

'biogeom': An R package for simulating and fitting natural shapes

Peijian Shi¹  | Johan Gielis²  | Brady K. Quinn³  | Karl J. Niklas⁴  |
David A. Ratkowsky⁵  | Julian Schrader^{6,7}  | Honghua Ruan¹ | Lin Wang¹ |
Ülo Niinemets^{8,9} 

¹Co-Innovation Centre for Sustainable Forestry in Southern China, Bamboo Research Institute, College of Biology and the Environment, Nanjing Forestry University, Nanjing, China

²Department of Biosciences Engineering, University of Antwerp, Antwerp, Belgium

³St. Andrews Biological Station, Fisheries and Oceans Canada, St. Andrews, New Brunswick, Canada

⁴School of Integrative Plant Science, Cornell University, Ithaca, New York, USA

⁵Tasmanian Institute of Agriculture, University of Tasmania, Hobart, Tasmania, Australia

⁶School of Natural Sciences, Macquarie University, Sydney, New South Wales, Australia

⁷Biodiversity, Macroecology & Biogeography, University of Göttingen, Göttingen, Germany

⁸Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Tartu, Estonia

⁹Estonian Academy of Sciences, Tallinn, Estonia

Correspondence

Peijian Shi, Bamboo Research Institute,
College of Biology and the Environment,
Nanjing Forestry University, Nanjing 210037,
China.

Email: pjshi@njfu.edu.cn

Julian Schrader, School of Natural Sciences,
Macquarie University, Sydney, NSW 2109,
Australia.

Email: jschrad@uni-goettingen.de

Funding information

German Research Foundation, Grant/Award
Number: SCHR1672/1-1; National Key
Research and Development Program of China,
Grant/Award Number: 2021YFD02200403

Abstract

Many natural objects exhibit radial or axial symmetry in a single plane. However, a universal tool for simulating and fitting the shapes of such objects is lacking. Herein, we present an R package called 'biogeom' that simulates and fits many shapes found in nature. The package incorporates novel universal parametric equations that generate the profiles of bird eggs, flowers, linear and lanceolate leaves, seeds, starfish, and tree-rings, and three growth-rate equations that generate the profiles of ovate leaves and the ontogenetic growth curves of animals and plants. 'biogeom' includes several empirical datasets comprising the boundary coordinates of bird eggs, fruits, lanceolate and ovate leaves, tree rings, seeds, and sea stars. The package can also be applied to other kinds of natural shapes similar to those in the datasets. In addition, the package includes sigmoid curves derived from the three growth-rate equations, which can be used to model animal and plant growth trajectories and predict the times associated with maximum growth rate. 'biogeom' can quantify the intra- or interspecific similarity of natural outlines, and it provides quantitative information of shape and ontogenetic modification of shape with important ecological and evolutionary implications for the growth and form of the living world.

KEYWORDS

boundary coordinates, Gielis equation, ovate leaf shape, sigmoid curve, symmetry

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals LLC on behalf of New York Academy of Sciences.

INTRODUCTION

Many natural objects exhibit some degree of symmetry or handedness, features that have garnered the attention of many biologists and geometers.^{1–6} Researchers have long attempted to develop special models for describing the shapes of certain types of biological objects (e.g., Refs. 4 and 6). However, efforts to build a universal geometric formula for a wide range of natural shapes have been largely unsuccessful. A notable exception is Gielis,³ who created a universal formula (henceforth referred to as the Gielis equation) that can simulate a wide range of diverse shapes, especially those with symmetrical features.⁷ However, the validity of the Gielis equation was not tested using actual data from natural geometries until 12 years later, when Shi *et al.*⁸ carried out a pioneering study modeling the leaf boundaries of four species of bamboo (Poaceae: Bambusoideae). The original version of the Gielis equation has seven parameters and it has been simplified for fitting different natural geometries to reduce its complexity. Subsequently, several studies have been published testing the validity of the Gielis equation using empirical data of leaves, seeds, and fruits, bird eggs, and sea stars.^{9–16} All of these studies have validated the universality of the Gielis equation in describing natural geometries. In addition, a large number of other natural geometries have been simulated by this equation, although the validity of the Gielis equation in fitting those shapes has not been examined thus far.¹⁷ In addition, when studying certain symmetrical geometries (i.e., sea stars with pentaradial symmetry and fruit capsules with triradial symmetry), a “twin” version of the Gielis equation has been proposed, which provides a better goodness-of-fit and a better representation of radial symmetry.^{12,15} Although neither the Gielis equation nor its twin version have been widely applied in ecological and evolutionary research, these equations are apparently useful as parametric models. Their simplified versions include model parameters that can be used as quantitative traits to explore potential intra- and interspecific variation in shapes and as such can provide key insights into the ecological role of shape variation and the evolution of shape.^{9,10,15} The difficulty in the application of all such geometric equations in biology is the limited availability of software packages to carry out data simulation and fitting.

Although the Gielis equation and its simplified versions can describe and fit many symmetrical or spirally arranged natural geometries, the models themselves cannot reveal the extent of the (a)symmetry exhibited by such shapes. For bilaterally symmetrical objects, previous studies have indicated that taking some landmarks on the boundary of an object to measure the extent of its (a)symmetry is particularly helpful.² Nevertheless, for many objects with unbroken boundaries, for example, bird eggs and ovate leaves, it is difficult to accurately define the corresponding landmarks along the axis of symmetry. Shi *et al.*¹⁸ previously proposed a standardized index for measuring the bilateral (a)symmetry of objects. However, this method has not been included in any software packages so far. In addition, because of the high degree of leaf-shape variation for ovate leaves (i.e., variations in the curvature of leaf boundaries across different plant species), this class of leaf shapes cannot be accurately represented by the

Gielis equation, and additional equations are needed to describe such variations.¹⁹

In thermal biology, many temperature-dependent development models are used to reflect the effect of temperature on organismal developmental rates (i.e., the reciprocal of the developmental time required to complete a specific developmental stage). Prior studies have shown that the integral forms of temperature-dependent developmental (or growth) rate models can produce sigmoid growth curves, which can be used to describe the growth trajectories of many animals and plants.^{20–22} In such cases, “time” replaces “temperature” as the independent (x-axis) variable. However, the integral forms of the more complex developmental rate models have no clear mathematical expressions. It is useful, therefore, to use the numerical integral approach to estimate the parameters of integrated developmental rate models. This also requires a software package to combine and carry out these complex calculations.

To aid researchers in simulating and fitting biological shapes, we have developed an R package called ‘biogeo’. In addition, this package has other functions related to biological geometries, including one allowing the calculation of the fractal dimension of a leaf venation network^{23,24} and a recently published egg-shape model.⁶

SUBJECT MODELS AND DATA FITTING METHODS

The Gielis equation and its twin version

The superellipse equation is an extension of the ellipse equation, but it can produce a more diverse array of variable curves, from diamonds to rectangles (Figure 1A). The superellipse equation was initially represented as:

$$\left|\frac{x}{A}\right|^n + \left|\frac{y}{B}\right|^n = 1, \quad (1)$$

where x and y represent the abscissa and ordinate of a curve transcribing a shape in the Cartesian coordinate system, and A and B are the major semi-axis and minor semi-axis of the superellipse, respectively. In the polar coordinate system, Equation (1) can be expressed as:

$$r = a \left(|\cos \varphi|^n + \left| \frac{1}{k} \sin \varphi \right|^n \right)^{-\frac{1}{n}}, \quad (2)$$

where r is the polar radius at the polar angle φ , $a = A$, and $k = B/A$. This reparameterized equation reduces the parameter-effects curvature and increases the close-to-linear behavior of the equation during nonlinear fitting.^{25,26}

