**Title:** An easy and automated calculation of the germination cardinal temperatures and thermal time using R

**Running Head:** Thermal Time & Cardinal Temperatures in R

Emma Ladouceur1,2, Hugh W. Pritchard3, Eduardo Fernández-Pascual3\*

1Museo Delle Scienze (Muse), Corso del Lavoro e Scienze, 3, Trento, Italy

2University of Pavia, Pavia, Italy

3Royal Botanic Gardens, Kew, Wellcome Trust Millennium Building, Wakehurst Place, West Sussex RH17 6TN, England

**\*Corresponding author:** [eduardofp.indurot@uniovi.es](mailto:eduardofp.indurot@uniovi.es); Tel.: +44(0)1444894184

**Keywords (7 max):** cardinal temperatures, thermal time models, thermal thresholds, seed traits, segmented model, piecewise regression, broken-stick regression

**Abstract (250 max)**

**Introduction**

Germination is the transition from seed to seedling, and occurs in response to a series of environmental cues including temperature, moisture, light and chemical signals (Bewley et al 2013). For each of these cues there is an optimal value at which the rate of germination is maximal. Below and above this optimum, the rate progressively decreases until the cue reaches base and ceiling thresholds beyond which germination stops. Thus, the rate of germination as a function of temperature can be described in non-dormant seeds as an accumulation of degrees-day above or below these thresholds (Garcia-Huidobro et al 1982). When a certain amount of degrees-day has been accumulated, the seed germinates. Every seed will require its own amount, but since experiments need to be done with a seed population, thermal time is expressed as the degrees-day needed to produce germination in a percentage of this population, usually the 50 %. Measuring the germination cue in degrees-day, i.e. in thermal time, has the advantage of integrating time and temperature (Romo and Eddelman 1995). Therefore, the thermal control of germination can be described in a mechanistic way using a handful of seed traits: the three cardinal temperatures, sub-optimal (base Tb), optimal (To) and supra-optimal (ceiling Tc); and the thermal time. Thermal time traits have two powerful applications. First, parameters calculated in a restricted set of experimental treatments can give good estimations of germination in more complex thermal environments (Hardegree et al 1999). Second, thermal time parameters can be compared in a standard way across seed individuals, populations and species (Trudgill et al 2000). As such, thermal time models provide the broadest base to conduct comparative studies of germination (Dürr et al 2015).

Here, we present a novel method to calculate the germination cardinal temperatures using segmented regression in the open software R. This method has three advantages over methods currently in use: 1. The computer identifies automatically the breaking point in the data (i.e. the separation between the suboptimal and supraoptimal germination temperature ranges), rather than the user estimating breaking points visually and on a case-by-case basis, which can lead to personal bias; 2. The analysis is computed in seconds, in contrast to the hours or days it may take using alternative methods; and 3. The method does not require the purchase of commercial statistical software. We test the method calculating cardinal temperatures in two mock species; one in which the three cardinal temperatures can be identified, and one in which experimental results did not lead to the successful identification of all the cardinal temperatures. We present here the format in which the data should be set up (Supplementary Information 1) and the script to analyse the data (Supplementary Information 2), so that out method can be completely transparent and rigorously reproducible by all seed researchers (Rocchini and Neteler 2012).

**Materials and Methods**

*Data preparation*

The two example datasets used here were assembled from the data presented by Fernández-Pascual et al. (2015, Annals of Botany). Describe here in more detail how the data needs to be formatted in the data file (columns, column names, etc). (Supplementary Information 1). The two species used here can also be substituted by different individuals, populations or experimental treatments, depending on the data being analysed. Although the example is presented with two species, the script is prepared to analyse larger numbers.

*R and R packages needed*

All analyses have been performed in the R language and environment for statistical computing (R Core Development Team (n.d.)) and all plots created using the ggplot2 package in R (Wickham 2009) (Supplementary Information 2).

*Step 1: Checking whether the data represent the full germination temperature range*

First, the package *dplyr* (Wickham and Francois 2016) was used to group data by germination treatment and replicate, and then filter the data into a new file to only represent the results of the final scoring date (i.e., the final germination proportions). Then, a function was created to estimate the mean final germination proportions and binomial confidence intervals using the package *binom* (Dorai-Raj (n.d.)), and *dplyr* used to apply the function per treatment. The first plots represent total mean germination of each treatment, plotted against temperature treatment (Figure 1). At this point, the *Figure 1* needs to be visually inspected for evidence that the temperature treatments used in the experiment resulted in representing the full germination temperature range of the study species species (Figure 1A). If this is the case, the segmented model we present here can be fitted and used to calculate the three cardinal temperatures. When the experimental treatments only gave results in either the sub- or the supraoptimal germination temperature range (Figure 1B), the segmented model cannot be fitted and only the base or ceiling temperatures can be calculated.

