuk P**, ***et al.* 2022**. Initial oak regeneration responses to experimental warming along microclimatic and macroclimatic gradients. *Plant Biology* **24**: 745–757. doi:10.1111/plb.13412

Oliveira GMD, **Silva FFSD**, **Araujo MDN**, ***et al.* 2019**. Environmental stress, future climate, and germination of *Myracrodruon urundeuva* seeds 1. *Journal of Seed Science* **41**: 32–43. doi:10.1590/2317-1545v41n1191945

Onofri A, **Gresta F**, **Tei F. 2010**. A new method for the analysis of germination and emergence data of weed species. *Weed Research* **50**: 187–198. doi:10.1111/j.1365-3180.2010.00776.x

Onofri A, **Mesgaran MB**, **Tei F**, **Cousens RD. 2011**. The cure model: an improved way to describe seed germination? *Weed Research* **51**: 516–524. doi:10.1111/j.1365-3180.2011.00870.x

Onofri A, **Mesgaran MB**, **Ritz C. 2022**. A unified framework for the analysis of germination, emergence, and other time-to-event data in weed science. *Weed Science* **70**: 259–271. doi:10.1017/wsc.2022.8

Ordóñez-Parra CA, **Dayrell RLC**, **Negreiros D**, ***et al.* 2023**. Rock n' Seeds: a database of seed functional traits and germination experiments from Brazilian rock outcrop vegetation. *Ecology* **104**: e3852. doi:10.1002/ecy.3852

Patilea V, **Van Keilegom I. 2020**. A general approach for cure models in survival analysis. *Annals of Statistics* **48**, 2323–2346. doi:10.1214/19-AOS1889

Peng Y, **Taylor JMG. 2017**. Residual-based model diagnosis methods for mixture cure models. *Biometrics* **73**: 495–505. doi:10.1111/biom.12582

Peng Y. *mixcure: mixture cure models. R package version 2.0*. 2020. doi:10.32614/CRAN.package.mixcure.

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026

Peng Y, **Yu B. 2021**. *Cure models: methods, applications, and implementation*. New York: Chapman and Hall/CRC. doi:10.1201/9780429032301

Piaia BB, **Rovedder APM**, **Giacomini IF**, ***et al.* 2023**. Direct seeding for forest restoration in southern Brazil: influence of soil conditions and subtropical climate. *Revista Acta Ambiental Catarinense* **20**: 01–12. doi:10.24021/raac.v20i1.6142

R Core Team**. 2026**. *R: a language and environment for statistical computing*. http://www.r-project.org (9 May 2026, date last accessed).

Revilla US, **Peña-Claros M**, **López-Mendoza RD**, **Meave JA**, **Bongers F. 2024**. Crown cover of a dominant pioneer legume affects tree species regeneration in a secondary tropical dry forest. *Botanical Sciences* **102**: 686–697. doi:10.17129/botsci.3382

Roesler GD, **Rodrigues J**, **Forti VA. 2023**. Bibliometric revision regarding the use of survival analysis in seed germination studies. *Ciência Rural* **53**: e20220223. doi:10.1590/0103-8478cr20220223

Romano A, **Stevanato P. 2020**. Germination data analysis by time-to-event approaches. *Plants* **9**: 617. doi:10.3390/plants9050617

Santos VS, **Duarte ML**, **Silvério S**, **Almeida LSD**, **Martins Filho S. 2025**. Do the elaiosome and seed coat color affect the seed quality of *Mabea fistulifera* Mart? *Journal of Seed Science* **47**: e202547005. doi:10.1590/2317-1545v47289194

Solarik KA, **Gravel D**, **Ameztegui A**, **Bergeron Y**, **Messier C. 2016**. Assessing tree germination resilience to global warming: a manipulative experiment using sugar maple (*Acer saccharum*). *Seed Science Research* **26**: 153–164. doi:10.1017/S0960258516000040

Downloaded from https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcag175/8713383 by Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) user on 07 August 2026