*Seed Science Research*

**cambridge.org/ssr**

## Technical Update

**Cite this article:** de Faria RQ, dos Santos ARP, Amorim DJ, Cantão RF, da Silva EAA, Sartori MMP (2020). Probit or Logit? Which is the better model to predict the longevity of seeds? *Seed Science Research* **30**, 49–58. https://doi.org/10.1017/S0960258520000136

Received: 23 September 2019
Accepted: 30 April 2020

**Key words:**
central limit theorem; link functions; residual normality; seed conservation; seed viability; soybean

***Author for Correspondence:**
Rute Q. de Faria, E-mail: rute.faria@ifgoiano.edu.br

© The Author(s), 2020. Published by Cambridge University Press

CAMBRIDGE UNIVERSITY PRESS

# Probit or Logit? Which is the better model to predict the longevity of seeds?

Rute Q. de Faria[1]*, Amanda R. P. dos Santos[2], Deoclecio J. Amorim[2], Renato F. Cantão[3], Edvaldo A. A. da Silva[2] and Maria M. P. Sartori[2]

[1]Department of Agricultural Engineering, Instituto Federal de Educação Ciência e Tecnologia Goiano, Campus Urutaí, Rod. Geraldo Silva Nascimento, Km-2,5, Zona Rural, Urutaí - GO 75790-000, Brazil; [2]Department of Production and Plant Breeding, School of Agriculture, Universidade Estadual Paulista (UNESP), Botucatu. Av. Universitária, n° 3780 - Altos do Paraíso, Botucatu, São Paulo 18610-034, Brazil and [3]Universidade Federal de São Carlos, Campus de Sorocaba (UFSCar), Rodovia João Leme dos Santos (SP-264), Km 110, Bairro do Itinga, Sorocaba, São Paulo 18052-780, Brazil

**Abstract**

The prediction of seed longevity (P50) is traditionally performed by the use of the Probit model. However, due to the fact that the survival data are of binary origin (0,1), the fit of the model can be compromised by the non-normality of the residues. Consequently, this leads to prediction losses, despite the data being partially smoothed by Probit and Logit models. A possibility to reduce the effect of non-normality of the data would be to apply the principles of the central limit theorem, which states that non-normal residues tend to be normal as the *n* sample is increased. The Logit and Probit models differ in their normal and logistic distribution. Therefore, we developed a new estimation procedure by using a small increase of the *n* sample and tested it in the Probit and Logit functions to improve the prediction of P50. The results showed that the calculation of P50 by increasing the *n* samples from 4 to 6 replicates improved the index of correctness of the prediction. The Logit model presented better performance when compared with the Probit model, indicating that the estimation of P50 is more adequate when the adjustment of the data is performed by the Logit function.

## Introduction

Seed longevity is important for the *ex situ* conservation and the survival of species in their natural environment. Longevity is a parameter closely linked to the physiological quality of the seeds and can be defined as the time required for the seed to reduce a sigma in viability during storage. Longevity includes vigour, since it can be reduced during the period of storage, being a key lever to increase crop production (Pereira Lima et al., 2017). The estimated seed longevity is performed by calculating P50, which indicates the time period in which a given batch loses 50% of its viability.

The predictive models of P50 are important tools in the management of seeds during *ex situ* conservation. Therefore, assertively defining this index can reduce analysis costs and facilitate commercial planning. In seed banks, prediction models are tools that can facilitate the estimation of seed viability. In these environments, the amount of seeds available to test viability is not always sufficient, comprising seed conservation. Thus, models to better predict seed longevity during storage may improve conservation strategies (Walters, 2015).

The data generated by the survival curve of a given seed lot have a sigmoidal behaviour. The Probit model is given by the inverse function of the cumulative distribution curve of the probability of survival of the seeds. Ellis and Roberts (1981) proposed this model to define the P50 in orthodox seeds, and since then, the Probit model has been proposed to predict longevity in seeds of different species and in different storage conditions (Ellis et al., 1990; Demir and Mavi, 2003; Obunyali et al., 2008; Marcondes et al., 2011).

Seed survival analysis data are binary responses, that is, the dependent variable ($Y$) is dichotomous, and therefore admits only two results. A non-viable seed is considered dead, and a viable seed is considered alive, being categorized as zero ($Y = 0$) or one ($Y = 1$), respectively. A binary response variable does not meet the assumptions of normality of the residuals necessary for an adequate adjustment of a symmetric function. Oral (2006) indicates that it is important to use appropriate non-normal distribution to obtain efficient estimators in cases of non-normality of the probability density function.

The models of Probit and Logit are options for the adjustment of dichotomous variables, such as binary variables, whose parameter estimations are performed using the ordinary least squares (OLS) adjustment method. The OLS method assumes the functions of probability distribution, which, in the case of Logit and Probit, would be logistic and normal, respectively. This is because the functions of Probit and Logit are symmetric, that is, they are equally

distributed around the mean of distribution. However, seed survival data have heavy tails that lead to the departure from normality, compromising the quality of prediction (Schützenmeister et al., 2012).

The Logit model is considered to be the most important for categorical variable data (Agresti, 2013). If compared to Probit, it is also mathematically simpler. The main difference between these two functions is due to the forms of the distribution curves that each one represents. Although both have dichotomous-dependent variables, the Logit and Probit linking functions are given by the inverse of the logistic and normal accumulated distributions, respectively. Thus, it is important to evaluate the behaviour of the probability distribution of each model, in order to establish the function that best represents the phenomenon.