Gielis³ extended the superellipse equation to take the form:

$$r = \left(\left| \frac{1}{A} \cos \frac{m}{4} \varphi \right|^{n_2} + \left| \frac{1}{B} \sin \frac{m}{4} \varphi \right|^{n_3} \right)^{-\frac{1}{n_1}}, \quad (3)$$

where m is a positive integer that determines the number of angles of the curve within $[0, 2\pi)$, and A , B , n_1 , n_2 , and n_3 are parameters to be

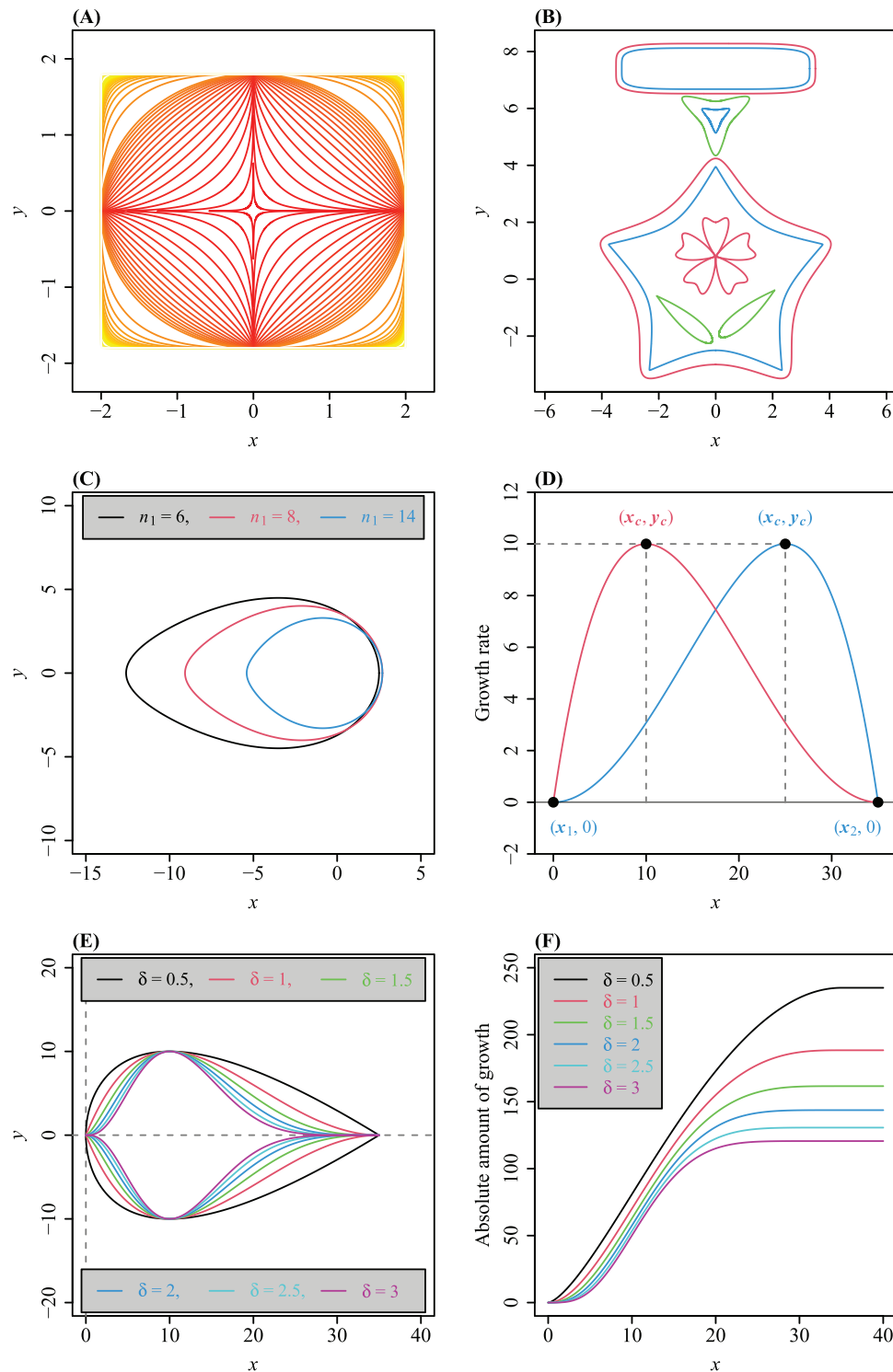


FIGURE 1 Shapes simulated by the 'biogeom' package. (A) A group of simulated superellipse curves based on Equation (2), that is, the superellipse equation. (B) Medal outlines simulated based on the Gielis equation and its simplified versions. (C) Three egg shapes simulated based on a simplified version of the Gielis equation. (D) The right-skewed (red) and left-skewed (blue) growth-rate curves based on Equation (9), that is, the modified Lobry–Rosso–Flandrois equation (MLRFE). (E) A group of ovate leaf shapes simulated based on two axisymmetrical curves generated by the MLRFE. (F) A group of sigmoid growth curves (representing the absolute growth amount of a measure of interest, e.g., organismal size, and population density) based on the numerical integration of the MLRFE. The detailed model parameters for each panel can be found in Appendix S1.

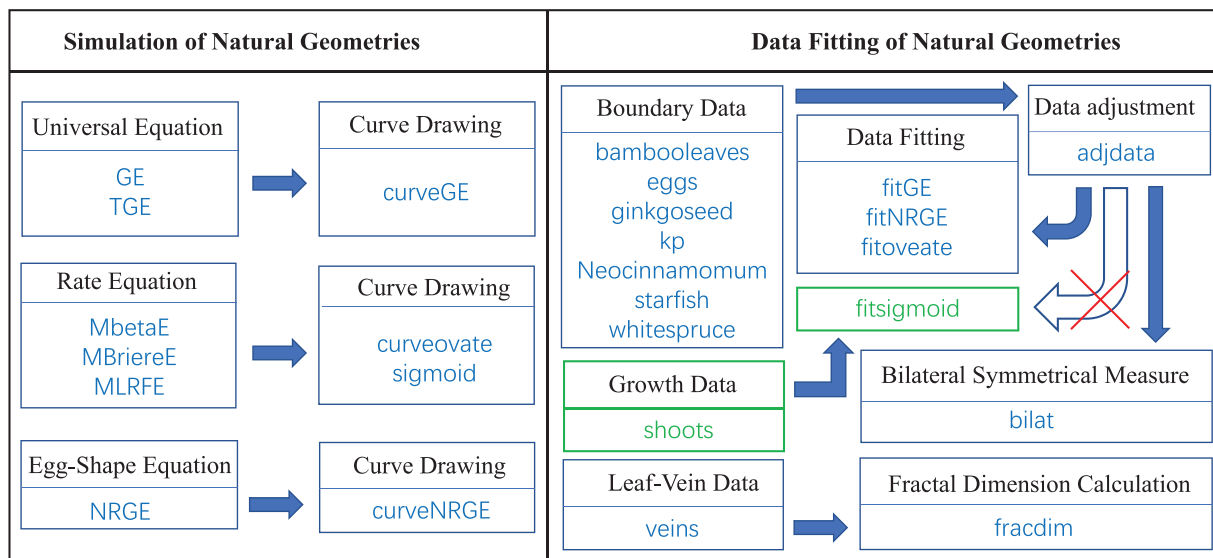


FIGURE 2 Framework of the 'biogeom' package. There are two modules: one for simulating natural geometries, and another for fitting natural geometries.

estimated. We have typically used this equation in a re-expressed form:

$$r = a \left(\left| \cos \frac{m}{4} \varphi \right|^{n_2} + \left| \frac{1}{k} \sin \frac{m}{4} \varphi \right|^{n_3} \right)^{-\frac{1}{n_1}}, \quad (4)$$

where $a = A^{n_2/n_1}$ and $k = B/A^{n_2/n_3}$. For convenience, Equation (4) is referred to as the *Gielis equation*. When $n_2 = n_3$, the Gielis equation can produce curves with rich axial or radial symmetry within $[0, 2\pi)$. Figure 2B shows the use of several Gielis curves to form a medal's profile, and Figure 2C shows three egg shapes produced by a simplified Gielis equation.¹⁶

The polar radius of the Gielis equation (r) is a linear function of the elementary polar radius (r_e) on a log-log scale, and the slope is equal to $-1/n_1$. This can be seen when we define an elementary polar radius (r_e) as:

$$r_e = \left| \cos \frac{m}{4} \varphi \right|^{n_2} + \left| \frac{1}{k} \sin \frac{m}{4} \varphi \right|^{n_3}. \quad (5)$$

Shi *et al.*¹² proposed a hyperbolic relationship between the Gielis and elementary polar radii to better describe the shapes of some sea stars:

$$\ln(r) = \frac{1}{\alpha + \beta \cdot \ln(r_e)} + \gamma, \quad (6)$$

where α , β , and γ are parameters to be estimated. We refer to Equation (6) as the twin Gielis equation.

