Seed Science Research

cambridge.org/ssr

Emergence speed comparison by non-linear regression and approached by time-to-event models for censored data

Thomas B. Michelon[1], Andreza C. Belniaki[1], Cesar A. Taconeli[2], Elisa S. N. Vieira[3] and Maristela Panobianco[1]

[1]Department of Plant Science, Federal University of Paraná – R, dos Funcionários, 1540, Curitiba, PR CEP 80035-050, Brazil; [2]Departament of Statistics, Federal University of Paraná – R, Evaristo F, Ferreira da Costa, 408, Curitiba, PR CEP 81530-015, Brazil and [3]Embrapa Forestry, Estrada da Ribeira, km 111, Colombo, PR CEP 83411-000, Brazil

Research Paper

Cite this article: Michelon TB, Belniaki AC, Taconeli CA, Vieira ESN, Panobianco M (2021). Emergence speed comparison by non-linear regression and approached by time-to-event models for censored data. *Seed Science Research* **31**, 319–325. https://doi.org/10.1017/S0960258521000258

Received: 19 July 2021
Revised: 14 October 2021
Accepted: 18 October 2021

Key words:
Bactris gasipaes; censoring; forestry; germination speed index; Maguire's index; non-linear regression; storage

Author for Correspondence:
Thomas B. Michelon,
E-mail: thomasnbrunomichelon@gmail.com

Abstract

Determining the germination speed is essential in experiments in the field of seed technology, as it allows the performance evaluation of a seed lot and the creation of predictive models. To this end, the literature addresses several methods and indexes. The objective of this study was to compare the main methods of emergence speed analysis in seeds, namely the non-linear regression models and the Emergence Speed Index (ESI), with the time-to-event models. The research was conducted with peach palm seeds (*Bactris gasipaes*) that were measured for viability and vigour through daily evaluations for 4 months. Vigour was evaluated by the quantification of the seed emergence speed, which was performed in three ways: ESI, non-linear regression and non-linear regression considering germination as a time-to-event event. From the results obtained, we conclude that the ESI is not a good indicator to evaluate the emergence speed; the non-linear regression model underestimates the errors and, thus, increases the probability of misclassifying treatments; the time-to-event model is more reliable in classifying treatments according to the emergence speed.

Introduction

The germination period for a seed lot means the time it takes for each seed to complete each stage of the germination process, from imbibition to radicle protrusion. However, germination is a complex physiological process, and each stage has different occurrence periods. Thus, the process as a whole presents a temporal pattern that comprises initial [latent period (lag phase)], limit fixed on total germination, and non-constant germination speed (Brown and Mayer, 1988a; McNair et al., 2012).

Determining the germination speed is essential in seed technology experiments, as it allows to evaluate the performance of a seed lot and create prediction models (Shafii and Price, 2001; Onofri et al., 2018). To this end, the literature addresses several methods and indexes. Among them, the most used is the Germination/Emergence Speed Index (ESI), proposed by Maguire (1962). Although it has been proposed as a seed vigour measure based on germination/emergence speed, criticism about its use claims that it confuses speed characteristics with viability, also they are strongly affected by the sample size (Brown and Mayer, 1986; McNair et al., 2012; Ribeiro-Oliveira and Ranal, 2016). For those inconsistencies, Pire and Varga-Simón (2019) pointed out that of the 124 papers that used this index, approximately 55% was considerably objectionable by misleading index application.

Other more recurrent methods to measure germination speed are through the fitting of non-linear regression curves. However, in addition to disregarding censored data (germination that occurs in the interval of the evaluations or seeds likely to germinate at the end of the experiment), they require normal data, which does not occur in the temporal pattern (Onofri et al., 2010; Ritz et al., 2013; Romano and Stevanato, 2020).

In this context, time-to-event analysis could be a more reliable alternative to describe the germination process and make inferences in seed analysis experiments (Onofri et al., 2011; Ritz et al., 2013), as it considers the presence of censored data and circumvents the need for normality. Thus, this study aimed to compare the main methods of emergence speed analysis in seeds, that is, the non-linear regression models and the ESI with time-to-event models.

Material and methods

We conducted the research with a peach palm (*Bactris gasipaes*) seed lot, which was collected in Nova California, Rondônia, Brazil (geographic coordinates: 09°51′06.00′ S and 64°35′52.41′ W) (Fig. 1). The seeds were treated with the systemic fungicide Carbendazim (benzimidazole)