However, because they are similar, the functions of Probit and Logit would require other adjustments in order to bring them to a condition closer to normality. The central limit theorem (CLT) proposes that a binomial variable can be approximated to normal when the $n$ of the sample is sufficiently large. The theorem considers that, being $X$ a random variable, with a mean $\mu$ and a standard deviation $\sigma$, when we withdraw a sample of size $n$ of $X$, where $\bar{X}$ is the average of the sample, when $n$ increases, we will have equation (1):

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma}} \approx N(0, 1) \quad (1)$$

That is, if the sample size is sufficiently large, we can assume that the sample mean has a normal distribution. In addition, the theorem indicates that with any distribution function, which has finite variance, the sample mean will have approximately normal distribution as the sample $n$ increases (Bobkov et al., 2014)

In this study, the Logit function will be tested for its robustness in predicting P50 in seeds, compared with the Probit model. A greater number of replicates of the sample data, as proposed by the CLT, will be evaluated for their capacity of normalization of the residues and consequent improvement in the prediction of the models.

## Material and methods

### Seed production and physiology

Soybean plants (*Glycine max* L.) were grown under natural conditions using standard planting and culture techniques in different regions of Brazil during the harvest season 2016/2017. The seeds were mechanically harvested and stored in bags. In April 2017, samples of approximately 1 kg were taken from 15 seed lots.

The literature does not specify a minimum number of cultivars necessary to fit the modelling. In this study, we sought to obtain as many samples as possible that would bring diversity and variability to the analysis of the models. Thus, cultivars adapted to different conditions of temperature and relative humidity produced in the central, southern and northern regions of the country were tested, as well as different levels of physiological quality.

The lots S01–S10, S12 and S15 were seeds produced in the central region of Brazil, Goiás Province. The cultivar S11 was produced in Paraná Province, southern region, and the lots S13 and S14 were seeds produced in Maranhão Province, northern region of Brazil. The lots are genetically modified cultivars, except for lots S11, S12 and S14. After carrying out the analyses in a completely randomized design (DIC), Fisher's LSD test was used at the level of 5% probability of error ($P \leq 0.05$) to differentiate the different levels of the vigour of the lots in this study. The cultivars were coded, and their regions of origin are shown in Table 1.

**Table 1.** Coding of soybean cultivars with their respective production region

| Code | Cultivars | Origin region |
|---|---|---|
| S01 | DS5916 IPRO | Centre |
| S02 | DS59716 IPRO | Centre |
| S03 | CD2591 IPRO (Lot 01) | Centre |
| S04 | CD2591 IPRO (Lot 2) | Centre |
| S05 | CD2737 RR (Lot 1) | Centre |
| S06 | CD2737 RR (Lot 2) | Centre |
| S07 | CD251 RR | Centre |
| S08 | CD2820 IPRO (Lot 1) | Centre |
| S09 | CD2820 IPRO (Lot 2) | Centre |
| S10 | DESAFIO (Lot 1) | Centre |
| S11 | BRS 284 | South |
| S12 | BRS 8581 | Centre |
| S13 | PARAGOMINAS | North |
| S14 | BRS333 | North |
| S15 | DESAFIO (Lot 2) | Centre |

In this study, wild cultivars were not addressed. It is known that these species naturally have a less homogeneous behaviour, which can directly interfere in the adjustment of the model. Despite this, the approach of this study focused on the assessment of the robustness of the models regarding the normality of the residues. Thus, the model that is more robust with agricultural cultivars, therefore more homogeneous, most likely will also have comparably better performance in material with less homogeneity.

Following this, the seed lots were placed in paper bags and stored at 12°C and 50% relative humidity. The water content of the whole seed was determined gravimetrically in two replications of six seeds by oven drying at 105°C for 24 h (ISTA, 1985) and expressed as grams of water per gram of dry weight (DW).

For the germination test, 6 replicates of 50 seeds were imbibed in moistened paper rolls at a constant temperature of 25°C in BOD (biochemical oxygen demand) chambers. The seeds were considered germinated when the protruded radicle showed a length of 2 mm or more. In addition, the first count of normal plants, 5 d after the germination test, was also used as a vigour test according to ISTA (1985).

The traditional accelerated ageing test was performed in individual germination boxes with the seeds suspended on wire meshes inside as an even layer of seeds. Within each individual compartment, 40 ml of distilled water was added and the boxes were placed in the BOD chambers at 45°C for 72 h. After this period, the seeds went through the same germination test as described above. The germination percentage was recorded after 5 d. The formation of a normal seedling was used as a criterion for germination (McDonald et al., 2008).

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

For the determination of longevity, the seeds were placed in a controlled environment with 75% RH (relative humidity) at 35°C using closed boxes containing a saturated solution of NaCl. Every 7 d, one of the boxes was removed to perform the germination test. Longevity was expressed as P50, defined as the time (days) in which the seeds lose 50% of viability (Ellis and Roberts, 1981). The equilibrium of water content was measured after the first 7 d of the longevity experiment. The water content was determined gravimetrically as described above and expressed as grams of water per gram DW.

### Fitting of the models

The survival curves were set up with the data obtained from the probability of germination during the time of longevity analysis. For transformation of the survival curve in the Probit and Logit linear scales, equations (2) and (3) were used, respectively Chen and Tsurumi (2011).

$$\text{Probit} \quad F(x) = \Phi^{-1} \tag{2}$$

$$\text{Logit} \quad F(x) = \ln\left(\frac{x}{1-x}\right) \tag{3}$$

where $\Phi$ is the normal density curve and $x$ is the probability of germination transformed into equivalent percentage indices ranging from 0 to 1. The survival data transformed into Probit and Logit linear scales represent the inverse of the normal and logistic accumulated distributions, respectively. The range between 5 at 95% of the survival data was select to fit the models.

The P50 index was obtained by the linear equation after models fitted, when Probit and Logit scales are equal to 0, following equations (4) and (5):

$$Y = \beta - \alpha X \tag{4}$$

$$X_{P50} = \frac{\beta}{\alpha} \tag{5}$$

where $Y$ is the Probit or Logit percentage viability, $X$ is the storage period (days), $\alpha$ is the slope of the model adjusted, $\beta$ is the Probit or Logit percentage viability at the beginning of the storage and is the time period (days) that a seed lot to reduce its viability by up to 50% (P50) (Ellis and Roberts, 1981).

