

Methods in Ecology and Evolution

MS NICOLA C STEER (Orcid ID : 0000-0002-6446-8601)

DR PAUL MICHAEL RAMSAY (Orcid ID : 0000-0001-7631-6480)

DR MIGUEL FRANCO (Orcid ID : 0000-0002-7249-4981)

Article type : Application

Editor : Samantha Price

nlstimedist: An R package for the biologically meaningful quantification of unimodal phenology distributions

Nicola C. Steer¹, Paul M. Ramsay¹ & Miguel Franco^{1*}

¹School of Biological and Marine Sciences, Plymouth University, Plymouth, PL4 8AA, UK

***Corresponding author:** School of Biological and Marine Sciences, Plymouth University, Plymouth, PL4 8AA, UK; M.Franco@plymouth.ac.uk; tel. +44-1752-584646.

Key words: phenology, time distribution, biologically interpretable parameters, model fitting, temporal dynamics, non-linear regression

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/2041-210X.13293

This article is protected by copyright. All rights reserved.

ABSTRACT

1. Phenological investigation can provide valuable insights into the ecological effects of climate change. Appropriate modelling of the time distribution of phenological events is key to determining the nature of any changes, as well as the driving mechanisms behind those changes.
2. Here we present the *nlstimedist* R package, a distribution function and modelling framework that describes the temporal dynamics of unimodal phenological events. The distribution function is derived from first principles and generates three biologically interpretable parameters.
3. Using seed germination at different temperatures as an example, we show how the influence of environmental factors on a phenological process can be determined from the quantitative model parameters.
4. The value of this model is its ability to represent various unimodal temporal processes statistically. The three intuitively meaningful parameters of the model can make useful comparisons between different time periods, geographical locations or species' populations, in turn allowing exploration of possible causes.

RESUMEN

1. La investigación de procesos fenológicos puede proporcionar información valiosa sobre los efectos ecológicos del cambio climático. El modelado de la distribución temporal de eventos fenológicos es clave para determinar la naturaleza de cualquier cambio, así como los mecanismos responsables.
2. Aquí presentamos el paquete `nlstimedist` R, una función de distribución y marco de modelado que describe la dinámica temporal de eventos fenológicos unimodales. La función es derivada a partir de principios básicos y consta de tres parámetros biológicamente interpretables.
3. Utilizando la germinación de semillas a diferentes temperaturas como ejemplo, ilustramos cómo la influencia de factores ambientales en un proceso fenológico es cuantificada por los parámetros del modelo.
4. El valor de este modelo es su capacidad para representar estadísticamente varios procesos temporales unimodales. Los tres parámetros del modelo tienen interpretación física simple y permiten hacer comparaciones útiles entre diferentes períodos, ubicaciones o poblaciones, lo que a su vez permite explorar posibles causas.

1 | INTRODUCTION

Periodically recurring, often seasonal, biological events (phenology) are influenced by environmental factors and interactions between organisms (Lieth 1974). Such phenomena are of particular interest because anthropogenic influences, such as climate change, might alter important ecological processes that are intimately correlated (Forrest & Miller-Rushing 2010). A biologically meaningful description of phenological events is essential to understanding their temporal dynamics, and offers an opportunity to assess the significance of its potential drivers (Rafferty *et al.* 2013).

Certain phenological events are recorded as a binary change from one recognisable state into another, either for a whole organism, as when a winter migrant has arrived for the breeding season (Gordo 2007) or for individual parts, as when individual leaf or flower buds on a plant burst (Cole & Sheldon 2017).

While it is clear that variability of individual events is expressed at the population level as a time distribution, phenological observations are often restricted to recording only extreme events, such as the date of the first flower to bloom (Fitter & Fitter 2002) or the first migrant of the season to arrive (Gordo & Sanz 2006). This approach ignores the population-level dynamics which contains a wealth of information regarding, for example, the duration of the phenomenon, its temporal skew, and its shape. Other scalar values may be conveniently-chosen thresholds (Zhang *et al.* 2003), such as the 50% of completion commonly used in the investigation of canopy phenology (Richardson *et al.* 2006), and varying percentage values are a key feature of the BBCH-scale used to identify phenological developmental stages in plants (Meier 2001). All of these approaches result in a single date, which is intended to capture useful information about the phenological process.

