

## RESEARCH NOTE

# SeedCalc, a new automated R software tool for germination and seedling length data processing<sup>1</sup>

Laércio Junio da Silva<sup>2\*</sup>, André Dantas de Medeiros<sup>2</sup> ,  
Ariadne Morbeck Santos Oliveira<sup>2</sup>

**ABSTRACT** – The need to optimize seed quality assessment using new, more accessible, and modern computational resources has led to the emergence of new tools. In this paper, we introduce SeedCalc, a new R software package developed to process germination and seedling length data. The functions included in SeedCalc allow fast and efficient data processing, offering greater reliability to the variables generated and facilitating statistical analysis itself since the data are already processed with the appropriate structure to be statistically analyzed in the R software. SeedCalc is available free of charge at <https://CRAN.R-project.org/package=SeedCalc>.

**Index terms:** R package, germination indexes, seedling uniformity, seed vigor, data analysis.

## SeedCalc, uma nova ferramenta automatizada do software R para processamento de dados de germinação e comprimento de plântula

**RESUMO** – A necessidade de otimizar a avaliação da qualidade de sementes utilizando novos recursos computacionais mais acessíveis e modernos levou ao surgimento de novas ferramentas. Neste artigo, introduzimos o SeedCalc, um novo pacote do software R desenvolvido para processar dados de germinação e comprimento de plântulas. As funções incluídas no SeedCalc permitem um processamento de dados rápido e eficiente, oferecendo maior confiabilidade às variáveis geradas e facilitando a própria análise estatística, uma vez que os dados já são processados com a estrutura apropriada para serem analisados estatisticamente no software R. O SeedCalc está disponível gratuitamente em <https://CRAN.R-project.org/package=SeedCalc>.

**Termos para indexação:** Pacote do R, índices de germinação, uniformidade de plântulas, vigor de sementes, análise de dados.

## Introduction

Electronic spreadsheets are used in a generalized manner to tabulate data, perform the most diverse experimental calculations, and process information. Nevertheless, errors in formulas or in handling of spreadsheets can significantly compromise the information generated (Powell et al., 2009). Thus, development of tools that minimize these risks and can increase the reliability of results is important.

Tests for evaluation of seed quality are routinely used in experimentation. One of the most prominent of these routine

tests is the germination test, nearly always associated with some vigor test, such as a seedling growth test (Nakagawa et al., 1999). In the germination test, daily counting of the number of seeds germinated is common, and these data are frequently used for calculation of the Germination Speed Index (Maguire, 1962) and Germination Speed (Edmond and Drapala, 1958). This same procedure is applied to the seedling emergence data. Nevertheless, various other indexes can be calculated based on these data, which may be useful for making inferences regarding different aspects of seed germination response, such as uniformity, coefficient of

<sup>1</sup>Submitted on 12/03/2018. Accepted for publication on 02/15/2019.

<sup>2</sup>Departamento de Fitotecnia, Universidade Federal de Viçosa, 36570-900 – Viçosa, MG, Brasil.

\*Corresponding author <laercio.silva@ufv.br>

variation, synchrony, among others.

Regarding seedling growth data, variables such as shoot, root, and total length are those most used, when image analysis software that contains other indexes that are automatically calculated based on these parameters is not used. Thus, several indexes may be calculated from these data, which could be a complicated task in dealing with a large number of lots/treatments, as occurs, for example, in phenotyping studies (Joosen et al., 2010), when a large number of seed lots or seeds from different cultivars are analyzed.

Among the tools that are able to serve the above purposes is SeedCalc, a new R package. R (R Core Team, 2019) is a powerful software used for statistical data analysis that has grown in importance in seed research. In this respect, several packages are available for this platform, which increases its applicability for analysis of the most varied types of data. SeedCalc was developed from the need to optimize acquisition of data from seed quality indicators. The package is available free of charge through CRAN (<https://CRAN.R-project.org/package=SeedCalc>), and it can be used on various platforms, such as Windows, MacOS, and Linux.