### Analysis of residue normality

The normality evaluation was carried out from the behaviour of the residues after the models were fitted. The residue analysis was performed using the tests of Shapiro–Wilk (S–W), Liliefors and Komogorov–Smirnov. According to Thode (2002), S–W is the most recommended test for the evaluation of fitted residues of the models; however, the author emphasizes that in smaller samples, the test of Komogorov–Smirnov could have greater sensitivity. Thus, the choice of the three tests allows a broader assessment of this analysis, and the Liliefors test is used for a tie-break criterion for non-normality of residues.

As the evaluation by the tests alone can be subjective, the graphical evaluation through the straightline of the normal Q–Q plot was also used as a complementary evaluation of the goodness of fit of the Probit and Logit models to the data. The analysis of normality through graphs and the application of normality tests are more accurate and less subjective, allowing a better interpretation of the residues.

### Proposal for model adjustment – CLT

With the expectation of bringing the response variable to a normality condition, this study applied the principles of the CLT. The theorem states that non-normal data tend to normality when the value in the sampled $n$ increases, and this occurs regardless of the type of distribution.

According to ISTA (1985), the recommendation of the $n$ sample in seed analysis is at least four replications, and this is the value commonly adopted. In this study, we suggested an increase of two more repetitions, that is, six replicates in all. This new value of $n$ aims to evaluate if the increment of two more replicates of the $n$ sample would be relevant for the improvement of the prediction by the models. The analysis of the relevance of sample increase was performed according to the proximity of the P50 obtained, with the interval of days in which the P50 was experimentally verified and with the residue analysis.

### Goodness of fit of the models

In order to evaluate the goodness of fit of each model, we considered the coefficient of determination $R^2$, $R_{adjusted}$ and the coefficient Pearson's $r$. Statistical analyses of this study were performed using statistical software R 3.3.3.

## Results

During the storage period under conditions of 75% RH and 35°C, the water content was measured periodically. The initial water content was relatively low with indices ranging from 6.84 to 11.74%. The equilibrium water (EW) content was measured after a period of 7 d of storage and again at 14 d of storage for confirmation. After conditioning the seeds at 75% RH and 35° C, there was an increase in water content, which after 7 d was in equilibrium with indices varying from 13.95 to 15.95%.

Germination displayed radicle emission indices between 90 and 100%, when only radicle emission was considered. The vigour analysis shows the differences between soybean lots (Table 2). The vigour evaluations, therefore, show that the selected seed group has a good representation of the differences naturally found among commercial seeds, but it can be inferred that the lots are subdivided into different levels of vigour. These differences are peculiar and enrich the study, since there is a good diversity of physiological quality among the lots evaluated. Therefore, the evaluation of the models proposed by this study had a significant amplitude of analysis.

During the longevity test, the seeds were evaluated as their viability every 7 d. That is, each week, the germination analysis was performed, and the expected interval (EI) indicates the time period, in days, in which the lots had a 50% germination index. The EI, therefore, is one of the main parameters in which we will first evaluate the prediction quality of the P50 determined by the Probit and Logit models.

In Table 3, the EI results of each cultivar are shown, and the P50 is calculated by the traditional methodology, using four replicates, followed by the P50 calculated with six replicates. The

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

**Table 2.** Characterization of the physiological quality of soybean seed lots, with their respective percentage of normal seedlings of the general germination index, of the first count of germination test (FC) and accelerated ageing (AA)

| Lot code | Germination (%)* Mean (SE) range | FC (%)* Mean (SE) range | AA (%)* Mean (SE) range |
|---|---|---|---|
| S10 | 100$^{a}$ (0.00) | 21.03$^{ef}$ (14.08) | 95.5$^{a}$ (4.43) |
| | 97.95–102.04 | 12.82–29.25 | 89.14–101.85 |
| S12 | 99.66$^{ab}$ (0.81) | 62.01$^{bc}$ (5.37) | 62.0$^{fgh}$ (7.12) |
| | 97.62–101.71 | 53.80–70.23 | 55.64–68.35 |
| S02 | 99.33$^{abc}$ (1.03) | 73.33$^{ab}$ (3.72) | 82.5$^{abcd}$ (3.41) |
| | 97.29–101.38 | 65.12–81.54 | 76.14–88.85 |
| S06 | 98.99$^{abc}$ (1.10) | 71.91$^{ab}$ (7.77) | 71.5$^{def}$ (6.61) |
| | 96.94–101.03 | 63.69–80.12 | 65.14–77.85 |
| S05 | 98.00$^{abcd}$ (1.26) | 72.10$^{ab}$ (6.12) | 93.5$^{ab}$ (3.78) |
| | 95.95–100.05 | 63.88–80.32 | 87.14–99.85 |
| S01 | 97.66$^{abcd}$ (0.61) | 26.0$^{de}$ (21.05) | 86.0$^{abcd}$ (5.16) |
| | 95.61–99.71 | 17.78–34.21 | 79.64–92.35 |
| S14 | 97.66$^{abcd}$ (0.615) | 4.7 $^{f}$ (7.76) | 37.16$^{i}$ (7.43) |
| | 95.62–99.71 | -3.54–12.88 | 30.80–43.51 |
| S13 | 96.97$^{abcd}$ (0.615) | 43.96$^{cd}$ (11.84) | 71$^{def}$ (5.29) |
| | 94.93–99.02 | 35.75–52.18 | 64.64–77.35 |
| S15 | 95.33$^{abcde}$ (3.26) | 15.14$^{ef}$ (17.43) | 12.0$^{j}$ (4.89) |
| | 93.28–97.38 | 6.92–23.35 | 5.64–18.35 |
| S08 | 94.66$^{bcdef}$ (0.843) | 85.0$^{a}$ (4.52) | 76.00$^{cdef}$ (4.32) |
| | 92.62–96.71 | 76.78–93.21 | 69.64–82.35 |
| S11 | 94.30$^{cdef}$(3.49) | 79.93$^{ab}$(6.07) | 90.0$^{abc}$ (7.12) |
| | 92.25–96.34 | 71.71–88.15 | 83.64–96.35 |
| S03 | 93.67$^{def}$ (1.20) | 83.33$^{a}$ (7.66) | 49.00$^{hi}$ (14.09) |
| | 91.62–95.71 | 75.12–91.54 | 42.64–55.35 |
| S09 | 93.00$^{def}$ (2.05) | 80.7$^{ab}$ (6.89) | 79.42$^{bcde}$ (4.32) |
| | 90.95–95.04 | 72.45–88.88 | 73.06–85.77 |
| S04 | 91.00$^{ef}$ (1.84) | 76.7$^{ab}$ (5.46) | 50.5$^{ghi}$ (1.91) |
| | 88.95–93.04 | 68.45–84.88 | 44.14–56.85 |
| S07 | 90.09$^{f}$ (0.52) | 73.99$^{ab}$ (7.11) | 65.5$^{efg}$ (5.74) |
| | 88.05–92.14 | 65.77–82.20 | 59.14–71.85 |