Rate equations and their extensions

'biogeom' combines three rate equations because the Gielis equation and its twin version cannot describe all natural shapes. For common

ovate leaves, we need to use a rate equation to form a leaf-shape model. When two axially symmetrical rate equations along the x-axis are combined into one, we find that it can describe and fit a variety of ovate leaf shapes especially well.¹⁹ Therefore, we added three rate equations to 'biogeom'. The three rate equations can be used to produce sigmoid curves by integration.^{20–22} However, there are no explicit analytical solutions for their integrated forms (i.e., the ontogenetic equations based on the rate equations). Therefore, we also provided a numerical fitting method to estimate the parameters of the rate equations.

The effect of temperature on the development (or growth) rate of organisms is represented as a left-skewed curve. In the low-temperature range, the rate increases exponentially from a lower threshold temperature above which development (or growth) begins. In the mid-temperature range, the rate increases linearly toward a maximum value. And in the high-temperature region, the rate sharply decreases until an upper (sometimes lethal) threshold temperature is reached, above which development (or growth) ceases (Figure 1C).^{27,28} Many temperature-dependent developmental (or growth) rate equations are used to describe the effect of temperature on the developmental rates of arthropods or the growth rates of microbial populations.^{29–33} The 'biogeom' package includes three of these models: the beta equation,³⁴ the Brière equation,³⁵ and the Lobry-Rosso-Flandrois equation,^{36,37} and their generalized versions. All three of these models include the lower and upper threshold temperatures as model parameters. Interestingly, the beta and Lobry-Rosso-Flandrois models have the same number of parameters that have similar biological and geometrical meanings.²¹ Herein, we used the original Lobry-Rosso-Flandrois equation^{36,37} to plot the shape of a curve (Figure 1D) using the equation:

$$y = \frac{y_c (x - x_2) (x - x_1)^2}{(x_c - x_1) [(x_c - x_1) (x - x_c) - (x_c - x_2) (x_c + x_1 - 2x)]}, \quad (7)$$

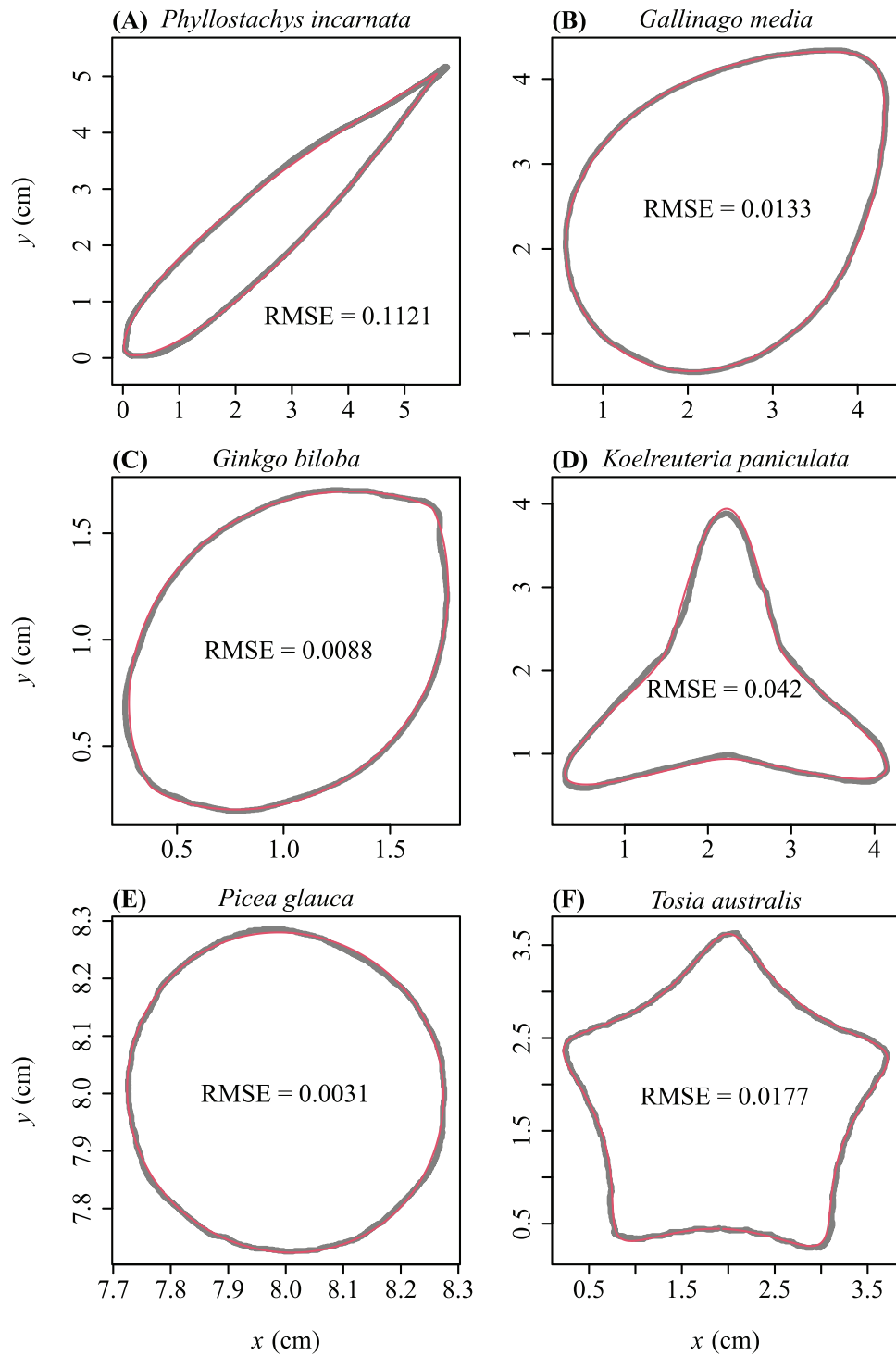


FIGURE 3 Observed (gray) and predicted (red) geometries (boundary coordinates). (A) A bamboo leaf. (B) A bird egg. (C) The side projection of a ginkgo seed. (D) The vertical projection of a fruit of the golden rain tree. (E) Tree rings in a cross-section of the stem of a white spruce tree. (F) A sea star. RMSE is the root-mean-square error between the observed and predicted polar radii. The estimated model parameters can be found in Appendix S1. These geometries were all considered to fit the Gielis equation, because the superellipse equation (panel A) is a special case of the Gielis equation (i.e., Equation 4).