When single dates are used to describe the timing of a phenological event, they are often compared across years or linked to changes in an environmental condition, such as temperature, using regression (Sparks & Tryjanowski 2010). As useful as these scalars may be to summarise key features and changes in phenology, they inevitably miss potentially important information about the shape of the overall time course (Clark & Thompson 2011; CaraDonna, Iler & Inouye 2014; Carter, Saenz & Rudolf 2018).

A more thorough assessment should aim to model the entire phenological time distribution (CaraDonna, Iler & Inouye 2014; Carter, Saenz & Rudolf 2018). This is frequently accomplished using classic growth functions, such as the logistic and Richards (Yin *et al.* 2003; Zhang *et al.* 2003; Richardson *et al.* 2006; Sun & Frelich 2011) as their sigmoid shape resembles the time course of a phenological event. The logistic model is symmetrical around its point of inflection which is always halfway along the asymptotes (Birch 1999), but there is no theoretical basis for a phenological event to be symmetrical around its mid-point. The Richards (or generalised logistic) model is more flexible due to an additional shape parameter but its parameters cannot be interpreted in a meaningful way (Richards 1959; Zeide 1993; Birch 1999; Damgaard & Weiner 2008).

An alternative approach builds on an existing body of work on niche overlap (Pleasants 1980; Fleming & Partridge 1984; Totland 1993; Castro-Arellano *et al.* 2010), and allows species interactions to be compared as measures of temporal overlap (e.g., Carter, Saenz & Rudolf 2018). These approaches take account of whole phenological distributions through time, accommodating multimodal or skewed responses. Temporal overlap is an outcome of interactions between distributions rather than a direct consideration of their shapes, but is a sensible approach where the comparison focuses on time alone and where there are multimodal, complex probability distributions.

For unimodal phenology distributions, a model that describes the entire phenological time distribution well, is sufficiently flexible to accommodate asymmetrical distributions, and generates biologically interpretable shape parameters would be more useful. In particular, the model should be derived from basic principles applicable to a wide spectrum of biological time distributions. Importantly, goodness-of-fit alone should not be used to justify model selection; it is always preferable to choose a model that has biologically meaningful parameters (Paine *et al.* 2012).

Here, we present a model for describing the temporal dynamics of unimodal phenological events. It has been derived from first principles and generates biologically meaningful parameters that can be compared and used to assess potential driving mechanisms.

2 | THE MODEL

A phenological process of events (y) unfolding over time (x) at a constant rate (r) would follow an exponential distribution (Franco 2018). Phenological processes, however, do not occur at a constant rate (Sparks & Tryjanowski 2010) and individual events are more likely to be distributed according to a probabilistic process described by the inverse logit governed by an additional parameter, c (Franco 2018). Finally, phenological processes do not occur instantly, but happen sometime after exposure to a specific set of conditions (Wu *et al.* 2015), which requires a third parameter, the time lag, t . By incorporating the lagged form of the inverse logit function into the exponential distribution, a suitable biological time distribution can be derived (Franco 2018). This cumulative distribution function (*cdf*) has the form:

$$y = \left(1 - \frac{r}{1 + e^{-c(x-t)}}\right)^x$$

The derivative of this function quantifies the probability density function (*pdf*):

$$\frac{dy}{dx} = \left(1 - \frac{r}{1 + e^{-c(x-t)}}\right)^x \left(\frac{rcxe^{-c(x-t)}}{(1 + e^{-c(x-t)})^2 \left(1 - \frac{r}{1 + e^{-c(x-t)}}\right)} - \ln \left(1 - \frac{r}{1 + e^{-c(x-t)}}\right) \right)$$

which describes the population-level rate at which the phenomenon occurs. Each of the function's parameters has clear meaning and units: r quantifies the maximum proportional rate at which the process occurs (it is dimensionless); c is the rate at which r converges on its maximum value (units: time^{-1}); and t is an overall measure of the process' time lag (units: time) (Fig. 1). Parameter t can also be thought of as a weighted measure of the process' duration – weighted in relation to the values of r and c , that is. It correlates with, but is not equivalent to any of the distribution's various measures of central tendency.

3 | THE R PACKAGE

nlstimedist is an R package that provides a convenient way to fit the time course of a unimodal phenological time distribution employing nonlinear regression. *nlstimedist* combines functions for data preparation, model fitting and data visualisation into one complete package, allowing efficient, accurate and meaningful analysis.