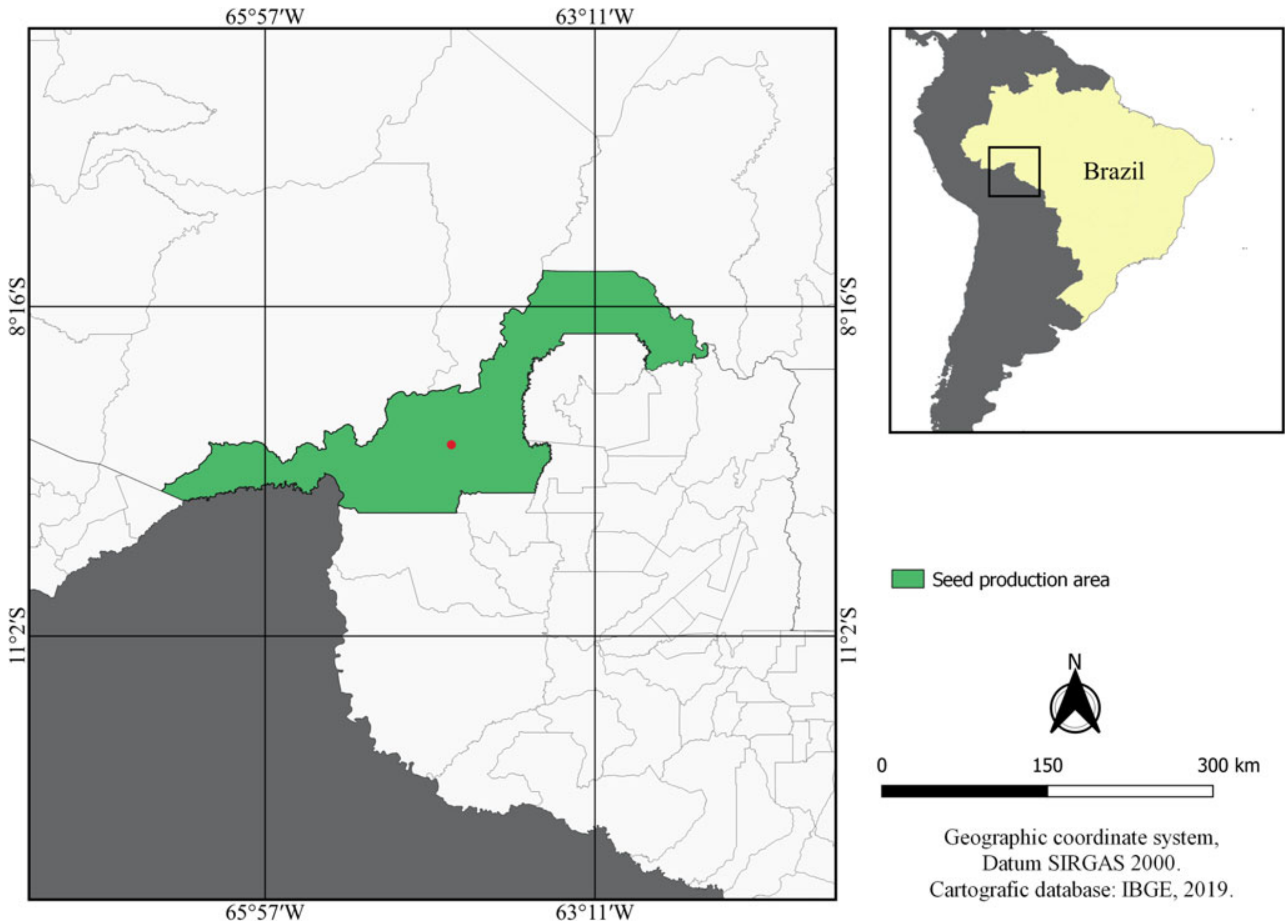

Fig. 1. Location of the peach palm seeds production municipality – Nova California, Rondônia, Brazil.

at a concentration of 100 ml/100 kg of seeds. They were then packed in polyethylene plastic bags and sent by air transport to the Universidade Federal do Paraná (UFPR) Seed Analysis Laboratory in Curitiba (Paraná, Brazil).

We manually homogenized the seeds and divided them into four subsamples with similar mass (230 seeds/package). They were stored in sealed polyethylene plastic bags (22×32 cm) with thicknesses of 0.10 and 0.20 mm at 15–17°C and relative air humidity of 58–66% (Merrow, 1991). We evaluated the viability and vigour of the seeds for 4 months, through the following tests.

Germination test (viability)

We conducted eight repetitions with 25 seeds each. They were sown at a depth equivalent to their diameter, with the fertile germ pore in contact with the substrate bed, in order to leave the other two sterile germ pores facing upwards. We used previously sterilized fine-grade vermiculite as substrate, which we placed inside plastic boxes ($17.5 \times 13.2 \times 11.5$ cm) perforated at the bottom and moistened with the equivalent volume of water to its retention capacity.

Lastly, we placed the plastic boxes in a *Mangelsdorf-type* germinator, at 25°C, under a light source. After 120 days, to evaluate the germination test, we removed the seedlings from the substrate and verified the presence of well-developed primary and adventitious roots, as well as an aerial part with a well-formed sheath. The results were expressed as a percentage of normal seedlings, and the maximum allowed tolerance for variation between repetitions was verified (Brasil, 2013).

Emergence speed (vigour)

The experimental scheme was a double factorial with additional treatment ($2 \times 4 + 1$), with the first factor being the type of package (0.10 and 0.20 mm), the second being the storage periods (1–4 months of storage) and the additional treatment as the control (unstored/unpackaged seeds). The design was completely randomized within each period and performed with eight replications.