where y represents the developmental (or growth) rate at temperature x ($^{\circ}\text{C}$); y_c represents the maximum rate, occurring at temperature x_c ; x_1 and x_2 represent the lower and upper threshold temperatures for development (or growth); and $y = 0$ when $x > x_2$ or $x < x_1$. When replacing temperature with time in Equation (7), we obtain a growth rate versus time curve. Equation (7) only produces a left-skewed curve when $\frac{(x_1+x_2)}{2} \leq x \leq x_2$.^{36,37} When $x_1 \leq x \leq \frac{x_1+x_2}{2}$, it is necessary to use the following equation (Figure 1C):²¹

$$y = \frac{y_c (x - x_1) (x - x_2)^2}{(x_2 - x_c) [(x_2 - x_c) (x - x_c) - (x_1 - x_c) (x_c + x_2 - 2x)]}. \quad (8)$$

To increase the versatility of the Lobry-Rosso-Flandrois equation in curve fitting (i.e., to increase the range of curves with different curvatures that can be fitted using Equation 8), an additional parameter, δ , was introduced:²¹

$$y = \begin{cases} y_c \left\{ \frac{(x-x_1)(x-x_2)^2}{(x_2-x_c)[(x_2-x_c)(x-x_c)-(x_1-x_c)(x_c+x_2-2x)]} \right\}^{\delta}, & \text{if } x \in \left[x_1, \frac{x_1+x_2}{2} \right]; \\ y_c \left\{ \frac{(x-x_2)(x-x_1)^2}{(x_c-x_1)[(x_c-x_1)(x-x_c)-(x_c-x_2)(x_c+x_1-2x)]} \right\}^{\delta}, & \text{if } x \in \left[\frac{x_1+x_2}{2}, x_2 \right]; \\ 0, & \text{if } x \notin [x_1, x_2]. \end{cases} \quad (9)$$

An ovate leaf shape can be formed by setting a symmetrical function of Equation (9) around the x -axis, that is, $f(x) = -y$. Figure 1E shows the effect of δ on the shape of this curve. Hereafter, we refer to Equation (9) as the modified Lobry-Rosso-Flandrois equation (MLRFE) for convenience. The modified beta and Br  re equations were obtained in a similar way by adding the parameter δ to improve the elasticity of the original beta and Bri  re equations in curve fitting.

We also derived a sigmoid growth equation to describe the growth trajectories of animals and plants using the integral form of a rate equation, that is:

$$y = \begin{cases} \int_{x_1}^x f(x) dx, & \text{if } x \in [x_1, x_2]; \\ 0, & \text{if } x < x_1; \\ \int_{x_1}^{x_2} f(x) dx, & \text{if } x > x_2. \end{cases} \quad (10)$$

Figure 1F shows the sigmoid growth curves plotted with different δ values based on the MLRFE. Here, the x -axis of any rate equation used to describe and fit ovate leaf shape does not represent temperature or time. Rather, it represents the distance from the leaf base (i.e., the connection between a leaf lamina and the petiole) to a point along the leaf length axis.

Data fitting methods

To estimate the parameters of the Gielis equation, the twin Gielis equation, and their simplified versions, it is necessary to know the planar coordinates of a scanned or photographed natural object. These coordinates can be obtained using the protocols proposed in

Refs. 8, 10, and 11. In a standard Gielis curve, the polar point is located at the origin (0, 0), and the major axis (defined as the straight line through the two points, $(r(\varphi = 0), 0)$ and $(r(\varphi = \pi), 0)$) is superimposed on the x -axis. However, many scanned or photographed shapes deviate from the standard forms predicted by the Gielis equation or its extended or simplified versions. Thus, three additional location parameters (i.e., the x - and y -coordinates of the polar point and the angle between the major axis and the x -axis) are needed when carrying out the nonlinear fitting procedure.¹⁶ We aimed to minimize the residual sum of squares (RSS) between the observed (r_i) and the predicted ($\hat{r}_i$) polar radii to estimate the model parameters of the Gielis equation or its extended or simplified versions:

$$\text{RSS} = \sum_{i=1}^N (r_i - \hat{r}_i)^2, \quad (11)$$

where the subscript i represents the i -th data point on the boundary of an image of interest, and N represents the number of data points on the boundary. The Nelder-Mead optimization approach³⁸ is used to minimize the RSS to obtain the numerical values of model parameters.

To estimate the parameters of the ovate leaf-shape models and the sigmoid growth equations, it is also necessary to extract the planar coordinates of the boundary of an image. The RSS between the observations of the y -coordinates (y_i) and the predicted y -coordinates ($\hat{y}_i$) are subsequently minimized using the Nelder-Mead optimization approach,³⁸ that is:

$$\text{RSS} = \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (12)$$

where i and N are the same as those in Equation (11).

FRAMEWORK OF BIOGEOM AND ITS FUNCTIONS

There are two main modules in 'biogeom', one for simulating and another for fitting natural geometries (Figure 2). The former module contains three classes of functions. The first class comprises universal equations and associated utilities. The 'GE' and 'TGE' functions are used to calculate the polar radii at different polar angles using the Gielis and twin Gielis equations, respectively (see Equations 4 and 6). Some simplified versions of the Gielis and twin Gielis equations can be specified by changing the option of the simpver argument. The 'curveGE' function is used to produce plots of the Gielis curves based on the 'GE' and 'TGE' functions. The second class of functions comprises growth-rate functions and their extensions, specifically the modified beta equation ('MbetaE'), modified Bri  re equation ('MBriereE'), and 'MLRFE'. The 'curveovate' function is used to generate the shape of an ovate leaf based on one of the above three rate equations. The 'sigmoid' function is used to generate sigmoid growth curves that reflect the growth trajectories (e.g., mass, height, or length) of animals or plants. The third function class in the shape module comprises the egg-shape model because bird eggs are a common natural shape that have attracted a

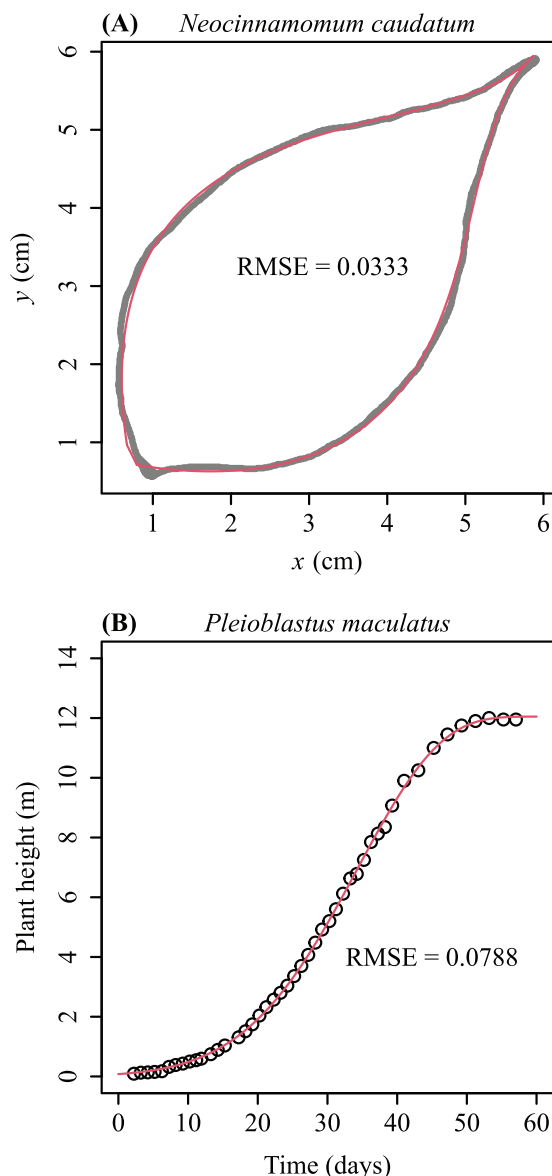


FIGURE 4 Observed (gray) and predicted (red) ovate leaf shape (A), and the observed and predicted height values of bamboo plants against time (B). The two curves both used the modified Lobry–Rosso–Flandrois equation (MLRFE; Equation 9). RMSE is the root-mean-square error between the observed and predicted y values. The estimated model parameters can be found in Appendix S1.

lot of attention.^{6,16,39,40} Although the 'GE' function can produce such a shape, it is necessary to provide a reference model for it. The 'NRGE' function⁶ is used to produce the y -coordinates when the x -coordinates are along the boundary of the upper part of an egg in the plane. The 'curveNRGE' function is used to simulate an integrated egg shape in the Euclidean coordinate system (i.e., a side view of the egg).