The model is fitted to data using the **timedist()** function. The function requires data in the form of the proportion of cumulative number of events through time, together with column identifiers (allowing the analysis of multi-column data) and starting values for r , c and t . If data are in their raw form of counts versus time, they can be cleaned and converted to proportions (range: 0-1) for model use, using the built-in tidy function **tdData()**.

The **timedist()** function returns an object which contains all of the fitted model information.

This includes the equation used to fit the estimated time distribution, estimated values for r , c and t , the model fit's residual sum of squares, and the number of iterations to convergence.

The object can be examined with all of the generic *nls* functions, such as **summary()**, and can also be used by packages such as 'nlstools' (Baty *et al.* 2015). Functions and packages such

as these can be used to assess how well the model fits the data and the reliability of the parameter estimates. The statistical moments and percentiles of the fitted distribution can be obtained from the model object. The *nlstimedist* package also has two built-in functions for plotting the estimated time distribution as either a cumulative distribution function (*cdf*) **tdCdfPlot(object, ...)** or a probability density function (*pdf*) **tdPdfPlot(object, ...)**.

nlstimedist is based on the framework provided by nlsLM from the *minpack.lm* package (Elzhov *et al.* 2016). nlsLM is a modification of the standard *nls* function that uses the Levenberg-Marquardt algorithm (Marquardt 1963) for model fitting (Elzhov *et al.* 2016). This fitting procedure was chosen because it is considered robust (Lourakis 2005). Because the method of non-linear regression fitting uses an iterative optimisation procedure to converge on the least squares solution, fairly accurate starting values need to be chosen (Ruckstuhl 2010). *nlstimedist* is not a self-starting model, therefore guidelines are provided to assist with the selection of appropriate starting values for the three parameters (see package vignette).

Fitting to the underlying cumulative distribution function (opposed to the more usual practice of fitting a probability density function to binned data) allows datasets with few observations to be analysed. The temporal resolution of the data must be sensible and representative of the whole phenology under investigation. This model cannot be applied to complex, multimodal phenologies.

As shown in Fig. 1, each parameter has a unique effect on three different aspects of the distribution's shape. In summary, *r* is a scaled rate of completion (without units), *c* is a measure of its temporal concentration (units: time^{-1}), and *t* is an overall measure of temporal delay (units: time). In combination, these parameters provide insight into potential drivers and mechanisms associated with the phenological process, such as rates of development, climate change, competition between species, genetic diversity, resource availability and

environmental heterogeneity. Exploring the relationships between model parameters and statistical moments with biological and environmental variables might offer additional understanding of possible determinants.

4 | APPLICATION OF THE MODEL

The model can be applied across a wide range of phenological studies, including aspects of reproduction and development (e.g., pollination, gestation, egg-laying, egg-hatching, germination, life stages), seasonal population dynamics (of leaves, flowers, whole organisms, etc.), species interactions (trophic mismatch, predator-prey dynamics, competition, pest outbreaks), migration and dispersal (in relation to cues and invasion dynamics), and mortality in response to environmental challenge (climate change, ecotoxicology). The model has also been fitted successfully to the distribution of reproductive value of perennial plants as a means of quantifying the duration (by parameter t) and the speed (parameter c) of life (Mbeau-Ache & Franco 2013).

As a worked example, we present data from a controlled seed germination experiment for *Puya raimondii*, a giant rosette plant from the Andes. The experiment tested the effect of temperature on germination along a temperature gradient ranging from 8.4° C to 23.7° C. We use this example to illustrate how the new function is able to quantify accurately the changing temporal dynamics of a phenological process. We also show how quantification of the models parameters can be used to determine the influence that an environmental factor, in this case temperature, has on seed germination.

The dataset used in this example is available on the Dryad Digital Repository. The file can be read directly into R using the following command.

```
> Puya <- read.csv ("PuyaGermination.csv", header = TRUE, sep = ",")
```

To obtain the estimated parameter values (r , c , t) at each temperature, the model was fitted to each column in the “Puya Germination” dataset separately using the **timedist()** function.

Starting values for parameter estimates are dependent on the length of the time course under investigation and as such, starting values were adjusted for each model fit.

```
> Puya1.1 <- timedist (Puya, x = "x", y = "T8.4", r = 0.04, c = 0.5, t = 40)
```

```
> Puya1.1
```

```
Nonlinear regression model
```

```
model: T8.4 ~ 1 - (1 - (r/(1 + exp(-c * (x - t))))))^x
```

```
data: data
```

```
      r      c      t
```

```
0.07347 0.44714 37.36754
```

```
residual sum-of-squares: 0.008334
```

```
Number of iterations to convergence: 7
```

```
Achieved convergence tolerance: 1.49e-08
```

Fitting accuracy was verified using a range of functions. The reliability of the parameter estimates was obtained for each fit using the generic **summary()** function for *nls* objects.