We quantified the emergence speed in three ways: Maguire's ESI, non-linear regression and non-linear regression considering germination in a time-to-event model. For the regression models, the speed parameter was the estimated emergence rate (ER: rapidity of emergence in 1/days) for the emergence of the g percentile from the total seeds sample (ER_g). The g value was chosen based on the treatment with the lower total emergence to make sure that all treatments reach at least that percentile (e.g. ED_{50}, ED_{40} or ED_{30}).

The ESI was performed along with the germination test from daily evaluations, according to the equation:

$$\mathrm{ESI} = \frac{E_1}{N_1} + \frac{E_2}{N_2} + \cdots + \frac{E_i}{N_i}$$

https://doi.org/10.1017/S0960258521000258 Published online by Cambridge University Press

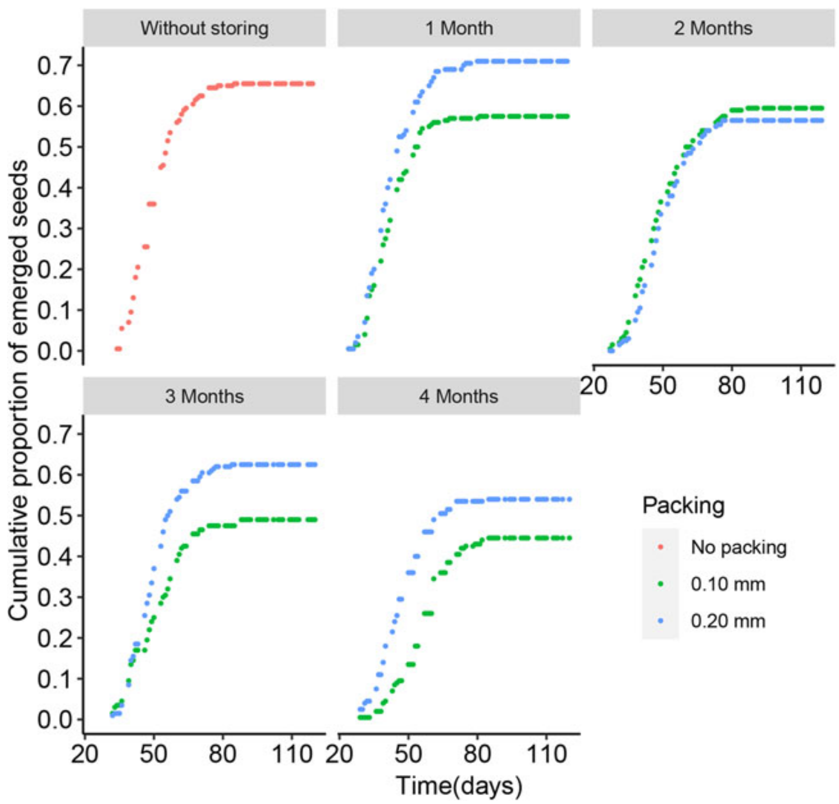

Fig. 2. Cumulative proportion of emergence in peach palm seeds without storing (control) and stored for 1–4 months, in packages of 0.10 and 0.20 mm, in relation to time.

where E_i and N_i refer, respectively, to the percentage of emergent seedlings (with a 3-mm sheath above the substrate) and the number of days elapsed from sowing until the ith count for $i = 1,2, \ldots ,n$.

To analyse the ESI, we used generalized linear models (GLMs), in which we tested three common distributions suitable for continuous response (normal, gamma and inverse normal) and different specifications for the linear predictor, based on experimental factors, to obtain the best-fit model (Olsson, 2002). We based the model selection on the Akaike Information Criterion (AIC), which is a goodness of fit measure that penalizes the model for its complexity (number of parameters). The AIC is defined by the following equation:

$$\mathrm{AIC} = -2 \log L + 2\, p$$

where p is the number of model parameters, andlog L is the logarithm value of the maximized likelihood under each model (Akaike, 1974). For the normal distribution, the normality of the residuals was verified through the Shapiro–Wilk test. Also, we used Tukey's range test to compare the mean ESI of the treatments, considering the significance level of 0.05.

We based the non-linear and time-to-event regressions on the following models: log-logistic, log-normal and Weibull, which are, respectively, given by the following equations:

$$F(t) = \frac{d}{1 + \exp\left[b\{\log(t) - \log(e)\}\right]} = \frac{d}{1 + ((t/e))^b}$$

$$F(t) = d\Phi(b(\log(x) - \log(e)))$$

$$F(t) = d \exp(-\exp(b(\log(x) - e)))$$

where $F(t)$ is the probability of emergence at the time of evaluation t_j (non-linear regression model) or at the time interval (t_{j-1}, t_j) between evaluations (time-to-event model), Φ is the normal distribution function, d is the parameter referring to the total emergence; b is the slope of the curve; e is the time (in days) elapsed to reach the 50% of the total emergence (Onofri et al., 2010; Ritz et al., 2013). For these, we tested the simplifications both in the effect of the covariates (simple effect, interaction among the factors and additional treatment) and in the parameters (one inflection point, asymptote or common slope between the curves) by the F-test. We used AIC to choose between the models fitted to each of the three distributions.