For calculations, SeedCalc uses data from daily counting of germination/emergence and seedling length measurements to automatically generate a series of variables related to seed physiological quality. It includes calculations of germination/emergence percentage, time required for germination or emergence (T10, T50, T90, and mean germination time), speed (germination speed index, mean germination rate), variability or heterogeneity (coefficient of variation of germination time, variance of germination), uncertainty, and synchrony. In addition, through seedling measurements, some indexes can be generated, such as growth, uniformity, and vigor indexes.

Thus, considering the importance of developing systems that allow calculation of seed quality indicators in a simple and automatic way, the aim of this study was to present the SeedCalc package and its applicability to analysis of data obtained from germination and seedling growth tests.

## Materials and Methods

All the functions developed were written in R programming language and, therefore, can be carried out in the R environment. R can be installed in the Windows, Linux, or Mac systems. Thus, this package can be freely used by the scientific community, regardless of the operating platform used.

The SeedCalc package can be installed simply by inserting the following command in the R software:

```
> install.packages("SeedCalc")
```

The following command must be typed in to load the package:

```
> library(SeedCalc)
```

The functions inserted in the package and the respective equations used, which were obtained from the scientific literature in the seed area, are described as follows.

**GermCalc:** applies all the functions related to data of seed germination and seedling emergence (Table 1).

**PlantCalc:** applies all the functions related to seedling analysis (Table 2).

To use these functions, the original data files can be saved in text files or electronic spreadsheets. However, they must meet some requirements. For the **GermCalc** function, the file must contain the first column with the time (any unit of time - days, hours, etc.), and the rest of the columns, identified with the treatment/lot, the number of seeds germinated in a cumulative manner (see Figure 1). For the **PlantCalc** function, the file must contain four columns in the following order: identification of the lot/treatment, seedling (identified from 1 to  $n$ , with  $n$  being the total number of seedlings), shoot length, and root length. The title of the columns, contained in the first line of the file, can receive any name, as long as the order of the columns respects the sequence of reference (see Figure 2A).

For generation of the corrected vigor index (Medeiros and Pereira, 2018), an additional file is necessary, composed of two columns: the first identifies the treatments/lots and the second, the germination percentages (see Figure 2B). The identifications of the treatments/lots must be identical in both files.

As an example of practical application of the SeedCalc package, real data were used from an experiment conducted with five commercial lots of soybean seeds. The seeds were tested for germination in accordance with the Rules for Seed Testing (Brasil, 2009). Four replications of 50 seeds each were placed on paper toweling (Germitest®) moistened with distilled water at the rate of 2.5 times the weight of the paper. Then rolls were made, and they were kept in BOD at 25 °C in a program of 8 hours of light and 16 hours in the dark. Daily counts of the number of normal seedlings were made. For the seedling growth test, the seeds were placed in a linear arrangement on the upper third of the paper toweling (Germitest®) and were maintained under the same conditions described for the germination test. At three days after the beginning of the test, the shoot and root length were measured using a millimeter ruler. The data obtained in the germination and seedling growth tests were processed through the SeedCalc package.

Table 1. Functions contained in the SeedCalc package for calculation of indexes, using data from daily counting of seedling on germination/emergence test.

