Non-linear mixed model implementation in InfoStat and interface to nlme library

Julio A. Di Rienzo
Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina –
dirienzo.julio@gmail.com

Raúl E. Macchiavelli
College of Agricultural Sciences, University of Puerto Rico, Mayagüez, Puerto Rico –
raul.macchiavelli@upr.edu

Fernando Casanoves*
Centro Agronómico Tropical de Investigación y Enseñanza, CATIE, Turrialba, Costa Rica –
casanoves@catie.ac.cr

Abstract

The nlme library of R implements linear and non-linear mixed models through the lme and nlme functions. In this paper we present an implementation of an interface of the nlme function in the framework of InfoStat Statistical Software. With this easy-to-use interface, InfoStat users can add non-linear mixed effects models to the statistics tools available in this free software.
The implementation includes several data handling options to obtain adjusted and predicted values as well as graphical tools for diagnostic purposes. When no random effects are specified, the interface utilizes the nls function in a way that is transparent for the end user. A tutorial containing several worked examples is included.

Keywords: non-linear regression; R software, mixed models; InfoStat software.

1. Introduction

Most statisticians have the expertise to handle complex models like nonlinear mixed effects models correctly using several software tools (mainly R, but also SAS, GenStat, etc.), but there is a large community of potential users that continue to use old approaches to fit nonlinear models in non-standard situations (for example, transforming the data, fitting individual curves and averaging them, ignoring possible correlations among data points, etc.). This may be due to the difficulty of understanding the theoretical background of these models and the complex and non-intuitive rules to fit them using statistical software.
The nlme library of R implements linear and non-linear mixed models through the lme and nlme functions. We developed an easy-to-use interface, in the framework of InfoStat, of the nlme function. This adds a non-linear mixed effects modelling tool to InfoStat. Previous work done with user-friendly R interfaces in InfoStat facilitates the use of linear mixed models (Di Rienzo et al. 2011), generalized linear mixed models (Di Rienzo et al. 2012), and DNA-Microarray analysis (Di Rienzo et al. 2016). All these include tutorials with well-developed examples to facilitate both the implementation and the correct interpretation of these methodologies.
The implementation of the nonlinear mixed model module is accompanied by a tutorial, which includes several worked examples. It also includes several data handling tools to obtain fitted and predicted values as well as graphical tools for diagnostic purposes. When no random effects are specified, the interface utilizes the nls function in a way that is transparent for the end user.

2. The interface

We developed a Windows implementation of a friendly interface to nlme functions within the framework of the InfoStat statistical software (www.infostat.com.ar). The application allows the user to declare random effects on fixed parameters model, obtain diagnostic graphics (Figure 1), and model heterogeneous variances (Figure 2) and spatially or temporally correlated data (Figure 3) using a friendly interface.

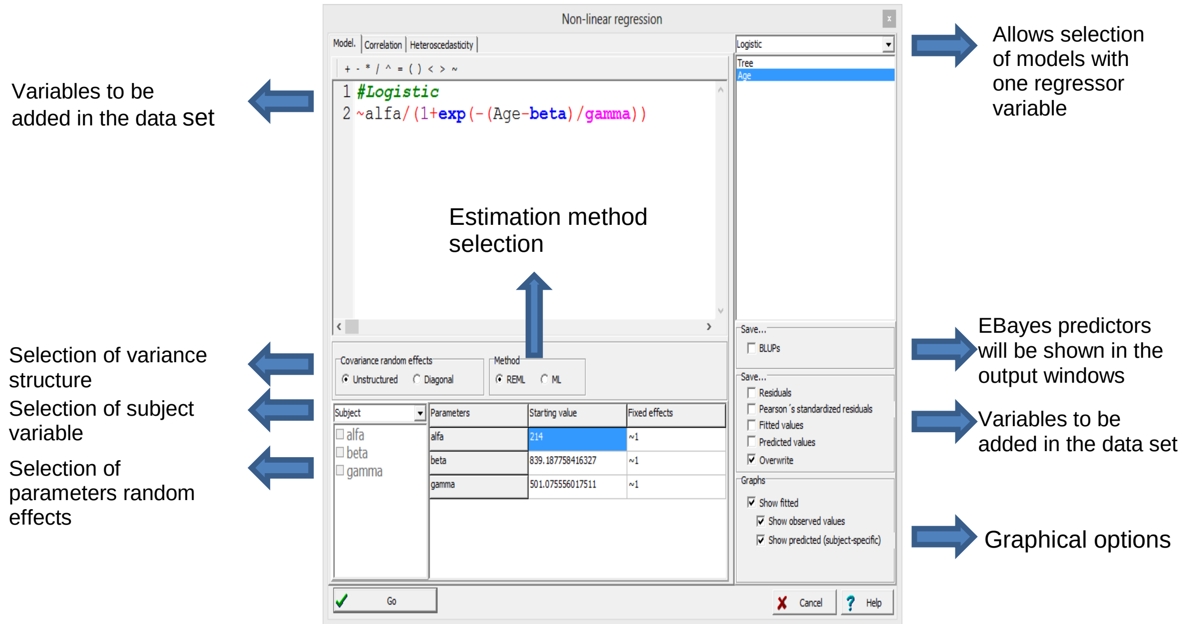

Figure 1. Mixed nonlinear models interface in InfoStat: Model definition.