We compared each treatment's $\mathrm{ER_g}$ estimated within the models by multiple pairwise comparisons, considering a significance level of 0.05 and adjusted P-value by the false discovery rate (Benjamini and Yekutieli, 2001).

Since the experiment was conducted by seeds clustered within germination boxes, the standard error was calculated using the cluster robust sandwich standard error method to guarantee the independence of the data for all analyses in both models (Carroll et al., 1998; Yu and Peng, 2008; Ritz et al., 2013; Onofri et al., 2018).

The goodness of fit of the non-linear and time-to-event regression models were evaluated graphically by the observed *versus* predicted values (Onofri et al., 2018). All the analyses were performed in the R software, version 3.5.2.

Results and discussion

Figure 2 shows the cumulative proportions of emerged peach palm seedlings according to storage periods and packaging types. The treatments' temporal patterns varied in terms of periods for the beginning of emergence, average variation rate and stabilization period. According to Brown and Meyer (1988a), the description of the germination process must be accurate, complete, unambiguous, easy to understand and amenable to statistical analysis; thus, it is considered that expressing it through cumulative proportion meets all these requirements. In addition, the germination time course provides the capability of describing germination speed, capability and uniformity at the same time, being this, the most common way found in the literature (Bradford, 2002; O'Neill et al., 2004; McNair et al., 2012).

It is worth noting that each point in Fig. 2 represents the period of observation of the emergence and not necessarily its occurrence, which happened at some point between observations (t_{j-1}, t_j). This type of uncertainty as to the exact moment of the event's occurrence is known as interval censoring in the field of survival analysis. In turn, seedlings that did not emerge at the end of the experiment but could have emerged at some point past the experimental period are characterized as right censored (O'Neill et al., 2004; Onofri et al., 2019).

According to O'Neill et al. (2004), to understand the germination pattern it is necessary to treat the germination time as a random event that can be explained by a probability distribution. The most recurrent in growth analysis are logistics, Weibull and log-normal, and they are used according to the cumulative distribution density to fit the sigmoidal pattern of cumulative germination (Brown and Mayer, 1988b; Romano and Stevanato, 2020).

For both non-linear regression and time-to-event models, the three-parameters log-normal distribution provides lower AIC (−8616.7 and 18,808.6, respectively) in contrast to the Weibull (−8580.6 and 18,890.3) and log-logistic distribution (−8609.7 and 18,832.3). By the *F*-test test, there was a difference between the additional treatment corresponding to seeds without storage and the other treatments ($P < 0.001$), as well as a significant effect of the interaction between storage period and type of package ($P < 0.001$). Therefore, a specific model fit is necessary for each of the nine treatments examined in this study (Table 1). For the time-to-event, the saturated model (with three parameters for all the nine curves) did not differ from the model with parameter b (relative to the slope) in common to all the nine curves ($P = 0.4063$) by the *F*-test as well; hence, the reduced model was employed.

The estimated parameters were similar in both models, but there was a higher difference when it came to standard errors, with the time-to-event model presenting errors up to six times larger, even by considering the robust standard error in both models.

The difference is caused by conceptual discrepancies between non-linear regression and time-to-event models. In non-linear regression, independent errors and homogeneous variance are assumed; however, the cumulative proportion of the emergence over time violates these assumptions. Since the emergence on t_j day is a sum of the emergence up to that particular moment, the errors become dependent; in the time-to-event model, on the other hand, this problem is solved once it considers only the seedlings that emerged between evaluation periods. Variance is not constant throughout the experimental period, since there is a lower increase in the proportion of emerged seedlings at the beginning and end of the experiment, resulting in a lower variance when compared to the intermediate period (Shafii and Price, 2001; Crane et al., 2002; McNair et al., 2012).

In addition to the conceptual divergences in the development of the non-linear regression and time-to-event models, the method of parameter estimation through least squares (Bates and Watts, 1988) considers that the emergence occurred at the moment of evaluation, which does not happen in practice. In the time-to-event model, the parameters are estimated by the maximum likelihood method (McCullagh and Nelder, 1989); thus, the presence of censored information is considered (Onofri et al., 2011; Ritz et al., 2015).