Function	Measure	Formula	Reference
FGP	Final germination percentage	$FGP = (n/N) \times 100$ $n$ is the number of seeds germinated, and $N$ is the total number of seeds.	ISTA (2015)
GI	Germination speed index	$GI = \sum_{i=1}^k (n_i / t_i)$ $n$ is the number of seeds germinated on each day of daily counting up to the last count, and $t$ is the number of days after the beginning of the test in each count.	Maguire (1962)
T10	Time required for germination of 10% of the seeds	$T_{10} = \frac{ti + \left\{ \left[ \frac{N}{\left( \frac{100}{10} \right)} \right] - ni \right\} (tj - ti)}{(nj - ni)}$ $N$ is the final number of seeds germinated, and $n_i$ and $n_j$ are the total number of seeds germinated in adjacent counts in time $t_i$ and $t_j$ , respectively, when $ni < \frac{N+1}{2} < nj$ .	Adapted from Farooq et al. (2005)
T50	Time required for germination of 50% of the seeds	$T_{50} = \frac{ti + \left\{ \left[ \frac{N}{\left( \frac{100}{50} \right)} \right] - ni \right\} (tj - ti)}{(nj - ni)}$ Same codification as for T10.	Farooq et al. (2005)
T90	Time required for germination of 90% of the seeds	$T_{90} = \frac{ti + \left\{ \left[ \frac{N}{\left( \frac{100}{90} \right)} \right] - ni \right\} (tj - ti)}{(nj - ni)}$ Same codification as for T10.	Adapted from Farooq et al. (2005)
MGT	Mean germination time	$MGT = \sum_{i=1}^k n_i t_i / \sum_{i=1}^k n_i$ $n_i$ is the number of seeds germinated per day (not the accumulated number, but the number corresponding to the $i$ -th observation), and $t_i$ is the time since the beginning of the germination test up to the $i$ -th observation.	Labouriau (1983)
MGR	Mean germination rate	$\bar{v} = CoVg / 100 = 1 / \bar{t}$ $\bar{t}$ is the mean germination time, and $CoVg$ is the germination speed coefficient.	Labouriau (1983)
VarGer	Variance of germination time	$s_t^2 = \sum_{i=1}^k n_i (t_i - \bar{t})^2 / \left( \sum_{i=1}^k n_i - 1 \right)$ $t$ is the mean germination time, $t_i$ is the time between the beginning of the experiment and the $i$ -th observation (day or hour), $n_i$ is the number of seeds germinated in time $i$ , and $k$ is the last count of the germination test.	Labouriau (1983)
CVt	Coefficient of variation of germination time	$CV_t = (S_t / \bar{t}) 100$ $S_t$ is standard deviation of the germination time, and $\bar{t}$ is mean germination time.	Carvalho et al. (2005)
Sinc	Germination synchrony	$Z = \sum Cni, 2 / N$ $Cni, 2 = ni(ni-1)/2$ and $N = \sum ni(\sum ni - 1)/2$ $Cni$ is the combination of the seeds germinated in time $i$ , two by two, and $ni$ is the number of seeds germinated in time $i$ .	Primack (1980)

Table 1. Continuation.

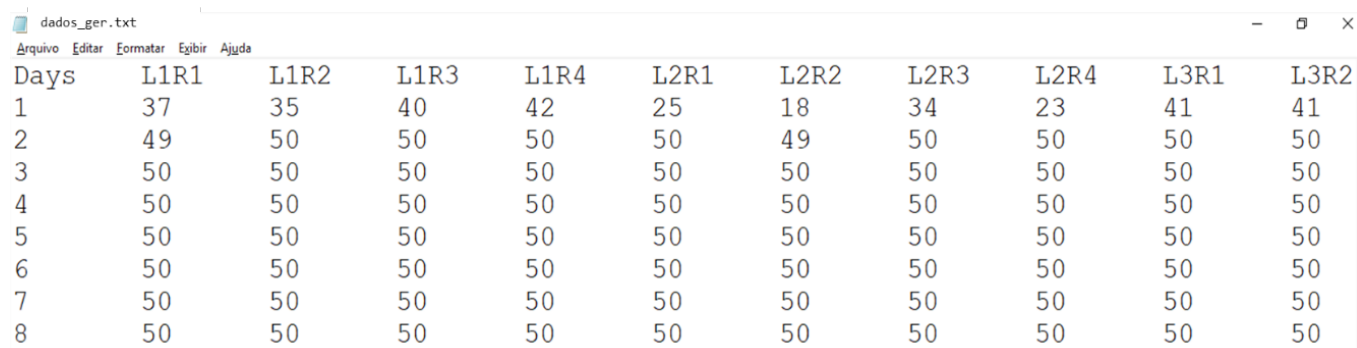
Function	Measure	Formula	Reference
Unc	Uncertainty	$\bar{E} = - \sum_{i=1}^k f_i \log 2 f_i$ <p>with <math>f_i</math> given by,</p> $f_i = n_i / \sum_{i=1}^k n_i$ <p><math>f_i</math> is the relative frequency of germination, and <math>n_i</math> is the number of seeds germinated on day <math>i</math>.</p>	Labouriau and Valadares (1976)
CVG	Germination speed coefficient	$CVG = \left( \sum_{i=1}^k f_i / \sum_{i=1}^k f_i x_i \right) 100$ <p><math>f_i</math> is the number of newly germinated seeds on day <math>i</math>, and <math>x_i</math> is the number of days from sowing.</p>	Nichols and Heydecker (1968)
UnifG	Uniformity of germination	$UnifG = (T90 - T10)$ <p>T90 is the time required for germination of 90% of the seeds, and T10 is the time required for germination of 10% of the seeds.</p>	Demilly et al. (2014)

Table 2. Functions contained in the SeedCalc package for obtaining indexes using seedling length measurement data.