*The averages followed by the same letter do not differ statistically between lots, by the Fisher's LSD test at the 5% probability level of error ($P \leq 0.05$).

values we found for the P50 range from 16 to 84 d for the condition of 75% RH and 35°C.

Fischer (2011), describing the principles of CLT, proved that the increase in the *n* sample is efficient to lead data distribution to normality, which of course could improve the goodness of fit of the models. However, it is limited due to the necessity of a high number of seeds and increased analysis costs. The P50 with the increase of the sample *n*, tested and presented in Table 3, showed a lower value for standard error in most samples analysed with six repetitions. The increase in the *n* sample was also necessary for a better fit in both models. This was observed for lots S03, S5, S10, S11 and S14, which presented values outside the interval, but with an increment of only two repetitions, it was already possible to bring the calculated P50 to the expected range.

As previously explained, the assumptions of the functions of Probit and Logit are similar. Therefore, to differentiate how one may be superior to the other depends on a more detailed analysis. Thus, the adjustment parameters for the quality of each model were compared from the parameters $R^2$, $R_{adjusted}$ and the Pearson's *r*, and are presented in Table 4.

The coefficients $R^2$ and $R_{adjusted}$ were very close to the Probit and Logit models, with a slight trend of higher indices for the Logit model. The coefficient *r* of Pearson also had higher indices for the Logit model. This index measures the intensity of the linear correlation between the variables; therefore, the larger it is, the greater the linear fit of the model. If we consider that the values resulting from the adjustment function must be as close as possible to a straight line, this index also indicates a proportionality

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

**Table 3.** Comparison of the experimentally obtained P50, and the P50 calculated by the Probit and Logit models with four and six repetitions, respectively

| Lots code | EI (d) | Germination of EI (%) | $P_{50}$ Probit 4 Rep (SE) | $P_{50}$ Probit 6 Rep (SE) | $P_{50}$ Logit 4 Rep (SE) | $P_{50}$ Logit 6 Rep (SE) |
|---|---|---|---|---|---|---|
| S01 | 56–63 | 58–33 | **56.49** (0.58) | **57.50** (0.57) | **56.36** (1.04) | **57.29** (1.02) |
| S02 | 49–56 | 50–29 | **51.79** (0.51) | **52.72** (0.57) | **51.83** (1.60) | **52.88** (1.04) |
| S03 | 35–42 | 58–37 | 32.47 (0.40) | **38.43** (0.28) | 32.24 (0.72) | **38,07** (0.48) |
| S04 | 35–42 | 71–37 | **36.24** (0.42) | **36.52** (0.38) | **35.82** (0.72) | **36.37** (0.73) |
| S05 | 42–49 | 64–44 | 50.17 (0.46) | **49.09** (0.35) | **49.27** (0.87) | **49.10** (0.64) |
| S06 | 42–49 | 70–45 | **47.72** (0.40) | **45.54** (0.35) | **47.28** (0.75) | **45.32** (0.64) |
| S07 | 63–84 | 67–13 | **65.75** (0.36) | **64.53** (0.40) | **64.84** (0.68) | **64.05** (0.71) |
| S08 | 63–84 | 79–13 | **72.94** (0.56) | **73.77** (0.47) | **71.38** (1.06) | **72.90** (0.85) |
| S09 | 63–84 | 69–23 | **75.30** (0.33) | **75.03** (0.35) | **73.27** (0.64) | **74.09** (0.61) |
| S10 | 35–42 | 74–34 | 33.82 (0.86) | **37.24** (0.44) | 33.39 (2.79) | **37.15** (0.80) |
| S11 | 42–49 | 90–48 | 50.39 (1.03) | **47.61** (0.47) | **49.68** (1.88) | **47.64** (0.86) |
| S12 | 28–35 | 83–32 | **33.62** (0.99) | **33.32** (0.41) | **34.00** (2.40) | **33.25** (0.71) |
| S13 | 28–35 | 57–46 | **35.59** (0.62) | **29.19** (0.23) | **35.48** (2.93) | **29.02** (0.44) |
| S14 | 14–21 | 69–29 | 12.04 (0.72) | **16.49** (0.27) | 13.31 (0.43) | **16.39** (0.45) |
| S15 | 14–21 | 67–30 | **16.83** (0.21) | **16.78** (0.22) | **16.88** (0.35) | **16.87** (0.38) |

EI, expected interval (time at which the P50 was experimentally identified); P50, viability of 50% of seed lot; In bold are the P50 values found for each model that coincide with the expected interval (IE).

behaviour regarding the linearity of the function, that is, the results indicate that the linear adjustment of the Logit function is higher than of the Probit function and consequently indicates more robustness of prediction.

The diagnosis of the normality of the residues indicates how closely the residues generated by the models are to a normal distribution. It also reveals if the dispersion of the data around the regression line is uniform. The normality of the residuals generated by the models is an efficient tool to verify the robustness of models regarding the quality of the prediction. The evaluation of normality can be performed from tests such as S–W, Lilliefors and Komogorov–Smirnov, and also be complemented by a graphical analysis to provide a better reading of the behaviour of the residues (Schützenmeister et al., 2012). We used the normal Q–Q plot to perform this graphical analysis.