Table 1. Estimated parameters and robust standard errors for the non-linear and time-to-event regression model, both with three-parameter log-normal distribution, of peach palm seeds without storage (control) and stored for 1–4 months, in 0.10- and 0.20-mm packages

Treatment	Parameter	NLM estimate	Time-to-event estimate
Without storing	B	5.028 (0.326)	4.426 (0.002)
	D	0.653 (0.008)	0.655 (0.009)
	E	48.322 (0.451)	48.638 (1.027)
1 month 0.10 mm	B	4.331 (0.212)	4.426 (0.002)
	D	0.575 (0.005)	0.575 (0.123)
	e	40.497 (0.370)	40.730 (2.966)
1 month 0.20 mm	b	4.016 (0.164)	4.426 (0.002)
	d	0.708 (0.004)	0.710 (0.009)
	e	40.09 (0.300)	40.625 (3.847)
2 months 0.10 mm	b	3.754 (0.262)	4.426 (0.002)
	d	0.595 (0.009)	0.595 (0.108)
	e	46.478 (0.704)	46.783 (1.243)
2 months 0.20 mm	b	4.469 (0.171)	4.426 (0.002)
	d	0.566 (0.005)	0.565 (0.114)
	e	48.044 (0.316)	48.096 (0.964)
3 months 0.10 mm	b	3.882 (0.249)	4.426 (0.002)
	d	0.493 (0.009)	0.490 (0.137)
	e	49.148 (0.685)	49.014 (1.131)
3 months 0.20 mm	b	4.893 (0.222)	4.426 (0.002)
	d	0.623 (0.006)	0.625 (0.097)
	e	47.823 (0.319)	48.162 (1.064)
4 months 0.10 mm	b	4.858 (0.357)	4.426 (0.002)
	d	0.448 (0.009)	0.445 (0.148)
	e	55.285 (0.587)	54.352 (1.403)
4 months 0.20 mm	b	4.140 (0.161)	4.426 (0.002)
	d	0.542 (0.004)	0.54 (0.122)
	e	45.571 (0.335)	45.306 (1.465)

Comparing both methods of parameter estimation, Ritz et al. (2013) demonstrated that, in seed analysis experiments, estimation by the maximum likelihood method will always produce higher standard errors, and the discrepancy will be higher according to the distribution of events through evaluation intervals. Thus, although non-linear regression models provide a good parameter estimation, they also provide underestimated standard errors that do not match experimental reality (O'Neill et al., 2004).

Due to the violation of assumptions, the underestimation of standard errors leads to problems in covering confidence intervals, as well as in the risk of type I error for hypothesis testing. This may result in erroneous inferences when comparing treatments. Moreover, in experiments with longer intervals between analyses, the standard error underestimation can be substantial (Onofri et al., 2011; Ritz et al., 2013; Onofri et al., 2018).