Function	Measurement	Formula	Reference
mean_pa	Mean shoot length	$Mean_{pa} = \frac{\sum_{i=1}^k HL}{n}$ <p>HL is the length of the shoot of each seedling, and <math>n</math> is the total number of seedlings evaluated.</p>	Nakagawa et al. (1999)
mean_raiz	Mean root length	$Mean_{raiz} = \frac{\sum_{i=1}^k RL}{n}$ <p>RL is the length of the root of each seedling, and <math>n</math> is the total number of seedlings evaluated.</p>	Nakagawa et al. (1999)
mean_total	Mean total length	$Mean_{total} = \frac{\sum_{i=1}^k SL}{n}$ <p>SL is the total length of each seedling, and <math>n</math> is the total number of seedlings evaluated.</p>	Nakagawa et al. (1999)
mean_razao	Mean of the root/shoot ratio	$Mean_{razao} = \frac{\sum_{i=1}^k RRA}{n}$ <p>RRA is the ratio between the root and the shoot of each seedling, and <math>n</math> is the total number of seedlings evaluated.</p>	Benincasa (2003)
Unif_1	Uniformity index	$Unif_{1\_1} = \left[ 1 - \frac{\sum_{i=1}^n  X_i - \bar{X} }{n \times X} \right] \times 1000 - \left[ n_{dead} \times \left( \frac{50}{n_{total}} \right) \right]$ <p><math>X_i</math> is the length of the seedling analyzed, <math>\bar{X}</math> is the mean length of seedlings of the seed lot analyzed, <math>n</math> is the variable of total number of seedlings evaluated, <math>n_{dead}</math> is the number of ungerminated or dead seeds present, and <math>n_{total}</math> is the total number of seedlings.</p>	Christiansen (1942), adapted by Castan et al. (2018)
Unif_2	Uniformity index	$Unif_{2\_2} = \max[1000 - (0.75 \times sh + 0.5 \times sr + 2.5 \times s_{total} + 50 \times srh) - 0 \times n_{dead}]$ <p>sh, sr, and srh are the standard deviations of the length of the shoot, primary root, and root/shoot ratio, respectively.</p>	Sako et al. (2001)
Growth	Growth index	$Growth = [(mean(h) \times wh) + (mean(r) \times wr)]$ <p>mean(h) and mean(r) are the arithmetic means of shoot length and root length, respectively. wh and wr are adjustable weights in the formula for shoot and root, however, with reference values of 10 and 90, respectively.</p>	Sako et al. (2001)

Table 2. Continuation.

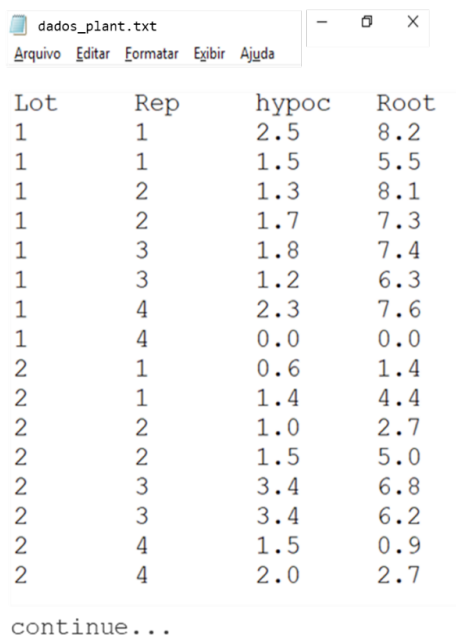
Function	Measurement	Formula	Reference
		$Vigor = (Growth \times wg) + (Uniformity \times wu)$	
Vigor	Vigor index	Growth is the growth index, and Uniformity is the uniformity index chosen by the user. wg and wu are adjustable weights in the formula for growth and uniformity, however, with standard values of 70 and 30, respectively.	Sako et al. (2001)
Vigor_corr	Corrected vigor index	$Vigor\_Corr = [(Growth \times wg) + (Uniformity \times wu)] \times \left(\frac{G}{100}\right)$ G is the percentage of germination of the seed lot.	Medeiros and Pereira (2018)