*Step 2: Estimating germination rates from the cumulative germination curves*

The package *drc* (Ritz et al 2015) was used to fit a dose-response model to the cumulative germination data. The package allows to fit different functions, and we have chosen by default the log-logistic because it gave the best fit in different trials. The user needs to check if this function gives the best fit to his/her own data. The script automatically tests the fit of different functions, and the user should use this information to change manually the chosen function (Supplementary material 2). Then the model with the best fit is used to plot the cumulative germination of each treatment against time, and the function fit needs to be checked visually (Figure 2). The models are used to calculate the time to reach successive deciles of germination (from 10 to 90 %) in each treatment. The inverse of the times is then calculated to have the germination rate of each temperature treatment.

*Step 3a: Fitting a segmented model to the full germination temperature range*

Where sub-optimal and supra-optimal temperatures are available in the data set, a segmented model can then be fitted (Figure 3a). A segmented models allows to identify breaking points in the data (i.e., the optimal germination temperature in which the germination rate is maximal). Then the model fits linear regressions separately to the segments defined by this breaking point (i.e., the suboptimal and supraoptimal germination temperature ranges) (Figure 3a). The package *segmented* (Vito (n.d.)) was used to set a function to fit a segmented regression and compute the intercept and the slope for each segmented relationship in the model. The functions requires the user to give an estimation of the breaking point, which can be based on the visual inspection of figures 1, 2 and 3. Once the model is fitted, the base temperature (Tb) is calculated by solving the suboptimal linear regression for the x-intercept (i.e., the temperature in which the value of the germination rate equals zero). An analogous procedure calculates the Tc using the supraoptimal linear regression. The inverse of the slopes of each regression gives the estimated thermal time for germination in degrees-day. The script calculates the cardinal temperatures and thermal times for each of the ten decile germination rates (10-90%). The fitted line for each segmented relationship were tested with the germination rate plotted against temperature for every decile. Figure 3a allows to the user to check visually the fit of the model to the data.

*Step 3b: Fitting a linear model to the sub- or supraoptimal germination temperature range*

It may be the case that the temperatures used in the experiment fall only on the suboptimal or supraoptimal germination temperature range (Figure 1b). In this case, the data does not allow to fit a segmented model and the three cardinal temperatures cannot be calculated. Nevertheless, a standard linear regression can be used to identify either the Tb or the Tc (Figure 3b). The inverse of the slope of this regression gives the thermal time for germination in degrees-day. We plot the germination rate against each temperature treatment using a smooth linear model for each decile in Figure 4b. Furthermore, it must be noted that this script can be used to calculate the germination base water potential. **Say somewhere that the script produces the figures 1 2 3 4 and the tables with the results**

**Results**

In *species A*, we generated data representing the full range of germination temperatures (Figure 1, 2, Supplementary Information 3 Table S1). This allowed to calculate germination rates in the suboptimal and supraoptimal temperature ranges (Table S2). A segmented regression could be fit to the data (Figure 3a), calculating the cardinal temperatures successfully (Table S3). In *species B,* the full range was not successfully identified (Table S1), but only the sub-optimal temperatures were represented in the data (Figure 1, 2). The supra-optimal temperatures were not identified, as after the best germination rate was reached at treatment X, higher temperatures did not produce any germination and had a total cumulative germination of zero (0). In this scenario it is impossible to determine the Tc nor the To. In *species B* the segmented model cannot be fit to the data (Figure 3 b), but a linear regression of the suboptimal temperature range can identify the Tb (Figure 4 b). The same script can be used if only supraoptimal temperatures are available, and in this case it will calculate the Tc.

**Discussion**

This method is appropriate for the analysis of germination data to identify cardinal temperatures and thermal time. The method we present here offers an opportunity to calculate the analyses rapidly, and to select the breaking point in the data without user bias. When using this script, we recommend that users become familiar with the basics to using R, and seek out one of the many introductory tutorials that exist online. We suggest to ensure all packages are installed, and libraries uploaded, and follow the notes in the script using the example data sets in the supplementary information, using this paper as a complimentary resource to the script.

To engage in an experiment that will set out to identify cardinal temperatures, one must plan an experimental design which will purposefully identify all the germination temperature range, with data points below the optimal temperature (sub-optimal) and above the optimal germination temperature (supra-optimal), pushing each species to its thermal thresholds. In order to fit a segmented curve, there must be a minimum of three suboptimal temperatures, and a minimum of three supra optimal temperatures. It is thus recommended that experimental designs include a minimum of 5-6 temperatures for each to effectively explore potential germination temperature ranges. The less that is known about a given species, the more temperatures that should be tested.