The data fitting module provides empirical datasets ('bambooleaves', 'eggs', 'ginkgoseed', 'kp', 'Neocinnamomum', 'starfish', 'whitespruce', 'shoots', and 'veins'), the data adjustment function ('adjdata'), functions for data-fitting ('fitGE', 'fitNRGE', and 'fitovate'), and a

function ('fracdim') for calculating the fractal dimension of leaf veins or leaf boundaries (Figure 2). With the exception of 'shoots' and 'veins', the datasets provide the boundary coordinate data of natural geometrical shapes: 'bambooleaves' consists of the boundary coordinate data of six leaves of the bamboo *Phyllostachys incarnata*; 'eggs' consists of the boundary coordinate data of nine species of avian eggs; 'ginkgoseed' consists of the boundary coordinate data of four side projections of the seeds of the gymnosperm *Ginkgo biloba* cv. Fozhi; 'kp' consists of the boundary coordinate data of four vertical projections of the fruits of the golden rain tree (*Koelreuteria paniculata*); 'Neocinnamomum' consists of the boundary coordinate data of the leaves of seven species of the genus *Neocinnamomum*; 'starfish' consists of the boundary coordinate data of eight individual sea stars from five different species; and 'whitespruce' consists of the planar coordinates of the tree rings of the evergreen conifer *Picea glauca*. In a similar manner, 'shoots' consists of the height growth data of four species of bamboo; and 'veins' provides a dataset of leaf vein coordinates of broad-leaved evergreen tree *Michelia compressa*. The data-fitting functions, including 'fitGE', 'fitNRGE', and 'fitovate', are used to fit the boundary coordinates of natural geometries using the Gielis or twin Gielis equations or their simplified versions, the egg-shape model, and the three ovate leaf-shape models. Nevertheless, before using these data-fitting functions, the boundary data should be adjusted by using the 'adjdata' function. This function changes an input dataset so that it contains points ordered counterclockwise along the shape boundary in the Euclidean coordinate system. For shapes having one apparent axis of symmetry, the 'bilat' function can be used to measure the extent of bilateral (a)symmetry based on the adjusted boundary data. In addition to the measure of bilateral (a)symmetry, the 'bilat' function can provide descriptive data for a polygon transcribed by a shape, including its area, length, width, perimeter, and so on. Height growth data (e.g., 'shoots') do not need to be adjusted by the 'bilat' function before fitting them using the 'fitsigmoid' function. In addition, the 'fracdim' function is used to calculate the fractal dimension of a complex leaf venation network (e.g., 'veins').

APPLICATIONS

'biogeom' was developed based on R (version 4.2.0),⁴¹ and requires installation of the R packages 'spatstat.geom' (version 2.4-0) and 'spatstat.data' (version 2.1-4).⁴²

```
library(spatstat.data)
library(spatstat.geom)
library(biogeom)
```

Below, we provide several examples of data simulation, data fitting, quantification of the extent of leaf bilateral (a)symmetry, and calculation of leaf vein fractal dimension. As emphasized before, the scope of the package application is not limited to leaves; the functions can simulate and fit other natural geometries.

1. An example of simulating two ovate leaf shapes based on two different rate equations.

```
P1 <- c(1, 1, pi/4, 2, 3, 10, 2)
RE1 <- curveovate(MLRFE, P = P1, x = seq(0,
  10, by = 0.1), fig.opt = TRUE)
RE2 <- curveovate(MbetaE, P = P1, x = seq(0,
  10, by = 0.1), fig.opt = TRUE)
dev.new()
par(mar = c(5,5,2,2))
plot(RE1$x, RE1$y, cex.lab = 1.5, cex.axis =
  1.5, xlab = expression(italic(x)), ylab =
  expression(italic(y)))
lines(RE2$x, RE2$y, col = 4)
```

2. An example of fitting the boundary data of a bamboo leaf.

```
data(bambooleaves)
uni.C <- sort(unique(bambooleaves$Code))
ind <- 1
Data <- bambooleaves[bambooleaves$Code ==
  uni.C[ind],]
x0 <- Data$x
y0 <- Data$y
Res1 <- adjdata(x0, y0, ub.np = 1000, times =
  1.2, len.pro = 1/20)
x1 <- Res1$x
y1 <- Res1$y
x0.ini <- min(x1)
y0.ini <- min(y1)
theta.ini <- pi/4
a.ini <- 8.4
n1.ini <- seq(0.01, 0.1, by = 0.01)
n2.ini <- 1
ini.val <- list(x0.ini, y0.ini, theta.ini,
  a.ini, n1.ini)
Res <- fitGE(GE, x = x1, y = y1, ini.val =
  ini.val, m = 1, simpver = 2, nval = 1,
  par.list = FALSE, stand.fig = TRUE, angle
  = NULL, fig.opt = TRUE, np = 2000, control =
  list(reltol = 1e-20, maxit = 20000))
RMSE0 <- sqrt(Res$RSS/Res$sample.size)
```

3. An example of fitting the height growth data of a species of bamboo.

```
data(shoots)
attach(shoots)
ind <- 3
x1 <- x[Code == ind]
y1 <- y[Code == ind]
```

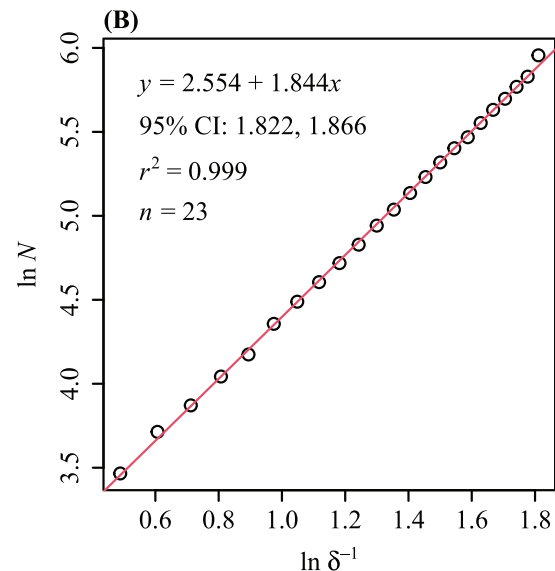
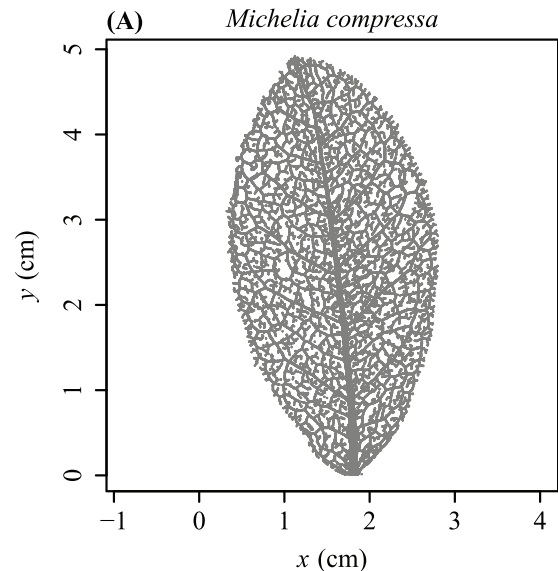