Days	L1R1	L1R2	L1R3	L1R4	L2R1	L2R2	L2R3	L2R4	L3R1	L3R2
1	37	35	40	42	25	18	34	23	41	41
2	49	50	50	50	50	49	50	50	50	50
3	50	50	50	50	50	50	50	50	50	50
4	50	50	50	50	50	50	50	50	50	50
5	50	50	50	50	50	50	50	50	50	50
6	50	50	50	50	50	50	50	50	50	50
7	50	50	50	50	50	50	50	50	50	50
8	50	50	50	50	50	50	50	50	50	50

Figure 1. Presentation of the germination count data organized for processing by SeedCalc.

A



Lot	Rep	hypoc	Root
1	1	2.5	8.2
1	1	1.5	5.5
1	2	1.3	8.1
1	2	1.7	7.3
1	3	1.8	7.4
1	3	1.2	6.3
1	4	2.3	7.6
1	4	0.0	0.0
2	1	0.6	1.4
2	1	1.4	4.4
2	2	1.0	2.7
2	2	1.5	5.0
2	3	3.4	6.8
2	3	3.4	6.2
2	4	1.5	0.9
2	4	2.0	2.7

continue...

B



Lot	Germination
1	100
2	88

continue ...

Figure 2. Presentation of the seedling data organized for processing by SeedCalc. In (A) are the seedling length data and in (B) are the germination values of the treatments/lots, corresponding to the ones in the first column of the data file of seedling length (A). The germination data (B) are optional and necessary for calculation of the corrected vigor index.

## Results and Discussion

The collection of functions available in the SeedCalc package implements various methods to describe the duration of the germination process in terms of germination indexes, as well as indexes related to seedling development. Thus, the data obtained in the germination and seedling growth tests were used to exemplify the application of the SeedCalc package. The use of the functions contained in the package and the interpretation of their outputs are better presented in the form of an example of application with real data.

As described, SeedCalc has two main functions: GermCalc and PlantCalc. The GermCalc function allows analysis of the daily counting data and receives the “Nseeds” argument, which refers to the number of seeds used in the germination test for each replication. The expression used for calculation of the germination indexes, with the use of the data presented in Figure 1, follows:

```
> dados_ger <- read.table('dados_ger.
txt', h=T)

> GermCalc(dados_ger, NSeeds = 50)

> GermCalc
```

Through this function, the variables presented in Figure 3 were generated.

The PlantCalc function allows analysis of the data of seedling length and, optionally, can receive the “Ger” argument, which refers to the mean of germination registered for the lots/treatments that is necessary to generate the Corrected vigor index. The expression used to carry out the analysis, using the data presented in Figure 2, follows:

```
> dados_plant <- read.table('dados_
plant.txt', h=T)

> germ <- read.table('per_ger.txt', h=T)

> PlantCalc(dados_plant, Ger = germ)

> PlantCalc
```

The variables presented in Figure 4 were generated through these commands. These variables (Figure 4) are normally generated automatically on specific software for image seedling analysis, such as Vigor-S® (Castan et al., 2018), SVIS® (Sako et al., 2001), Groundeye® (previously referred to as SAS) (Pinto et al., 2015) and SAPL® (Medeiros and Pereira, 2018). However, the few systems developed have the disadvantage to be limited to analysis of restricted number of species. Thus, making these quality indexes available would facilitate their use for these and other species.