Standard errors of parameter estimates were very small, and model fit was highly significant in all cases ($p < 0.001$; Table 1).

```
> summary (Puya1.1, correlation = TRUE, symbolic.cor = FALSE)
```

```
Formula: T8.4 ~ 1 - (1 - (r/(1 + exp(-c * (x - t))))))^x
```

Parameters:

	Estimate	Std. Error	t value	Pr(> t)	
r	0.073472	0.005016	14.65	4.51e-11	***
c	0.447144	0.024582	18.19	1.40e-12	***
t	37.367538	0.344737	108.39	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02214 on 17 degrees of freedom

Correlation of Parameter Estimates:

r	c
c	-0.64
t	0.92 -0.82

Number of iterations to convergence: 7

Achieved convergence tolerance: 1.49e-08

Nonlinear regression has no direct R^2 . However, a pseudo R^2 calculated as $1 - [\sum(y - \hat{y})^2 / \sum(y - \bar{y})^2]$, which defines a similar quantity for nonlinear regression and is able to describe the proportion of variance explained by the model (Kvålseth 1985; Cameron & Windmeijer 1997). Extracting this quantity from each model object provided another measure of how well the model fitted the data. R^2 was over 0.99 for all temperature treatments (Table 1), although we recommend caution in the interpretation of this statistic, as it provides an over-optimistic measure of fit (Spiess & Neumeyer 2010).

```
> Puya1.1$rss()
[1] 0.99693
```

The statistical moments and the percentiles of the distribution can also be extracted from each model object. These facilitate comparison of different temperature treatments throughout time (Table 1).

```
> Puya1.1$getMoments()
      mean variance      sd      skew kurtosis  entropy
1 35.66111 34.71189 5.891679 4.156171 36.20471 4.096621
```

Plotting the model fits as both cumulative distribution functions and probability density functions provides a useful summary of how germination is affected across a range of temperatures (Fig. 2). These plots provide an informative visual summary of the maximum per capita rate of germination, temporal spread and time delay of seed germination at each temperature.

A key feature of the model is the production of numerically meaningful parameter values. These parameters, when plotted against biological or environmental variables, allow potential driving mechanisms to be tested. In this example, temperature affected all three parameter estimates in a curvilinear fashion (Fig. 3). Parameters r , c and t displayed significant quadratic relationships with temperature, helping to identify the temperature at which germination was fastest, more concentrated and least delayed after sowing. This optimal temperature was remarkably similar for all three parameters: $r = 15.6^\circ \text{C}$, $c = 15.5^\circ \text{C}$ and $t = 15.9^\circ \text{C}$.

Although the quadratic relationship with temperature was significant and each parameter predicted similar optimal temperatures, there is no reason to expect either a similar optimum for all three parameters or a symmetrical response on either side of the optima. The analysis of other phenological processes may yield different statistical relationships. Temperature was used in this example to illustrate the effect that an environmental factor has on the time course of seed germination. However, the same principles would apply to other environmental conditions that vary on a continuous scale.

5 | CONCLUSIONS

The *nlstimedist* package was built to facilitate the application of Franco's (2018) distribution function to phenological data. The model adequately describes a unimodal phenological process of events that are usually recorded as completions, i.e., on a binary scale. It is conceptually simple and is able to capture the essence of a phenological process because its three parameters quantify properties of the distribution with known units: a maximum net per capita rate (dimensionless), a rate at which this maximum rate is achieved (units: time^{-1}) and an overall measure of the process' time lag (units: time). Both biological and environmental variables have been shown to affect the individual parameters in a predictable way (Mbeau-Ache & Franco 2013; Franco 2018), and examples provided here. The flexibility of the model in representing various continuous distributions, the interpretability of its parameters and its ability to estimate the underlying statistical distribution of an often highly asymmetrical temporal process make it a useful tool in the analysis of unimodal phenological phenomena.

6 | AUTHOR'S CONTRIBUTIONS

M.F. conceived the idea and designed the distribution function; P.R. collected the data; N.S. analysed the data and interpreted the results; N.S. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

7 | CITATION OF NLSTIMEDIST

Studies using *nlstimedist* should cite this article.