When the residue normality tests have a *P*-value higher than the significance level (0.05), this indicates that there is normality in the generated residues. Otherwise, they reject the normality (Table 5).

The Logit model had a lower rejection of normality when compared with the Probit model for most of the evaluated lots. It had only two rejections for only one of the three tests applied in lots S03 and S04. On the other hand, the Probit model obtained 13 rejections regarding normality, with five lots being rejecting by more than two tests, that is, around 33% of the samples.

The graphical analysis of the normal Q–Q plot indicates the robustness of the Logit model more clearly when compared with the Probit model. Figures 1–4, referring to lots S01, S02, S03 and S11, respectively, allow us to evaluate this difference. The Probit model for the data from lot S01 has the tails far from normality, indicating the difficulty of this model to adjust to this behaviour. The Logit model, on the other hand, clearly has greater proximity to the normal line.

Figures 1–4 present normality behaviour close between the Probit and Logit models. However, the Logit model had closest to the standard line, in most samples analysed. The normal Q–Q plot shows a lack of normality in both models, especially in the extreme points of tails, but it is more pronounced in the Probit model. Another critical issue is the central axis of the graph (0,0), where it is predicted the P50. At this point, the Probit model stays a greater distance from the standard line, which could increase the prediction errors in this model.

Lots S04, S05, S06, S07, S08, S09, S10, S12, S13, S14 and S15 presented similar behaviour to lots S01 and S02, S03 and S11, with the Logit model performing better in all of them. In the case of lot S11, the Logit model also clearly has greater proximity to the normal line, as can be observed in Fig. 4.

## Discussion

Chen and Tsurumi (2011) reported the difficulty of distinguishing Probit and Logit models. The authors suggest that this is only possible with a relevant number of samples. Therefore, the classical parameters evaluation criteria have little efficiency in distinguishing them, and we suggested that the analysis of the normality of the residues could function as a tool to elucidate the differences, which was evidenced by the tests of S–W, Komogorov–Smirnov and Lilliefors, and by the normal Q–Q plot described.

The diagnosis of normality from the normal Q–Q plot showed very important to distinguish the models. Because if there is confluence of the observed data to those predicted by the normal line at the point of interest, in this case the axis (0,0), the P50 was predicted correctly.

The evaluation of the normality of the residues showed the robustness of the Logit model in adjusting the sigmoid curve

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

**Table 4.** Comparison of adjustment parameters of the Probit and Logit models, with *n* sample of six replicates

| | | Parameters of model** | | Parameters of adjust of the model | | |
|---|---|---|---|---|---|---|
| Lots | Model | Intercept (SE) | Slope (SE) | Person's *r* | $R^2$ | $R_{adj}$ |
| S01 | Probit | 2.63(0.16) | 0.16(0.004) | −0.70 | 0.73 | 0.73 |
| | Logit | 4.75 (0.29) | −0.08 (0.007) | −0.82 | 0.74 | 0.73 |
| S02 | Probit | 2.66 (0.23) | −0.05(0.006) | −0.91 | 0.66 | 0.65 |
| | Logit | 4.80 (0.42) | −0.09 (0.010) | −0.91 | 0.66 | 0.65 |
| S03 | Probit | 1.68 (0.09) | −0.04 (0.003) | −0.91 | 0.81 | 0.81 |
| | Logit | 2.90 (0.15) | −0.07 (0.006) | −0.91 | 0.81 | 0.81 |
| S04 | Probit | 1.84 (0.11) | −0.05 (0.003) | −0.89 | 0.79 | 0.78 |
| | Logit | 3.20 (0.19) | −0.09 (0.006) | −0.89 | 079 | 0.79 |
| S05 | Probit | 2.43 (0.11) | −0.05 (0.003) | −0.93 | 0.84 | 0.84 |
| | Logit | 4.34 (0.20) | −0.09 (0.006) | −0.92 | 0.84 | 0.83 |
| S06 | Probit | 2.86 (0.13) | −0.06 (0.003) | −0.95 | 0.89 | 0.89 |
| | Logit | 5.08 (0.24) | −0.11 (0.240) | −0.95 | 0.89 | 0.88 |
| S07 | Probit | 1.99 (0.10) | −0.03 (0.002) | −0.84 | 0.76 | 0.75 |
| | Logit | 3.48 (0.18) | −0.05 (0.004) | −0.84 | 0.75 | 0.75 |
| S08 | Probit | 2.27 (0.12) | −0.03 (0.002) | −0.77 | 0.70 | 0.69 |
| | Logit | 4.07 (0.23) | −0.05 (0.005) | −0.82 | 0.70 | 0.70 |
| S09 | Probit | 1.85 (0.09) | −0.02 (0.002) | −0.76 | 0.73 | 0.72 |
| | Logit | 3.23 (0.16) | −0.04 (0.003) | −0.84 | 0.73 | 0.72 |
| S10 | Probit | 2.76 (0.15) | −0.07 (0.004) | −0.82 | 0.88 | 0.88 |
| | Logit | 4.93 (0.27) | −0.13 (0.007) | −0.91 | 0.88 | 0.88 |
| S11 | Probit | 4.45 (0.27) | −0.09 (0.005) | −0.92 | 0.87 | 0.87 |
| | Logit | 8.13 (0.48) | −0.17 (0.010) | −0.92 | 0.87 | 0.87 |
| S12 | Probit | 3.72 (0.29) | −0.11 (0.011) | −0.89 | 0.83 | 0.82 |
| | Logit | 6.79 (0.50) | −0.20 (0.019) | −0.89 | 0.84 | 0.83 |
| S13 | Probit | 2.25 (0.09) | −0.07 (0.003) | −0.97 | 0.93 | 0.93 |
| | Logit | 3.99 (0.17) | −0.13 (0.006) | −0.96 | 0.93 | 0.93 |
| S14 | Probit | 1.04 (0.10) | −0.06 (0.004) | −0.92 | 0.88 | 0.88 |
| | Logit | 1.77 (0.17) | −0.11 (0.006) | −0.92 | 0.89 | 0.88 |
| S15 | Probit | 1.74 (0.07) | −0.10 (0.004) | −0.89 | 0.96 | 0.96 |
| | Logit | 3.03 (0.13) | −0.18 (0.007) | −0.97 | 0.96 | 0.96 |

**Significant by the *t*-test at the 99% probability level.

formed by the longevity of soybean seeds. Although this study did not present wild species, it expects that they would behave less homogeneously than agricultural species in general, which would probably increase the incidence of outliers and heavier tails. Therefore, it can be inferred that the Logit model would also be superior to Probit in these cases; however, it is necessary to further investigation for confirmation.