With SeedCalc, the indexes are implemented automatically, and can be used to generate more detailed information regarding the vigor of seed lots of any species, regardless of the way the data were acquired and without

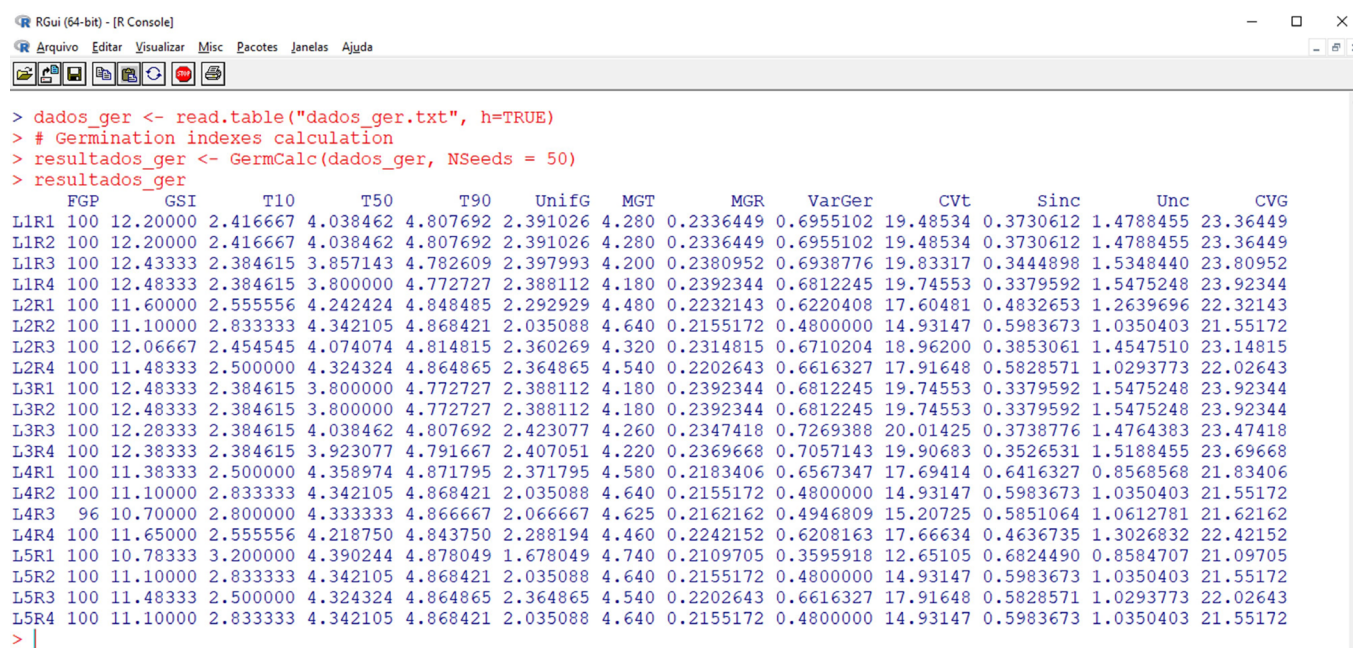


Figure 3. Example of results output of the functions applied to the daily counting data of the germination test of soybean seeds by SeedCalc.



the need to use specific seedling analysis systems. That way, this information becomes more accessible to the scientific community, and has a direct impact on the seed sector.

To illustrate, in Figure 5, soybean seedlings at three days after the beginning of the germination test can be seen from two samples, with their respective quality indexes. Lot B exhibits seedlings of shorter length and an irregular pattern, which is reflected in lower growth, uniformity, and vigor indexes compared to Lot A.

SeedCalc constitutes an innovative and efficient analysis tool to calculate indexes of seed germination and of seedling performance. The functions developed allow fast and efficient data processing, with a view to offering greater reliability to the variables generated and to facilitating statistical analysis itself, since the processed data have a suitable structure for analysis in R software (R Core Team, 2019). The use of these functions on the R software ensures they can be used freely by the scientific community.