8 | DATA ACCESSIBILITY

The package is available on CRAN <https://cran.r-project.org/package=nlstimedist> and the data and R script used in this study are available on the Dryad Digital Repository <https://doi.org/10.5061/dryad.f01pr47> (Steer, Ramsay & Franco 2019).

9 | REFERENCES

- Baty, F., Ritz, C., Charles, S., Brutsche, M., Flandrois, J.-P. & Delignette-Muller, M.-L. (2015) A toolbox for nonlinear regression in R: the package nlstools. *Journal of Statistical Software*, **66**, 1-21.
- Birch, C.P.D. (1999) A new generalized logistic sigmoid growth equation compared with the Richards growth equation. *Annals of Botany*, **83**, 713-723.
- Cameron, C.A. & Windmeijer, F.A.G. (1997) An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics*, **77**.
- CaraDonna, P.J., Iler, A.M. & Inouye, D.W. (2014) Shifts in flowering phenology reshape a subalpine plant community. *Proceedings of the National Academy of Sciences*, **111**, 4916-4921.
- Carter, S.K., Saenz, D. & Rudolf, V.H.W. (2018) Shifts in phenological distributions reshape interaction potential in natural communities. *Ecology Letters*, **21**, 1143-1151.

- Castro-Arellano, I., Lacher Jr, T.E., Willig, M.R. & Rangel, T.F. (2010) Assessment of assemblage-wide temporal niche segregation using null models. *Methods in Ecology and Evolution*, **1**, 311-318.
- Clark, R.M. & Thompson, R. (2011) Sampling bias in the determination of first and last occurrences. *Plant Ecology & Diversity*, **4**, 201-211.
- Cole, E.F. & Sheldon, B.C. (2017) The shifting phenological landscape: Within- and between-species variation in leaf emergence in a mixed-deciduous woodland. *Ecology and Evolution*, **7**, 1135-1147.
- Damgaard, C. & Weiner, J. (2008) Modeling the growth of individuals in crowded plant populations. *Journal of Plant Ecology*, **1**, 111-116.
- Elzhov, T.V., Mullen, K.M., Spiess, A.-N. & Bolker, B. (2016) Minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in Minpack, Plus Support for Bounds.
- Fitter, A.H. & Fitter, R.S.R. (2002) Rapid changes in flowering time in British plants. *Science*, **296**, 1689-1691.
- Fleming, T.H. & Partridge, B.L. (1984) On the analysis of phenological overlap. *Oecologia*, **62**, 344-350.
- Forrest, J. & Miller-Rushing, A.J. (2010) Toward a synthetic understanding of the role of phenology in ecology and evolution. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **365**, 3101-3112.
- Franco, M. (2018) The time distribution of biological phenomena – illustrated with the London marathon. *PeerJ Preprints* 6:e27175v1.
- Gordo, O. (2007) Why are bird migration dates shifting? A review of weather and climate effects on avian migratory phenology. *Climate Research*, **35**, 37-58.
- Gordo, O. & Sanz, J.J. (2006) Climate change and bird phenology: a long-term study in the Iberian Peninsula. *Global Change Biology*, **12**, 1993-2004.
- Kvålseth, T.O. (1985) Cautionary Note about R2. *The American Statistician*, **39**, 279-285.
- Lieth, H. (1974) *Phenology and Seasonality Modeling*. Springer-Verlag, Berlin.
- Lourakis, M.I.A. (2005) A brief description of the Levenberg-Marquardt algorithm implemented by levmar. *Foundation of Research and Technology*, **4**, 1-6.
- Marquardt, D.W. (1963) An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, **11**, 431-441.
- Mbeau-Ache, C. & Franco, M. (2013) The time distribution of reproductive value measures the pace of life. *Journal of Ecology*, **101**, 1273-1280.
- Meier, U. (2001) Growth stages of mono- and dicotyledonous plants: BBCH-Monograph. *Federal Biological Research Centre for Agriculture and Forestry*.
- Paine, C.E.T., Marthens, T.R., Vogt, D.R., Purves, D., Rees, M., Hector, A. & Turnbull, L.A. (2012) How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. *Methods in Ecology and Evolution*, **3**, 245-256.