The Probit model proposed by Ellis and Roberts (1981) can show significant errors, which, at high temperatures and relative humidity in the case of soybean, reached 802% of prediction uncertainty (Sinício, 2004). Therefore, through the diagnostic of residue normality generated by the Probit and Logit models, it is possible to observe a greater robustness of the Logit model. However, we indicated that the model be adjusted with an increment of a minimum *n* equal to six replicates.

## Conclusions

The Logit model is more adequate than the Probit function for the prediction of longevity in soybean seeds (P50). It presents greater robustness to the non-normality of the residues after the transformation by the application of the model. The increase in the *n* sample from four to six replicates is necessary to improve the goodness of fit of the model to predict the P50 in soybean seeds. The diagnosis of normality of the residues by

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

**Table 5.** Comparison of three normality tests applied to the goodness of fit, the Probit and Logit models, with six replicates

| | | Probit Decision to normality at level (5%) | | Logit Decision to normality at level (5%) | |
|---|---|---|---|---|---|
| Lot | Test | *P*-value | Decision | *P*-value | Decision |
| S01 | S–W | $4.12 \times 10^{-7}$ | Reject normality | 0.4 | Cannot reject |
| | Lilliefors | $5.62 \times 10^{-7}$ | Reject normality | 0.09 | Cannot reject |
| | K–S | 0.03 | Reject normality | 0.50 | Cannot reject |
| S02 | S–W | 0.02 | Reject normality | 0.06 | Cannot reject |
| | Lilliefors | 0.13 | Cannot reject | 0.12 | Cannot reject |
| | K–S | 0.13 | Cannot reject | 0.12 | Cannot reject |
| S03 | S–W | 0.23 | Cannot reject | 0.04 | Reject normality |
| | Lilliefors | 0.20 | Cannot reject | 0.70 | Cannot reject |
| | K–S | 0.70 | Cannot reject | 0.52 | Cannot reject |
| S04 | S–W | 0.16 | Cannot reject | 0.16 | Cannot reject |
| | Lilliefors | 0.007 | Reject normality | 0.02 | Reject normality |
| | K–S | 0.20 | Cannot reject | 0.27 | Cannot reject |
| S05 | S–W | 0.56 | Cannot reject | 0.056 | Cannot reject |
| | Lilliefors | 0.20 | Cannot reject | 0.16 | Cannot reject |
| | K–S | 0.83 | Cannot reject | 0.59 | Cannot reject |
| S06 | S–W | 0.36 | Cannot reject | 0.41 | Cannot reject |
| | Lilliefors | 0.2 | Cannot reject | 0.2 | Cannot reject |
| | K–S | 0.64 | Cannot reject | 1 | Cannot reject |
| S07 | S–W | 0.37 | Cannot reject | 0.21 | Cannot reject |
| | Lilliefors | 0.2 | Cannot reject | 0.2 | Cannot reject |
| | K–S | 1 | Cannot reject | 0.97 | Cannot reject |
| S08 | S–W | 0.00001 | Reject normality | 0.37 | Cannot reject |
| | Lilliefors | 0.02 | Reject normality | 0.20 | Cannot reject |
| | K–S | 0.27 | Cannot reject | 0.89 | Cannot reject |
| S09 | S–W | $4.95 \times 10^{-9}$ | Reject normality | 0.55 | Cannot reject |
| | Lilliefors | $7.75 \times 10^{-4}$ | Reject normality | 0.20 | Cannot reject |
| | K–S | 0.09 | Cannot reject | 1 | Cannot reject |
| S10 | S–W | $3.51 \times 10^{-6}$ | Reject normality | 0.94 | Cannot reject |
| | Lilliefors | 0.002 | Reject normality | 0.2 | Cannot reject |
| | K–S | 0.14 | Cannot reject | 1 | Cannot reject |
| S11 | S–W | 0.68 | Cannot reject | 0.90 | Cannot reject |
| | Lilliefors | 0.20 | Cannot reject | 0.20 | Cannot reject |
| | K–S | 0.88 | Cannot reject | 1 | Cannot reject |
| S12 | S–W | 0.36 | Cannot reject | 0.34 | Cannot reject |
| | Lilliefors | 0.20 | Cannot reject | 0.18 | Cannot reject |
| | K–S | 0.78 | Cannot reject | 0.64 | Cannot reject |
| S13 | S–W | 0.91 | Cannot reject | 0.90 | Cannot reject |
| | Lilliefors | 0.2 | Cannot reject | 0.2 | Cannot reject |
| | K–S | 1 | Cannot reject | 1 | Cannot reject |
| S14 | S–W | 0.09 | Cannot reject | 0.07 | Cannot reject |
| | Lilliefors | 0.12 | Cannot reject | 0.16 | Cannot reject |

(*Continued*)

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

**Table 5.** (*Continued.*)

| | | Probit Decision to normality at level (5%) | | Logit Decision to normality at level (5%) | |
|---|---|---|---|---|---|
| Lot | Test | *P*-value | Decision | *P*-value | Decision |
| | K–S | 0.55 | Cannot reject | 0.62 | Cannot reject |
| S15 | S–W | $3.81 \times 10^{-4}$ | Reject normality | 0.19 | Cannot reject |
| | Lilliefors | $2.03 \times 10^{-4}$ | Reject normality | 0.20 | Cannot reject |
| | K–S | 0.058 | Cannot reject | 0.75 | Cannot reject |

S–W, Shapiro–Wilk test; K–S, Komogorov–Smirnov test.