FIGURE 5 Leaf venation network (A) and the calculation of its fractal dimension (B) for the leaf veins of *Michelia compressa* (Magnoliaceae). In panel (B), the box-counting method was used to calculate the leaf-vein fractal dimension. Here, δ represents the side length of a square box; N is the number of square boxes that include at least one pixel of leaf veins; the slope is the estimated fractal dimension; 95% CI represents the 95% confidence interval of the slope; r^2 is the coefficient of determination; and n represents the sample size used for the linear regression.

```
delta0 <- c(0.5, 1, 2, 5, 10, 20)
ini.val <- list(600, 25, 0, 40, delta0)
resu2 <- fitsigmoid(MLRFE, x = x1, y = y1,
  ini.val = ini.val, simpver = NULL, fig.opt =
  TRUE, control = list(reltol = 1e-20, maxit
  = 20000))
```

4. An example of getting the measure of the extent of a leaf's bilateral (a)symmetry.

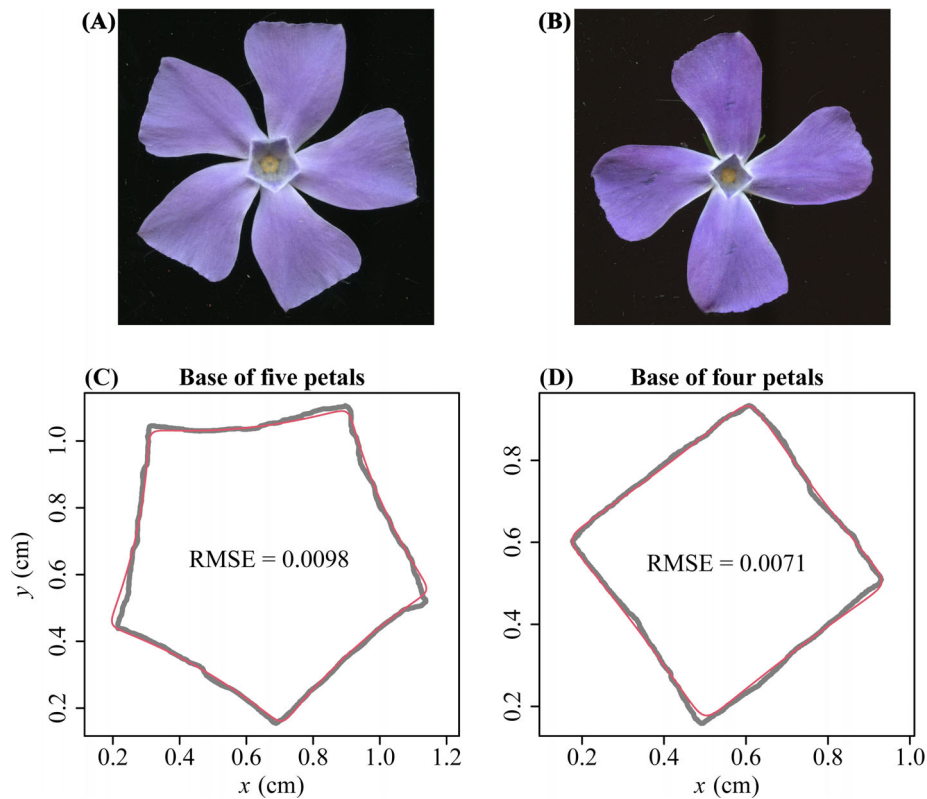


FIGURE 6 Examples of normal and mutant petal bases of *Vinca major* (A, B) and results of fitting their shapes using the Gielis equation (C, D). In panels C and D, the gray curves represent the observed boundaries of the petal bases; the red curves represent the boundaries predicted by the Gielis equation; RMSE is the root-mean-square error between the observed and predicted polar radii.

```
data(Neocinnamomum)
uni.C <- sort(unique(Neocinnamomum$Code))
ind <- 2
Data <- Neocinnamomum[Neocinnamomum$Code ==
  uni.C[ind],]
x0 <- Data$x
y0 <- Data$y
Res1 <- adjdata(x0, y0, ub.np = 2000, len.pro
  = 1/20)
x1 <- Res1$x
y1 <- Res1$y
Res2 <- bilat(x = x1, y = y1, time.interval
  = 0.00045, peri.np = 500, n.loop = 30,
  auto.search = TRUE, fd.opt = TRUE, frac.fig
  = TRUE)
Res2$AR
Res2$SI
Res2$scan.area
Res2$scan.length
Res2$scan.area
Res2$scan.perimeter
```

Note: before using the 'bilat' function, we used the 'adjdata' function to adjust the data points on the leaf boundary coordinates to strictly follow a counterclockwise order. The 'bilat' function can obtain the extent of leaf bilateral (a)symmetry (the area ratio of the left side to the

right side of a leaf lamina [AR] and the standardized index for bilateral (a)symmetry [SI]) and other leaf measurements, including the leaf area, length, width, and perimeter.

5. An example of calculating the fractal dimension of a leaf vein network.

```
data(veins)
par(mar = c(5,5,2,2))
plot(veins$x, veins$y, cex = 0.01, asp = 1,
  cex.lab = 1.5, cex.axis = 1.5, xlab
  = expression(italic(x)), ylab =
  expression(italic(y)))
fracdim(veins$x, veins$y)
```

Figure 3 shows additional fitted examples, and includes the boundary coordinate data of a bamboo leaf, an avian egg, a ginkgo seed, a vertical projection of the fruit of the golden rain tree (*K. paniculata*), tree rings (in cross-sections) of the stem of a white spruce (*P. glauca*), and a sea star. The gray curves represent the scanned (actual) boundaries, and the red curves represent the boundaries predicted by the Gielis equation. The root-mean-square errors (RMSEs, which equal the square-root of the quotient of the RSS and the sample size) between the observed and predicted polar radii are also listed in the panels of Figure 3. Figure 4 shows the application of the MLRFE in fitting an ovate leaf of a species of Lauraceae, and the height versus time



FIGURE 7 The aboveground architectural structure of *Myricaria prostrata* in Pulan County, Tibet, China. Dr. Yongping Yang took the photo on July 2, 2012.

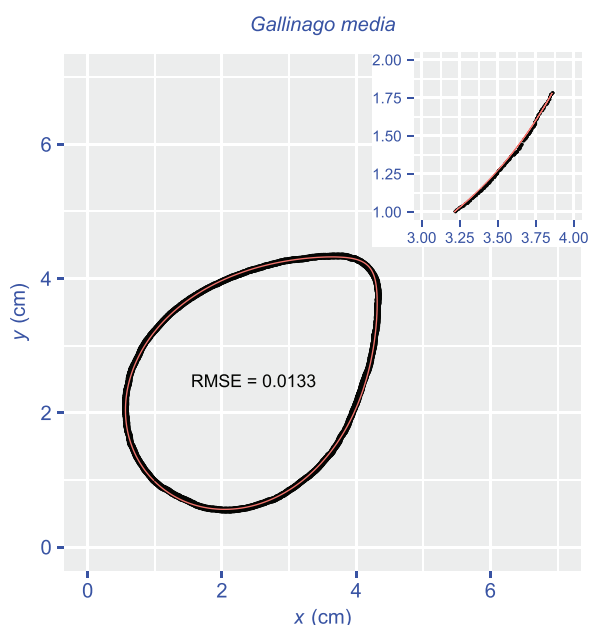


FIGURE 8 An example of using the package 'ggplot2' to redraw the results of fitting the boundary coordinate data of a bird egg. The observed (black) and predicted (red) geometries (boundary coordinates) predicted by the Gielis equation are shown here. The partial comparison between the observed and predicted geometries of the egg's boundary is shown in the top-right inset subplot.

data of a growing bamboo shoot. The RMSEs presented in this case represent the prediction errors between the observed and predicted y -values because the rate equation was not expressed in the polar coordinate system. Figure 5 shows the calculation of the fractal dimension of the leaf venation network of a species of Magnoliaceae based on the box-counting method.^{23,24} The estimated fractal dimension is the estimated slope of the straight line fitting the relationship between the number of square boxes that include one leaf-vein pixel and the side length of a box on a log-log scale (Figure 5B).