Constant temperatures are known to produce promising results for testing cardinal temperatures, particularly for agricultural varieties which have been bred for uniformity, and stable germination temperatures (Covell et al 1986, Ellis et al 1986). However many wild species, particularly alpine species, may require alternating temperatures, or alternating temperatures may better represent real ambient conditions found in nature (Carol C Baskin and Baskin 2014). When testing with alternating temperatures, it is recommended for plotting that the average temperature be taken of the two alternating temperatures for each treatment (Ellis and Barret 1994). Thus, alternating temperature regimes should be carefully planned to still result in temperature gradients when averaged, and may lead to odd results. Constant temperatures should never be compared to alternating temperatures within the same analysis, but if testing both, the two temperature regimes should be set up as two experiments, to compare the groups against each other. Ideally, in setting up an experimental design for wild species, a regime of both constant and alternating temperatures could be tested separately, and the results compared. Experiments are recommended to be run until cumulative germination stops or reaches a plateau. Scoring dates should be adjusted to the speed of germination of the study species, so a cumulative germination curves with good resolution are produced. Furthermore, it is important to consider that the cardinal temperatures are not fixed values but depend on the dormancy state of the seeds (Pritchard et al 1996).

We have found and confirmed that this method is useful to calculate thermal time and cardinal temperatures in germination experiments which have tested a full range of temperatures across a given species germination gradient, both rapidly and rigorously. We believe that this method will be of great value to generate cardinal temperature data for an ever increasing number of species. Physiological thresholds such as the cardinal temperatures are highly informative plant traits, and a key tool to integrate plant regeneration into multi-species community studies (Jiménez-Alfaro et al 2016, Funk and Larsson 2016).

**Acknowledgements**

Special thanks to the NAtive Seed Science TEchnology and Conservation (NASSTEC) Initial Training Network (ITN) consortium. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n°607785. E.F.P. had the financial support of the Government of Asturias and the FP7 – Marie Curie - COFUND programme of the European Commission (Grant ‘Clarín’ ACA14-19). The Royal Botanic Gardens, Kew, receive grant-in-aid from Defra.

**References**

Baskin CC and Baskin JM (2014) *Seeds: Ecology, Biogeography, and Evolution of Dormancy and Germination* (Second Edition.). Elsevier Inc., 1–1602.

Bewley DJ, Bradford KJ, Hilhorst HWM and Nonogaki H (2013) *Seeds Physiology of development Germination and Dormancy* (Third.). New York: Springer, 1–407.

Covell S, Ellis RH, Roberts EH and Summerfield RJ (1986) The Influence of Tmperature of Seed Germination Rate in Grain Legumes. *Journal of Experimental Botany* 37(178): 705–715.

Dorai-Raj S ((n.d.)) binom: Binomial Confidence Intervals For Several Parameterizations (1st edition). Available at: http://CRAN.R-project.org/package=binom.

Dürr C, Dickie JB, Yang XY and Pritchard HW (2015) Ranges of critical temperature and water potential values for the germination of species worldwide: Contribution to a seed trait database. *Agricultural and Forest Meteorology*. Elsevier B.V. 200: 222–232.

Ellis RH and Barret S (1994) Alternating Temperatures and Rate of Seed Germination in Lentil. *Annals of Botany* 74: 519–524.

Ellis RH, Covell S, Roberts EH and Summerfield RJ (1986) The Influence of Temperature on Seed Germination Rate in Grain Legumes. *Journal of Experimental Botany* 37(183): 1503–1515.

Garcia-Huidobro J, L MJ and Squires GR (1982) Time, Temperature and Germination of Pearl Millet (Pennisetum typhoides S. & H.). *Journal of Experimental Botany* 33(133): 288–296.

Hardegree SP, Van Vactor SS, Pierson FB and Palmquist DE (1999) Predicting variable-temperature response of non-dormant seeds from constant-temperature germination data. *Journal of Range Management* 52(1): 83–91.

R Core Development Team ((n.d.)) R: language and environment for statistical computing. Vienna, Austria. Available at: http://www.R-project.org/.

Ritz C, Baty F, Sreibig JC and Gerhard D (2015) Dose-Response Analysis Using R. *PLoS ONE*.

Rocchini D and Neteler M (2012) Let the four freedoms paradigm apply to ecology. *Trends in Ecology & Evolution*. Elsevier Ltd 27(6): 310–311.

Romo JT and Eddelman LE (1995) Use of Degree-Days in Multiple-Temperature Experiments. *Journal of Range Management* 48(5): 410–416.

Trudgill DL, Squire GR and Thompson K (2000) A thermal time basis for comparing the germination requirements of some British herbaceous plants. *New Phytologist*. Cambridge University Press 145(1): 107–114.

Vito MR ((n.d.)) Segmented: An R Package to Fit Regression Models with Broken-Line Relationships. *R News* 8(1): 20–25. Available at: http://cran.r-project.org/doc/Rnews/.

Wickham H (2009) *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wickham H and Francois R (2016) dplyr: A Grammar of Data Manipulation. Available at: http://CRAN.R-project.org/package=dplyr.



**Figure 1:** Final Germination proportions across all temperature treatments for species A) and species B)



**Figure 2:** Cumulative Germination Curves across all temperature treatments for species A) and species B)



**Figure 3:** Time to germination across each decile (%) of total germination, across each treatment using a segmented model for species A) and species B).



**Figure 4:** Time to germination across each decile (%) of total germination across each temperature treatment, using a smooth linear model for species A) and species B)