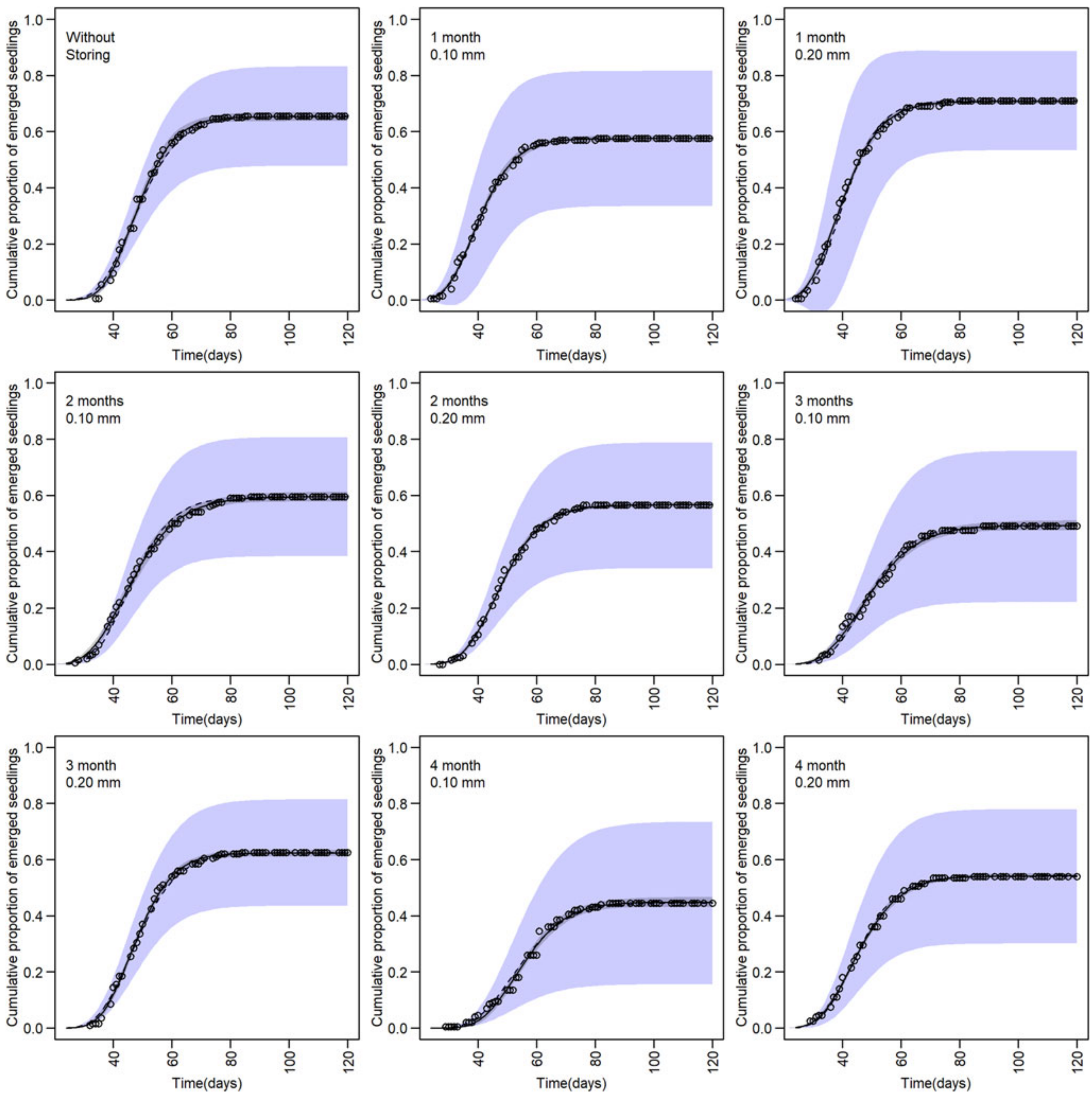

Fig. 3. Cumulative proportion of peach palm seedlings emergence in relation to time adjusted to non-linear regression (dotted line) and time-to-event (continuous line) models, both with three-parameter log-normal distribution, for seeds without storing (control) and stored for 1–4 months in 0.10- and 0.20-mm packages. The 95% confidence interval of NLR, TTT models are displayed in grey and blue, respectively. The circles represent the observed values.

In Fig. 3, it is possible to observe that the log-normal distribution was well fitted to the data, both in non-linear regression and time-to-event analysis, by comparison to the observed *versus* adjusted data (Bradburn et al., 2003; Dey and Kundu, 2010). The similarity between estimated parameters with different standard errors for non-linear regression and time-to-event models is graphically reflected in close-fitting curves, but with discrepant confidence bands.

For each curve, the respective *x*-axis value for inflection point (i.e. middle of the curve and the parameter *e* from the equation) which can be interpreted as the time to reach 50% of the total emergence (T50) does not necessarily coincide with the 50th percentile of the total number of seeds and the T50 based on the total number of seeds is the common parameter for germination/emergence speed in seed science (Soltani et al., 2015).

The emergence of a seed lot can be deemed faster when the latent period and/or the variation rate is lower; such characteristics can be observed in the temporal patterns with the period for the beginning of emergence and sigmoid slope, respectively. A challenge in conveying the germination/emergence speed is breaking down this single characteristic and conveying it in a measure that is not influenced by the others from the temporal pattern, so that it may be possible to compare two seed lots with completely different patterns and for this T50 can be an

Table 2. Speed comparison by estimated time to reach 40% of total seeds (T40), robust standard error (in brackets) and confidence interval of 95% (bellow) for the non-linear and time-to-event regression model and the ESI (Maguire's index) for peach palm seeds without storage (additional/control treatment) and stored for 1–4 months in 0.10- and 0.20-mm packages

Treatment	Time-to-event ER40* estimated	Non-linear regression ER40 estimated	Maguire's ESI index estimated
1 month 0.20 mm	0.024a	0.024a	1.756a
	(0.019; 0.028)	(0.024; 0.024)	(1.58; 1.936)
1 month 0.10 mm	0.022a	0.022b	1.412ab
	(0.019; 0.025)	(0.021; 0.022)	(1.236; 1.592)
2 months 0.10 mm	0.019a	0.019cd	1.288b
	(0.018; 0.02)	(0.018; 0.02)	(1.108; 1.464)
Without storing	0.019a	0.02c	1.348ab
	(0.018; 0.02)	(0.019; 0.02)	(1.168; 1.528)
3 months 0.20 mm	0.019a	0.019c	1.296b
	(0.018; 0.02)	(0.019; 0.02)	(1.12; 1.476)
4 months 0.20 mm	0.019a	0.019cd	1.196bc
	(0.018; 0.02)	(0.018; 0.019)	(1.016; 1.376)
2 months 0.20 mm	0.018a	0.018d	1.18bc
	(0.018; 0.019)	(0.018; 0.019)	(1; 1.356)
3 months 0.10 mm	0.017b	0.016e	1.008bc
	(0.016; 0.017)	(0.015; 0.017)	(0.828; 1.188)
4 months 0.10 mm	0.014c	0.014f	0.206c
	(0.013; 0.014)	(0.013; 0.015)	(0.161; 0.25)

*Lower case letters represent a comparison of the parameters in the column, considering a significance level of 0.05.

easily used, as well its inverse – emergence rate (ER50), to directly measure the speed. (Shafii et al., 1991; Gardarin et al., 2011; Soltani et al., 2015). As it is based on the total number of seeds, for treatments that do not reach at least 50% of emergence, other percentiles can be used (e.g. T20 and T40) (Soltani et al., 2015).