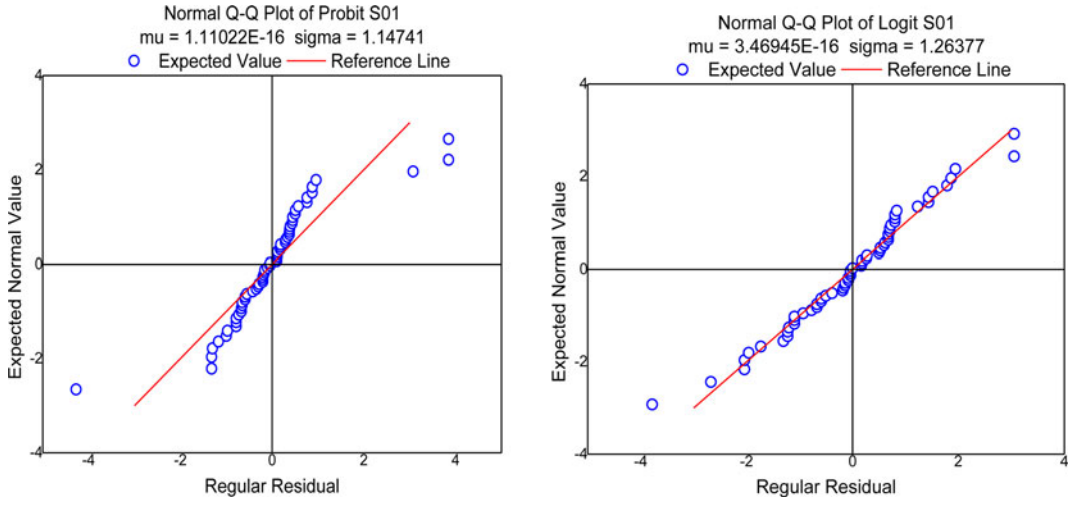



**Fig. 1.** Normal Q–Q plot for the Probit (a) and Logit (b) models, respectively, referring to lot S01, with six replicates.

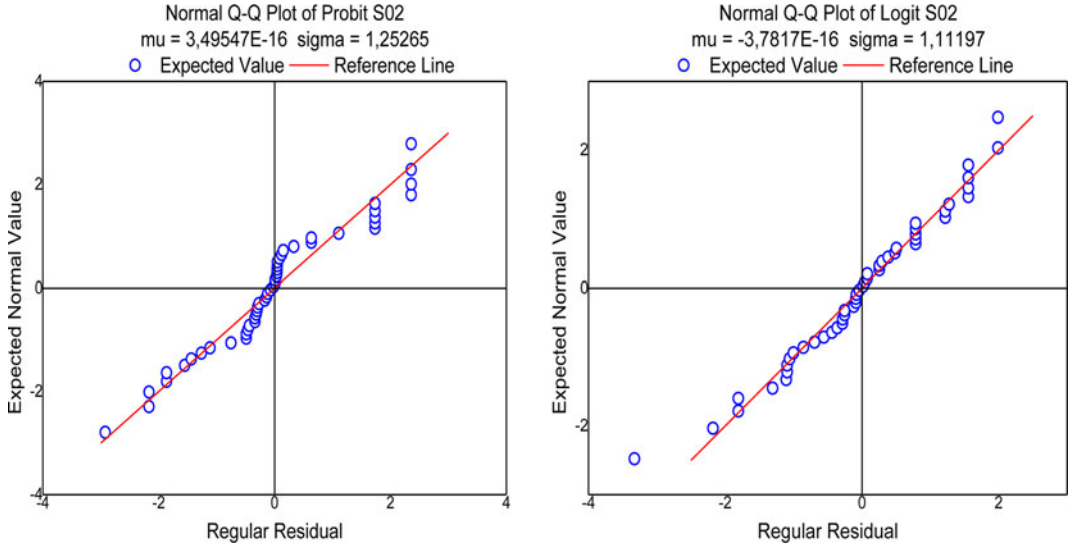



**Fig. 2.** Normal Q–Q plot for the Probit (a) and Logit (b) models, respectively, referring to lot S02, with six replicates.

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

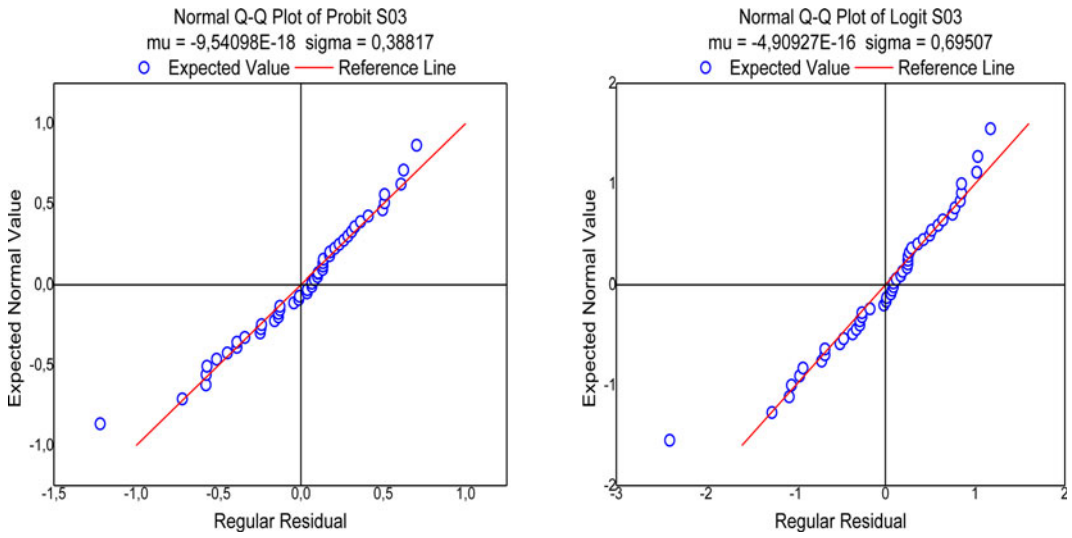



**Fig. 3.** Normal Q–Q plot for the Probit (a) and Logit (b) models, respectively, regarding lot S03, after increase the *n* sample to six replicates.