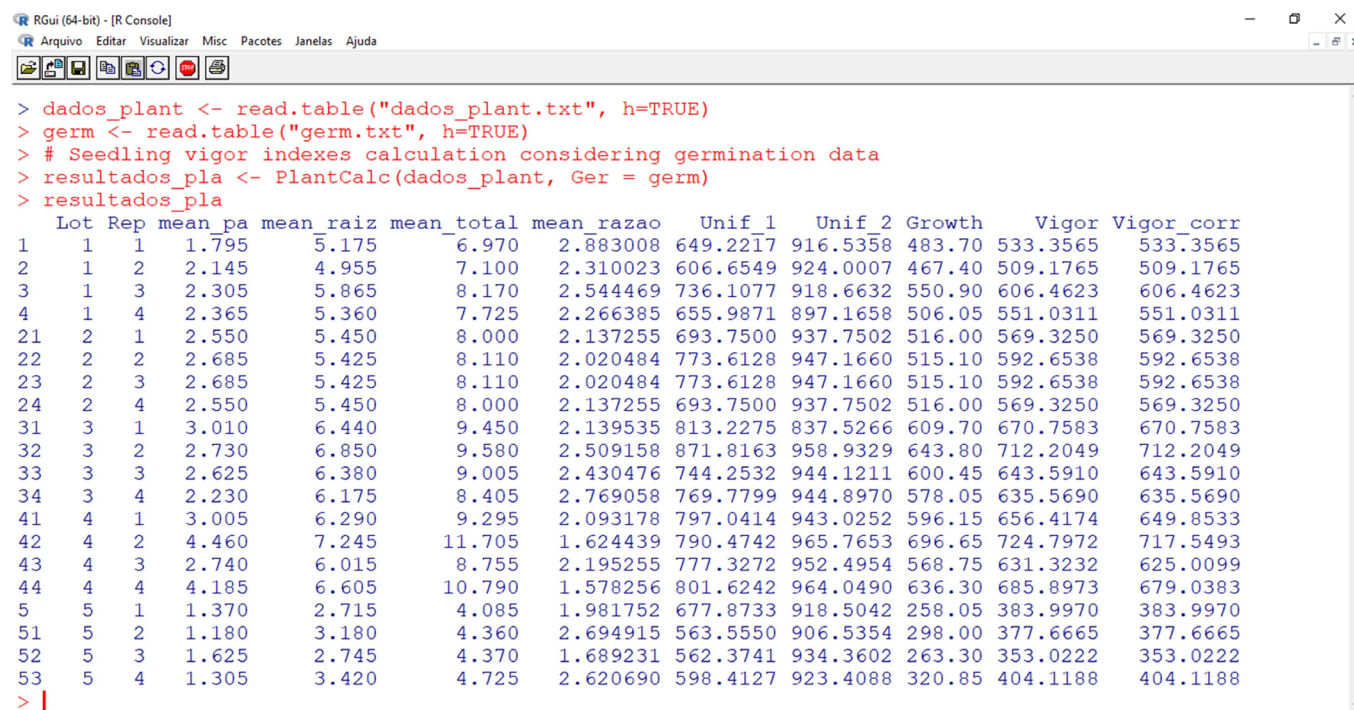


Figure 4. Example of results output of the functions applied to seedling lengths of soybean seeds by SeedCalc.



Figure 5. Soybean seedlings originated from seeds with different levels of vigor and their respective quality indexes. (A) higher vigor and (B) lower vigor.

## Conclusions

The SeedCalc package is a free access tool and generates indexes based on daily counting data from germination/emergence and seedling growth tests. It represents a powerful tool for research and will establish new computational approaches within the seed technology sector.

## Acknowledgements

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil* (CAPES) - Finance code 001.