The estimative velocity to emerge 40% (ER40) of total seeds both in non-linear regression and time-event (Table 2) was similar; nevertheless, while the seeds stored for 1 month in 0.20-mm package displayed the fastest emergence according to the non-linear regression, for the time-event model, they did not show a significant difference from those stored in 0.10-mm packaging for 1 and 2 months, 0.20 mm for 2, 3 and 4 months as well the non-stored ones. Which corroborates the fact that non-linear regression produces smaller standard so tends to the alternative hypothesis; for example, for this model the treatments were classified in six classes, while for the time-event there were just three.

As in the regression analysis, the ESI values were significant for the interaction between storage period and packaging type (P-value = 0.036); however, they were not significant for the additional treatment (P-value = 0.290). It is worth highlighting that the non-significance occurs only between the additional treatment and the average of the interaction, which does not imply equality between the additional treatment and each level of interaction.

For the ESI contrasts (Table 2), the seeds stored for 1 month and the non-stored ones showed the highest emergence speed, similar to the time-event model, but while the seeds stored in 0.20-mm packaging for 4 months did not present a significant difference for the fastest treatment (i.e. 0.10 mm for 1 month) according to time-event model, for the ESI they did not differ from the seeds stored in 0.10-mm packaging for 4 months, this treatment being considered as the slower by the time model. Despite the differences in slope and latent period in the treatments' temporal patterns between the packages of seeds stored for 4 months, Maguire's index shows similar values for those. Such behaviour was characterized by Brown and Mayer (1986) as an anomalous behaviour of the Maguire index, in which different temporal patterns can present similar speed indexes.

A criticism regarding the ESI is related to the correlation with the final emergence; that is, the index combines questions of viability and emergence speed. This was also observed in our study where the ESI and the emergence presented a linear correlation coefficient of 0.93 ($P < 0.001$), thus ratifying the criticism regarding the ambiguity of the measure. It is worth noting that the correlation between the ESI and the total emergence is not always present in the same proportion in experiments with seeds, so care should be taken in their application (Throneberry and Smith, 1955; Brown and Meyer, 1986; Shafii and Price, 2001; Kader, 2005; Ranal and Santana, 2006; McNair et al., 2012).

The ESI was proposed as a measure of seed vigour based on the seed lot emergence speed. Given that vigour is the sum of the properties of a seed that determine its level of activity and performance during germination/emergence (ISTA, 2020), the ambiguity between the computation of the properties of speed and the total germination may justify its popularization. For McNair et al. (2012), computing the speed and germination range in conjunction is not a problem in seed technology, as it helps to discriminate between more and less vigorous lots.

By this, it is worth to emphasize that there are several methods to discriminate seed lot speed (e.g. germination/emergence indexes), the literature is vast about them and by the power of simplicity, one is encouraged to use them. But when considering to evaluate the emergence speed by estimating the curve parameters, which are important for prediction models, and which the hydrothermal theory is based on, the approach by the classical non-linear regression may lead to substantial inferential errors, and by that, the choice of time-to-event can be justified (Bradford, 2002).

Conclusion

It is important to be aware that the choice of evaluation method may drastically influence the results; thus, the experimental design must critically evaluate the choice of the data analysis method. Therefore, it is possible to conclude that the ESI is not a good method to evaluate emergence speed, because it mixes characteristics of the temporal pattern such as emergence speed and total emergence. On the other hand, the adjustment of the accumulated emergence by non-linear regression models underestimates the errors and increases the probability of misclassifying treatments, while the time-to-event models are more reliable in classifying treatments according to the emergence speed.

Acknowledgements. The author(s) would like to thank the company Klabin for donating the seeds to the experiment and the Academic Publishing Advisory Center (Centro de Assessoria de Publicação Acadêmica, CAPA – www.capa.ufpr.br) of the Federal University of Paraná (UFPR) for assistance with English language translation and editing. The authors thank the

Coordination for the Improvement of Higher Education Personnel (CAPES) for the granting of scholarships.

Authors' contributions. T.B.M. is involved in the conceptualization. A.C.B. and E.S.N.V. are involved in the data acquisition. T.B.M. is involved in the data analysis. T.B.M., C.A.T., M.P., E.S.N.V. and A.C.B. are involved in the design of methodology. T.B.M., C.A.T., M.P. and A.C.B. are involved in the writing and editing.

Conflicts of interest. None declared.