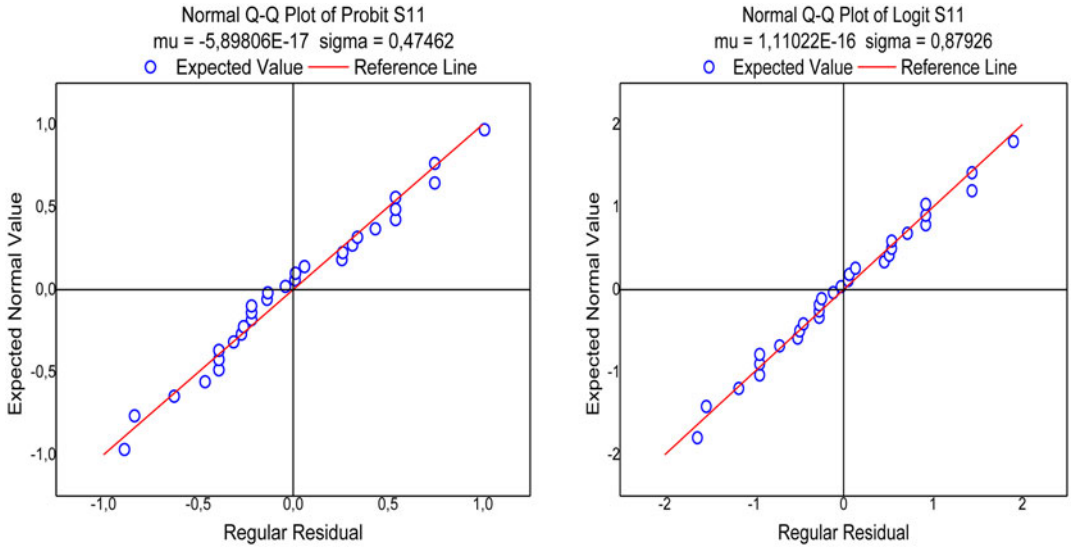



**Fig. 4.** Normal Q–Q plot for the Probit (a) and Logit (b) models, respectively, referring to lot S11, after increasing the *n* sample to six replicates.

the classic tests and the normal Q–Q plot can be considered as a selection criterion for the goodness of fit of the Probit and Logit models.

**Funding.** This work was supported by FAPESP (Foundation for Research Support of the Province of São Paulo), grant no. FAPESP n° 2016/13126-0; English language adjustments by Roger Hutchings.

## References

**Agresti A** (2013) *Categorical data analysis*. Hoboken, NJ, USA, John Wiley & Sons.

**Bobkov, SG, Chistyakov, GP, Götze, F** (2014) *Probability theory and related fields*, vol. **159**, pp. 1–59. Springer-Verlag Berlin Heidelberg.

**Chen G and Tsurumi H** (2011) Probit and Logit model selection. *Communications in Statistics – Theory and Methods* **40**, 159–175.

**Demir I and Mavi K** (2003) Effect of controlled hydration treatments on storage longevity of aubergine seeds during development. *Israel Journal of Plant Sciences* **51**, 291–295.

**Ellis RH and Roberts EH** (1981) The quantification of aging and survival in orthodox seeds. *Seed Science and Technology* **9**, 373–409.

**Ellis RH, Hong TD, Roberts EH and Tao K-L** (1990) Low moisture content limits to relations between seed longevity and moisture. *Annals of Botany* **65**, 493–504.

**Fischer H** (2011) *History of the central limit theorem: from classical to modern probability theory*. New York, Springer.

**International Seed Testing Association (ISTA)** (1985) International rules for seed testing. *Seed Science & Technology* **13**, 300–520.

**Marcondes MC, Andreoli C and Miglioranza E′** (2011) Viability equation to determine the longevity of fungicide-treated seeds of wheat stored in a conventional warehouse. *Acta Scientiarum* **33**, 539–544.

**McDonald MB, Evans AF, Bennett MA, Fujimura K, Sako Y, Xu L and Hoffmaster, A** (2008) Using the seed vigor imaging system for improving stand establishment. *Acta Horticulturae* **782**, 83–91.

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press

**Obunyali CO, Muasya RM, Nyamongo DO and Van Rheenen HA** (2008) Study on comparative longevity of banked and freshly collected seeds of two wild sesame species. *South African Journal of Botany* **74**, 764–767.

**Oral E** (2006) Binary Regression with Stochastic Covariates, Communications in Statistics - Theory and Methods **35**(8), 1429–1447.

**Pereira Lima JJ, Buitink J, Lalanne D, Rossi RF, Pelletier S, da Silva EAA and Leprince O** (2017) Molecular characterization of the acquisition of longevity during seed maturation in soybean. *PLoS ONE* **12**, e0180282.

**Schützenmeister A, Jensen U and Piepho H-P** (2012) Checking normality and homoscedasticity in the general linear model using diagnostic plots. *Communications in Statistics: Simulation and Computation* **41**, 141–154.

**Sinício R** (2004) Generalised longevity model for orthodox seeds. *Biosystems Engineering* **89**, 85–92.

**Thode HC** (2002) *Testing for normality*. New York: Marcel Dekker.

**Walters C** (2015) Genebanking seeds from natural populations. *Natural Areas Journal* **35**, 98–105.

https://doi.org/10.1017/S0960258520000136 Published online by Cambridge University Press