The parameters of data simulation and estimated parameters for Figures 1 and 3–5 can be found in Appendix S1.

POTENTIAL USES IN ECOLOGICAL AND EVOLUTIONARY RESEARCH

Prior studies have confirmed that many natural shapes, especially symmetrical shapes, can be expressed by the Gielis equation or its twin version,^{9–16} which shows the feasibility of using the Gielis equation in exploring ecological or evolutionary trajectories in the morphological variation among closely related species or taxa, different individuals of the same species (as in phenotyping), or changes attending ontogeny. A recent study showed that a simplified version of the Gielis equation can describe the shapes of real bird eggs well.¹⁶ The estimated value of the parameter controlling the symmetry of the curve is of potential value for quantifying the intraspecific and interspecific variations in avian egg profiles. However, doing so requires the use of a large sample for each species. (An exploration of this topic exceeds the scope of the present work.) In addition to the egg-shape example, the Gielis equation can depict the intraspecific variation in leaf and flower geometries. We provide an example of the variation in the shape of the petal base of *Vinca major*. The normal petal bases of this plant are pentagonal (Figure 6A). However, quadrangular petal bases also occasionally occur rarely (Figure 6B) (but less than 2% according to a survey of 380 flowers of *V. major*, unpublished data). Nevertheless, the two types of petal bases both follow a simplified version of the Gielis equation (Figure 6C,D) that is also valid in describing bird egg shapes.¹⁶ The 'fracdim' function can be used to calculate the fractal dimension of leaf veins, but it can also be used to examine the planar projections of objects with branched structures, for example, arteries, veins, and trees. Figure 7 shows a plant species growing in the alpine regions of Tibet that has evolved a low spreading growth form, possibly as an adaptation to very strong winds and low temperatures. The aboveground architecture (in top view) of the plant is comparable with

those of closely related species with upright architectures (inside view) using their fractal dimensions calculated by the 'fracdim' function of 'biogeom'.

In addition to describing the egg shapes of birds and reptiles, 'biogeom' can also describe the eggs of some insects, which can have large deformations from axial symmetry. Therefore, efforts to increase the deformation versatility of the Gielis equation are worth of attention. From our recent study on simulating and fitting the cross-sections of a species of square bamboo (*Chimonobambusa utilis*),⁴³ it appears to be at least technically feasible to combine a deformation function with one or more equation parameter(s).

Finally, the 'biogeom' package uses the base graphics system (from the 'graphics' package) to produce graphical outputs that are traditional S-like graphics.⁴⁴ In comparison with the currently popular R package 'ggplot2', which uses a perfect set of the grammar of graphics and statistical models to conveniently visualize data and embodies a deep philosophy of visualization,⁴⁵ the base graphics system has its drawbacks since it lacks these features. However, to reduce the dependencies of 'biogeom' on other R packages and the frequency of software updates needed, we suggest using the base graphics system instead of the 'ggplot2' grammar of the graphics system in our package. Fortunately, the graphical output is only a visualization of the values returned by the functions in 'biogeom', and the user can easily use these values and the functions in 'ggplot2' to redraw the graphics. In Figure 8, we illustrate an example of using the package 'ggplot2' to redraw the results of fitting the boundary coordinate data of a bird egg.

CONCLUSIONS

The R package 'biogeom' provides several novel ready-to-use functions to simulate and fit natural geometries. It can be used to quantify the morphological characteristics of eggs, fruits, seeds, and other organic objects that are apparently or approximately symmetrical and is a promising tool for exploring intra- and interspecific variation in shape and obtaining relevant quantitative trait information for evolutionary and ecological studies. In particular, the evolutionary trajectories of different organisms can be reflected by their morphological (dis)similarity. The influence of external factors (e.g., temperature, light intensity, and humidity) can also lead to a variation in biological geometries due to the morphological differences among organs and tissues. To a certain degree, the R package 'biogeom' can analyze the similarity and variation among organismal shapes using a universal equation. In addition, 'biogeom' provides candidate sigmoid growth equations, and a function to produce an index of the extent of leaf bilateral (a)symmetry. Basic botanical information on leaf area, length, width, and perimeter can also be provided by this package, which means that it is suitable for calculating leaf-shape indices, such as the leaf roundness index, leaf dissection index, and leaf ellipticalness index.¹⁴

ACKNOWLEDGMENTS

We are deeply thankful to Dr. Douglas Braaten for pre-evaluating this work and encouraging us to submit it to this journal, and we thank the reviewer for providing us with positive and constructive comments. We

also thank Cang Hui, Lei Chen, Meiling Fan, Jianguo Huang, Yabing Jiao, Yang Li, Yirong Li, Mengdi Liu, Yu Song, Rolf Turner, Victoria Wimmer, Yongping Yang, and Kexin Yu for their valuable help during the development of this package. J.S. was supported by the German Research Foundation (DFG; Grant No. SCHR1672/1-1); H.R. was supported by the National Key Research and Development Program of China (Grant No. 2021YFD02200403).

Open access funding enabled and organized by Projekt DEAL.

AUTHOR CONTRIBUTIONS

All authors contributed integrally to the construction of all the points of the manuscript, giving their consent for its publication; P.S., J.G., K.J.N., J.S., and Ü.N. designed this work; P.S., J.G., and D.A.R. provided methods; P.S., B.K.Q., and J.S. wrote the initial draft; and J.G., K.J.N., D.A.R., H.R., L.W., B.K.Q., and Ü.N., commented on and revised the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

Documentation, source code, and the datasets used herein are freely available on CRAN (<https://cran.wustl.edu/web/packages/biogeom/index.html>) and GitHub (<https://github.com/cran/biogeom>). The R package biogeom can be installed in R version 4.2.0 or higher with the following command: `install.packages("biogeom")`.

PEER REVIEW

The peer review history for this article is available at: <https://publons.com/publon/10.1111/nyas.14862>.

ORCID

Peijian Shi  <https://orcid.org/0000-0003-4696-0130>

Johan Gielis  <https://orcid.org/0000-0002-4536-3839>

Brady K. Quinn  <https://orcid.org/0000-0001-5781-7087>

Karl J. Niklas  <https://orcid.org/0000-0003-3491-1286>

David A. Ratkowsky  <https://orcid.org/0000-0002-7514-3570>

Julian Schrader  <https://orcid.org/0000-0002-8392-211X>

Ülo Niinemets  <https://orcid.org/0000-0002-3078-2192>

REFERENCES

1. Thompson, D. W. (1917). *On growth and form*. London: Cambridge University Press.
2. Palmer, A. R., & Strobeck, C. (1986). Fluctuating asymmetry: Measurement, analysis, patterns. *Annual Review of Ecology, Evolution, and Systematics*, 17, 391–421.
3. Gielis, J. (2003). A general geometric transformation that unifies a wide range of natural and abstract shapes. *American Journal of Botany*, 90, 333–338.
4. Dornbusch, T., Watt, J., Baccar, R., Fournier, C., & Andrieu, B. (2011). A comparative analysis of leaf shape of wheat, barley and maize using an empirical shape model. *Annals of Botany*, 107, 865–873.
5. Buschmann, H., & Borchers, A. (2020). Handedness in plant cell expansion: A mutant perspective on helical growth. *New Phytologist*, 225, 53–69.
6. Narushin, V. G., Romanov, M. N., & Griffin, D. K. (2021). Egg and math: Introducing a universal formula for egg shape. *Annals of the New York Academy of Sciences*, 1505, 169–177.