References

Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723.
Bates D and Watts D (1988) *Nonlinear regression and its applications*. New York, NY, USA, John Wiley and Sons Inc.
Benjamini Y and Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29**, 1165–1188.
Bradburn MJ, Clark TG, Love SB and Altman DG (2003) Survival analysis part III: multivariate data analysis – choosing a model and assessing its adequacy and fit. *British Journal of Cancer* **89**, 605–611.
Bradford K (2002) Applications of hydrothermal time to quantifying and modeling seed germination and dormancy. *Weed Science* **50**, 248–260.
Brasil (2013) *Instruções para Análise de Sementes de Espécies Florestais*. Brasília, DF, Brazil, Ministério da Agricultura Pecuária e Abastecimento (in Portuguese).
Brown RF and Mayer DG (1986) A critical analysis of Maguire's germination rate index. *Journal of Seed Technology* **10**, 101–110.
Brown RF and Mayer DG (1988a) Representing cumulative germination. 1. A critical analysis of single-value germination indices. *Annals of Botany* **61**, 117–125.
Brown RF and Mayer DG (1988b) Representing cumulative germination. 2. The use of the Weibull function and other empirically derived curves. *Annals of Botany* **61**, 127–138.
Carroll RJ, Wang S, Simpson DG, Stromberg AJ and Ruppert D (1998) The sandwich (robust covariance matrix) estimator. Technical Report, Department of Statistics. *A & M*, pp. 1–16.
Crane M, Newman MC, Chapman FP and Felon J (2002) *Risk assessment with time to event models*. New York, USA, Lewis Publishers.
Dey AK and Kundu D (2010) Discriminating between the log-normal and log-logistic distributions. *Communications in Statistics – Theory and Methods* **39**, 280–292.
Gardarin A, Dürr C and Colbach N (2011) Prediction of germination rates of weed species: relationships between germination speed parameters and species traits. *Ecological Modelling* **222**, 626–636.
ISTA – The International Seed Testing Association (2020) *International rules for seed testing*. Bassersdorf, CH, International Rules for Seed Testing.
Kader MA (2005) A comparison of seed germination calculation formulae and the associated interpretation of resulting data. *Journal & Proceeding of the Royal Society of New South Wales* **138**, 65–75.
Maguire JD (1962) Speed of germination – aid in selection and evolution for seedling emergence and vigor. *Crop Science* **2**, 176–177.
Mccullagh P and Nelder JA (1989) *Generalized linear models* (2nd edn). Florida, USA, Chapman & Hall.
Mcnair JN, Sunkara A and Frobish D (2012) How to analyze seed germination data using statistical time-to-event analysis: non-parametric and semi-parametric methods. *Seed Science Research* **22**, 77–95.
Meerow AW and Broschat TK (1991) Palm seed germination. *University of Florida, Bulletin* **274**, 1–9.
Olsson U (2002) *Generalized linear models: an applied approach*. Lund, Sweden, Studentlitteratur.
O'Neill ME, Thomson PC, Jacobs BC, Brain P, Butler RC, Turner H and Mitaka B (2004) Fitting and comparing seed germination models with a focus on the inverse normal distribution. *Australian and New Zealand Journal of Statistics* **46**, 349–366.
Onofri A, Benicasa P, Mesgaran MB and Ritz C (2018) Hydrothermal-time-to-event models for seed germination. *European Journal of Agronomy* **101**, 129–139.
Onofri A, Gresta F and Tei F (2010) A new method for the analysis of germination and emergence data of weed species. *Weed Research* **50**, 187–198.
Onofri A, Mesgaran MB, Tei F and Cousens RD (2011) The cure model: an improved way to describe seed germination? *Weed Research* **51**, 516–524.
Onofri A, Piepho HP and Kozak M (2019) Analyzing censored data in agricultural research: a review with examples and software tips. *Annals of Applied Biology* **174**, 3–13.
Pire R and Vargas-Simón G (2019) Recurrent inconsistencies in publications that involve Maguire's germination rate formula. *Forest Systems* **28**, 1–5.
Ranal MA and Santana DG (2006) How and why to measure the germination process? *Revista Brasileira de Botanica* **29**, 1–11.
Ribeiro-Oliveira JP and Ranal MA (2016) Sample size in studies on the germination process. *Botany* **94**, 103–115.
Ritz C, Baty F, Streibig JC and Gerhard D (2015) Dose-response analysis using R. *PLoS ONE* **10**, 1–13.
Ritz C, Pipper CB and Streibig JC (2013) Analysis of germination data from agricultural experiments. *European Journal of Agronomy* **45**, 1–6.
Romano A and Stevanato P (2020) Germination data analysis by time-to-event approaches. *Plants* **9**, 2–5.
Shafii B and Price WJ (2001) Estimation of cardinal temperatures in germination data analysis. *Journal of Agricultural, Biological, and Environmental Statistics* **6**, 356–366.
Shafii B, Price WJ, Swensen JB and Murray GA (1991) Nonlinear estimation of growth curve models for germination data analysis. *Conference on Applied Statistics in Agriculture* **3**, 19–36.
Soltani E, Ghaderi-Far F, Baskin CC and Baskin JM (2015) Problems with using mean germination time to calculate rate of seed germination. *Australian Journal of Botany* **63**, 631–635.
Throneberry GO and Smith FG (1955) Relation of respiratory and enzymatic activity to corn seed viability. *Plant Physiology* **30**, 337–344.
Yu B and Peng Y (2008) Mixture cure models for multivariate survival data. *Computational Statistics & Data Analysis* **52**, 1524–1532.