- Pleasants, J.M. (1980) Competition for bumblebee pollinators in Rocky Mountain plant communities. *Ecology*, **61**, 1446-1459.
- Rafferty, N.E., Caradonna, P.J., Burkle, L.A., Iler, A.M. & Bronstein, J.L. (2013) Phenological overlap of interacting species in a changing climate: an assessment of available approaches. *Ecology and Evolution*, **3**, 3183-3193.
- Richards, F.J. (1959) A Flexible Growth Function for Empirical Use. *Journal of Experimental Botany*, **10**, 290-300.
- Richardson, A.D., Bailey, A.S., Denny, E.G., Martin, C.W. & O'Keefe, J. (2006) Phenology of a northern hardwood forest canopy. *Global Change Biology*, **12**, 1174-1188.
- Ruckstuhl, A. (2010) Introduction to nonlinear regression. *IDP Institut für Datenanalyse und Prozessdesign, Zürcher Hochschule für Angewandte Wissenschaften*.
- Sparks, T.H. & Tryjanowski, P. (2010) Regression and causality. *Phenological Research* (eds I. Hudson & M. Keatley), pp. 123-145. Springer, Dordrecht.
- Spieß, A.-N. & Neumeyer, N. (2010) An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC Pharmacology*, **10**, 6-6.
- Steer, N.C., Ramsay, P.M. & Franco, M. (2019) Data from: nlstimedist: an R package for the biologically meaningful quantification of unimodal phenology distributions. *Dryad Digital Repository*. doi:10.5061/dryad.f01pr47.
- Sun, S. & Frelich, L.E. (2011) Flowering phenology and height growth pattern are associated with maximum plant height, relative growth rate and stem tissue mass density in herbaceous grassland species. *Journal of Ecology*, **99**, 991-1000.
- Totland, Ø. (1993) Pollination in alpine Norway: flowering phenology, insect visitors, and visitation rates in two plant communities. *Canadian Journal of Botany*, **71**, 1072-1079.
- Wu, D., Zhao, X., Liang, S., Zhou, T., Huang, K., Tang, B. & Zhao, W. (2015) Time-lag effects of global vegetation responses to climate change. *Global Change Biology*, **21**, 3520-3531.
- Yin, X.Y., Goudriaan, J., Lantinga, E.A., Vos, J. & Spiertz, H.J. (2003) A flexible sigmoid function of determinate growth. *Annals of Botany*, **91**, 361-371.
- Zeide, B. (1993) Analysis of growth equations. *Forest Science*, **39**, 594-616.
- Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C., Gao, F., Reed, B.C. & Huete, A. (2003) Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, **84**, 471-475.

Table 1. Number of seeds germinated (N) and percentage of germination (y_{max}) in each temperature category, estimated parameter values (with standard errors in parenthesis), *** $p < 0.001$, proportion of variance explained by the model (R^2) and statistical moments for each of the predicted distributions.

Temp. (°C)	N	y_{max} (%)	r (s.e.)	Sig.	c (s.e.)	Sig.	t (s.e.)	Sig.	R^2	Mean	SD	Skew	Kurtosis	Entropy
8.4	148	74.0	0.073 (0.005)	***	0.447 (0.025)	***	37.368 (0.345)	***	99.7	35.661	5.892	4.156	36.205	4.097
9.3	156	78.0	0.075 (0.003)	***	0.653 (0.032)	***	29.532 (0.158)	***	99.8	29.461	6.334	4.848	37.191	3.823
12.5	161	80.5	0.112 (0.008)	***	0.806 (0.062)	***	22.018 (0.230)	***	99.6	21.421	3.925	4.621	38.424	3.354
13.8	164	82.0	0.129 (0.004)	***	1.485 (0.083)	***	16.133 (0.071)	***	99.9	16.360	3.571	5.188	39.817	2.748
14.7	160	80.0	0.126 (0.006)	***	1.418 (0.117)	***	15.113 (0.104)	***	99.7	15.580	3.988	4.774	33.338	2.911
16.7	147	73.5	0.134 (0.003)	***	2.230 (0.108)	***	13.992 (0.037)	***	99.9	14.597	3.639	5.116	36.541	2.393
17.6	157	78.5	0.139 (0.005)	***	1.917 (0.158)	***	14.028 (0.074)	***	99.8	14.452	3.418	5.144	37.802	2.500
19.5	159	79.5	0.121 (0.008)	***	0.801 (0.080)	***	15.970 (0.226)	***	99.3	16.061	4.431	3.893	25.498	3.526
20.0	155	77.5	0.090 (0.004)	***	0.487 (0.041)	***	17.896 (0.259)	***	99.4	18.775	7.089	3.218	17.477	4.307
21.7	146	73.0	0.080 (0.003)	***	0.504 (0.025)	***	25.638 (0.212)	***	99.8	25.516	6.712	3.898	26.187	4.154
22.4	144	72.0	0.058 (0.003)	***	0.283 (0.024)	***	30.436 (0.560)	***	99.1	30.989	10.882	3.133	17.629	4.992
23.7	88	44.0	0.052 (0.005)	***	0.201 (0.015)	***	43.426 (0.992)	***	99.2	41.159	11.452	2.641	16.558	5.242