## References

- BENINCASA, M.M.P. Análise do crescimento em plantas – noções básicas. Jaboticabal: FUNEP, 412p. 2003.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. *Regras para análise de sementes*. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. Brasília: MAPA/ACS, 2009.
- CARVALHO, M.P.; SANTANA, D.G.; RANAL, M.A. Emergência de plântulas de *Anacardium humile* A. St.-Hil. (Anacardiaceae) avaliada por meio de amostras pequenas. *Revista Brasileira de Botânica*, v.28, n.3, p.627-633, 2005. <http://dx.doi.org/10.1590/S0100-84042005000300018>.
- CASTAN, D.O.C.; GOMES-JUNIOR, F.G.; MARCOS-FILHO, J. Vigor-S, a new system for evaluating the physiological potential of maize seeds. *Scientia Agricola*, v.75, n.2, p.167-172, 2018. <http://dx.doi.org/10.1590/1678-992x-2016-0401>.
- CHRISTIANSEN, J.E. Irrigation by Sprinkling. University of California, Berkeley, CA, USA. (California Agriculture Experiment Station Bulletin, 670). 1942.
- DEMILLY, D.; DUCOURNAU, S.; WAGNER, M.H.; DÜRR, C. Digital imaging of seed germination. In: GUPTA, S.D.; IBARAKI, Y. (Eds.). *Plant Image Analysis: Fundamentals and Applications*. 1. ed. CRC Press, 2014. p.147-164.
- EDMOND, J.B.; DRAPALA, W.J. The effects of temperature, sand and soil, and acetone on germination of okra seeds. *Proceedings of American Society of Horticultural Science*, v.71, n.2, p.428-434, 1958.
- FAROOQ, M.; BASRA, S.M.A.; AHMAD, N.; HAFEEZ, K. Thermal Hardening: A New Seed Vigor Enhancement Tool in Rice. *Journal of Integrative Plant Biology*, v.47, n.2, p.187-193, 2005. <http://dx.doi.org/10.1111/j.1744-7909.2005.00031.x>.
- ISTA. The germination test. In: *International rules for seed testing*. Zurich, Switzerland: International Seed Testing Association, 2015. <http://dx.doi.org/10.15258/istarules.2015.05>. p. i-5-56.
- JOONSEN, R.V.L.; KODDE, J.; WILLEMS, L.A.J.; LIGTERINK, W.; VAN DER PLAS, L.H.W.; HILHORST, H.W.M. *Germinator*: A software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *Plant Journal*, v.62, n.1, p. 148-159, 2010. <http://dx.doi.org/10.1111/j.1365-313X.2009.04116.x>
- LABOURIAU, L.G. Uma nova linha de pesquisa na fisiologia da germinação das sementes. *Anais do XXXIV Congresso Nacional de Botânica*. SBB, Porto Alegre, 11-50, 1983.
- LABOURIAU L.G.; VALADARES M.B. On the germination of seeds of *Calotropis procera*. *Anais da Academia Brasileira de Ciências* v.48, p.174-186, 1976.
- MAGUIRE, J.D. Speed of germination-aid selection and evaluation for seedling emergence and vigor. *Crop Science*, v.2, p.176-177, 1962. <http://dx.doi.org/http://dx.doi.org/10.2135/cropsci1962.0011183X0002000200033x>.
- MEDEIROS, A.D.; PEREIRA, M.D. SAPL ® : a free software for determining the physiological potential in soybean seeds. *Pesquisa Agropecuária Tropical*, v.48, n.3, p.222-228, 2018. <http://dx.doi.org/10.1590/1983-40632018v4852340>.
- NAKAGAWA, J.; KRZYZANOWSKI, F.C.; VIEIRA, R.D.; FRANÇA-NETO, J.B. Testes de vigor baseados no desempenho das plântulas. In: KRZYZANOWSKI, F.C.; VIEIRA, R.D.; FRANÇA-NETO, J.B. (Eds.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 1999. p. 9-13.
- NICHOLS, M.A.; HEYDECKER, W. Two approaches to the study of germination data. *Proceedings of the International Seed Testing Association*, v.33, p.531-540, 1968.
- PINTO, C.A.G.; CARVALHO, M.L.M.; ANDRADE, D.B.; LEITE, E.R.; CHALFOUN, I. Image analysis in the evaluation of the physiological potential of maize seeds. *Revista Ciência Agronômica*, v.46, n.2, p.319-328, 2015. <http://dx.doi.org/10.5935/1806-6690.20150011>.
- POWELL, S.G.; BAKER, K.R.; LAWSON, B. Errors in operational spreadsheets. *Journal of Organizational and End User Computing*, v.21, n.3, p.24-36, 2009. <http://dx.doi.org/10.4018/joeuc.2009070102>.
- PRIMACK, R.B. Variation in the phenology of natural populations of montane shrubs in New Zealand. *Journal of Ecology*, v.68, n.3, p.849-862, 1980.
- R CORE TEAM. R Development Core Team. *R: A Language and Environment for Statistical Computing*, 2019. <http://dx.doi.org/http://www.R-project.org>
- SAKO, Y.; MCDONALD, M.B.; FUJIMURA, K.; EVANS, A.F.; BENNETT, M.A. A system for automated seed vigor assessment. *Seed Science and Technology*, v.29, n.3, p.625-636, 2001. <https://www.eurofinsus.com/media/162083/seed-vigor-imaging-system.pdf>



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.