7. Matsuura, M. (2015). Gielis' superformula and regular polygons. *Journal of Geometry*, 106, 383–403.
8. Shi, P., Xu, Q., Sandhu, H. S., Gielis, J., Ding, Y., Li, H., & Dong, X. (2015). Comparison of dwarf bamboos (*Indocalamus* sp.) leaf parameters to determine relationship between spatial density of plants and total leaf area per plant. *Ecology and Evolution*, 5, 4578–4589.
9. Lin, S., Zhang, L., Reddy, G. V. P., Hui, C., Gielis, J., Ding, Y., & Shi, P. (2016). A geometrical model for testing bilateral symmetry of bamboo leaf with a simplified Gielis equation. *Ecology and Evolution*, 6, 6798–6806.
10. Shi, P., Ratkowsky, D. A., Li, Y., Zhang, L., Lin, S., & Gielis, J. (2018). A general leaf area geometric formula exists for plants—Evidence from the simplified Gielis equation. *Forests*, 9, 714.
11. Su, J., Niklas, K. J., Huang, W., Yu, X., Yang, Y., & Shi, P. (2019). Lamina shape does not correlate with lamina surface area: An analysis based on the simplified Gielis equation. *Global Ecology and Conservation*, 19, e00666.
12. Shi, P., Ratkowsky, D. A., & Gielis, J. (2020). The generalized Gielis geometric equation and its application. *Symmetry*, 12, 645.
13. Tian, F., Wang, Y., Sandhu, H. S., Gielis, J., & Shi, P. (2020). Comparison of seed morphology of two ginkgo cultivars. *Journal of Forestry Research*, 31, 751–758.
14. Li, Y., Niklas, K. J., Gielis, J., Niinemets, Ü., Schrader, J., Wang, R., & Shi, P. (in press). An elliptical blade is not a true ellipse, but a superellipse – Evidence from two *Michelia* species. *Journal of Forestry Research*, <https://doi.org/10.1007/s11676-021-01385-x>
15. Li, Y., Quinn, B. K., Gielis, J., Li, Y., & Shi, P. (2022). Evidence that supertriangles exist in nature from the vertical projections of *Koeleruteria paniculata* fruit. *Symmetry*, 14, 23.
16. Shi, P., Gielis, J., & Niklas, K. J. (in press). Comparison of a universal (but complex) model for avian egg shape with a simpler model. *Annals of the New York Academy of Sciences*, <https://doi.org/10.1111/nyas.14799>
17. Gielis, J. (2017). *The geometrical beauty of plants*. Paris: Atlantis Press.
18. Shi, P., Zheng, X., Ratkowsky, D. A., Li, Y., Wang, P., & Cheng, L. (2018). A simple method for measuring the bilateral symmetry of leaves. *Symmetry*, 10, 118.
19. Shi, P., Yu, K., Niklas, K. J., Schrader, J., Song, Y., Zhu, R., Li, Y., Wei, H., & Ratkowsky, D. A. (2021). A general model for describing the ovate leaf shape. *Symmetry*, 13, 1524.
20. Yin, X., Goudriaan, J., Lantinga, E. A., Vos, J., & Spiertz, H. J. (2003). A flexible sigmoid function of determinate growth. *Annals of Botany*, 91, 361–371.
21. Shi, P., Fan, M., Ratkowsky, D. A., Huang, J., Wu, H., Chen, L., Fang, S., & Zhang, C. (2017). Comparison of two ontogenetic growth equations for animals and plants. *Ecological Modelling*, 349, 1–10.
22. Cao, L., Shi, P., Li, L., & Chen, G. (2019). A new flexible sigmoidal growth model. *Symmetry*, 11, 204.
23. Vico, P. G., Kyriacos, S., Heymans, O., Louryan, S., & Cartilier, L. (1998). Dynamic study of the extraembryonic vascular network of the chick embryo by fractal analysis. *Journal of Theoretical Biology*, 195, 525–532.
24. Shi, P., Yu, K., Niinemets, Ü., & Gielis, J. (2021). Can leaf shape be represented by the ratio of leaf width to length? Evidence from nine species of *Magnolia* and *Michelia* (Magnoliaceae). *Forests*, 12, 41.
25. Ratkowsky, D. A. (1983). *Nonlinear regression modeling: A unified practical approach*. New York: Marcel Dekker.
26. Ratkowsky, D. A. (1990). *Handbook of nonlinear regression models*. New York: Marcel Dekker.
27. Uvarov, B. P. (1931). Insects and climate. *Transactions of the Entomological Society of London*, 79, 1–232.
28. Campbell, A., Frazer, B. D., Gilbert, N., Gutierrez, A. P., & Mackauer, M. (1974). Temperature requirements of some aphids and their parasites. *Journal of Applied Ecology*, 11, 431–438.
29. Ratkowsky, D. A., Lowry, R. K., McMeekin, T. A., Stokes, A. N., & Chandler, R. E. (1983). Model for bacterial culture growth rate throughout the entire biokinetic temperature range. *Journal of Bacteriology*, 154, 1222–1226.
30. Régnière, J., Powell, J., Bentz, B., & Nealis, V. (2012). Effects of temperature on development, survival and reproduction of insects: Experimental design, data analysis and modeling. *Journal of Insect Physiology*, 58, 634–647.
31. Shi, P., Reddy, G. V. P., Chen, L., & Ge, F. (2016). Comparison of thermal performance equations in describing temperature-dependent developmental rates of insects: (I) empirical models. *Annals of the Entomological Society of America*, 109, 211–215.
32. Ratkowsky, D. A., & Reddy, G. V. P. (2017). Empirical model with excellent statistical properties for describing temperature-dependent developmental rates of insects and mites. *Annals of the Entomological Society of America*, 110, 302–309.
33. Rebaudo, F., Struelens, Q., & Dangles, O. (2018). Modeling temperature-dependent development rate and phenology in arthropods: The devRate package for R. *Methods in Ecology and Evolution*, 9, 1144–1150.
34. Yin, X., Kropff, M. J., McLaren, G., & Visperas, R. M. (1995). A nonlinear model for crop development as a function of temperature. *Agricultural and Forest Meteorology*, 77, 1–16.
35. Brière, J.-F., Pracros, P., Le Roux, A.-Y., & Pierre, J.-S. (1999). A novel rate model of temperature-dependent development for arthropods. *Environmental Entomology*, 28, 22–29.
36. Lobry, J. R., Rosso, L., & Flandrois, J. P. (1991). A Fortran subroutine for the determination of parameter confidence limits in non-linear models. *Binary*, 3, 86–93.
37. Rosso, L., Lobry, J. R., & Flandrois, J. P. (1993). An unexpected correlation between cardinal temperatures of microbial growth highlighted by a new model. *Journal of Theoretical Biology*, 162, 447–463.
38. Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7, 308–313.
39. Stoddard, M. C., Yong, E. H., Akkaynak, D., Sheard, C., Tobias, J. A., & Mahadevan, L. (2017). Avian egg shape: Form, function, and evolution. *Science*, 356, 1249–1254.
40. Biggins, J. D., Thompson, J. E., & Birkhead, T. R. (2018). Accurately quantifying the shape of birds' eggs. *Ecology and Evolution*, 8, 9728–9738.
41. R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
42. Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. Chapman and Hall/CRC Press.
43. Huang, W., Li, Y., Niklas, K. J., Gielis, J., Ding, Y., Cao, L., & Shi, P. (2020). A superellipse with deformation and its application in describing the cross-sectional shapes of a square bamboo. *Symmetry*, 12, 2073.
44. Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis*. Belmont, CA: Wadsworth & Brooks/Cole.
45. Wickham, H. (2009). *ggplot2. Elegant graphics for data analysis*. New York: Springer.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Shi, P., Gielis, J., Quinn, B. K., Niklas, K. J., Ratkowsky, D. A., Schrader, J., Ruan, H., Wang, L., & Niinemets, Ü. (2022). 'biogeom': An R package for simulating and fitting natural shapes. *Ann NY Acad Sci.*, 1–12. <https://doi.org/10.1111/nyas.14862>