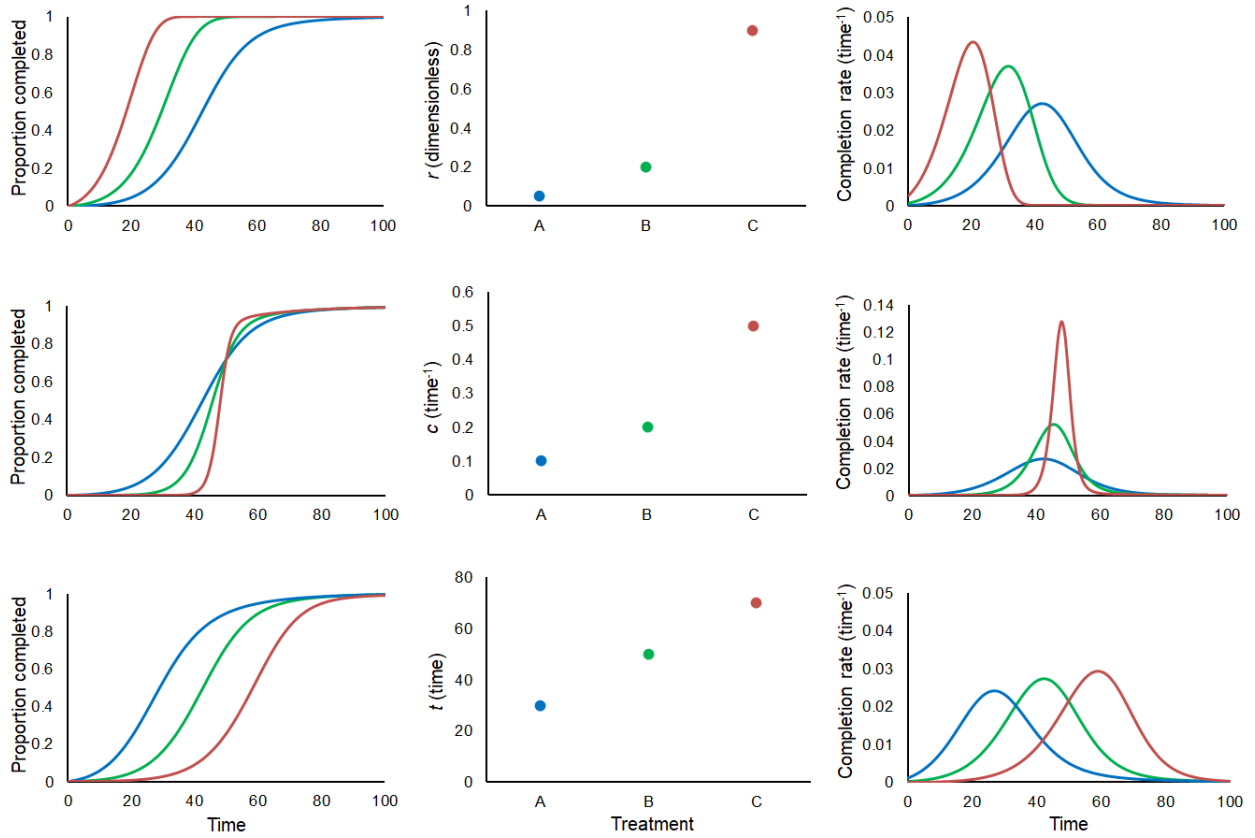


Figure 1. The influence of the three model parameters (r , c , and t) on the cumulative distribution function (left panels) and probability density function (right panels). The central panels show how each parameter varies, while the other two are held constant.

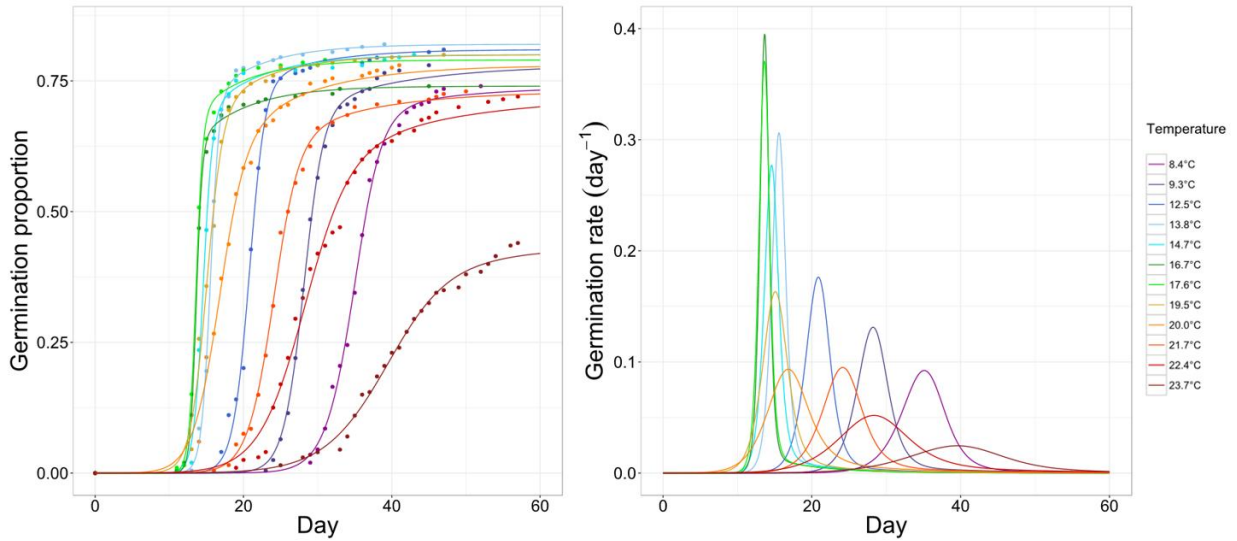


Figure 2. Cumulative distribution functions (left) and corresponding probability density functions (right) for *Puya raimondii* germination occurring along a temperature gradient ranging from 8.4°C to 23.7°C. Probability density functions describe the population-level rate of germination and the area under each curve is equal to the maximum percentage of germination.

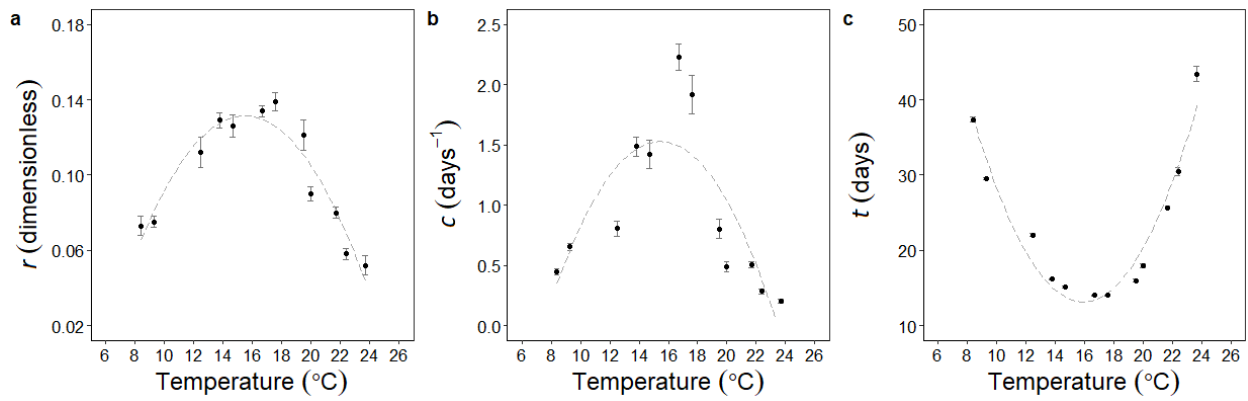


Figure 3. The relationship between temperature and the values of parameter estimates

produced from each model fit (a) parameter r , (b) parameter c , and (c) parameter t . All three

quadratic relationships were significant (a) $R^2 = 0.915$, $p < 0.000$, (b) $R^2 = 0.672$, $p = 0.007$,

(c) $R^2 = 0.945$, $p < 0.000$. Error bars represent the standard errors of parameter estimates.