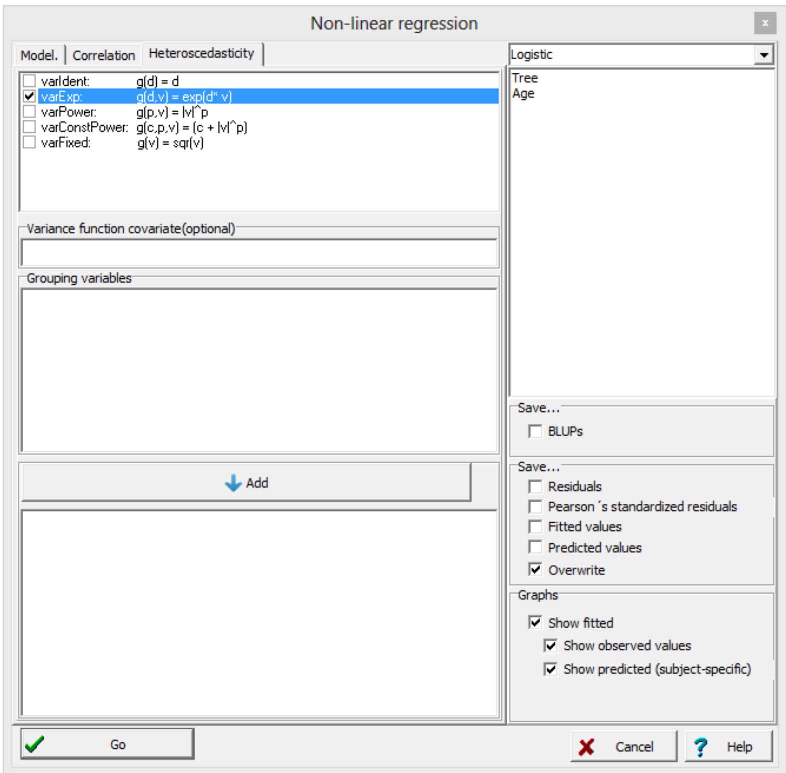

Figure 2. Mixed nonlinear models interface in InfoStat: Heteroscedasticity options.

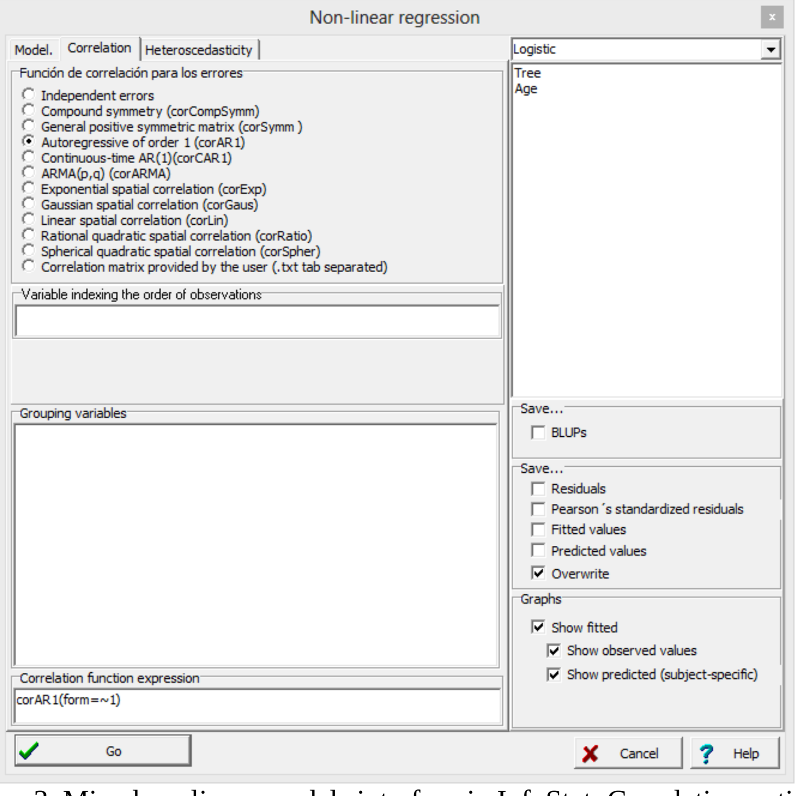

Figure 3. Mixed nonlinear models interface in InfoStat: Correlation options.

3. Example

Data set description

The data describe the growth of orange trees (Figure 4). The trunk circumference of 5 trees is measured at 7 different ages, giving a total of 35 datapoints (Draper and Smith, 1998; Pinheiro and Bates, 2004). From Figure 5 it is clear that a three-parameter logistic regression model fits the individual tree data, and also could fit the average growth curve:

$$\frac{\alpha}{1+\exp\left(-(\mathrm{Age}-\beta)/\gamma\right)}$$

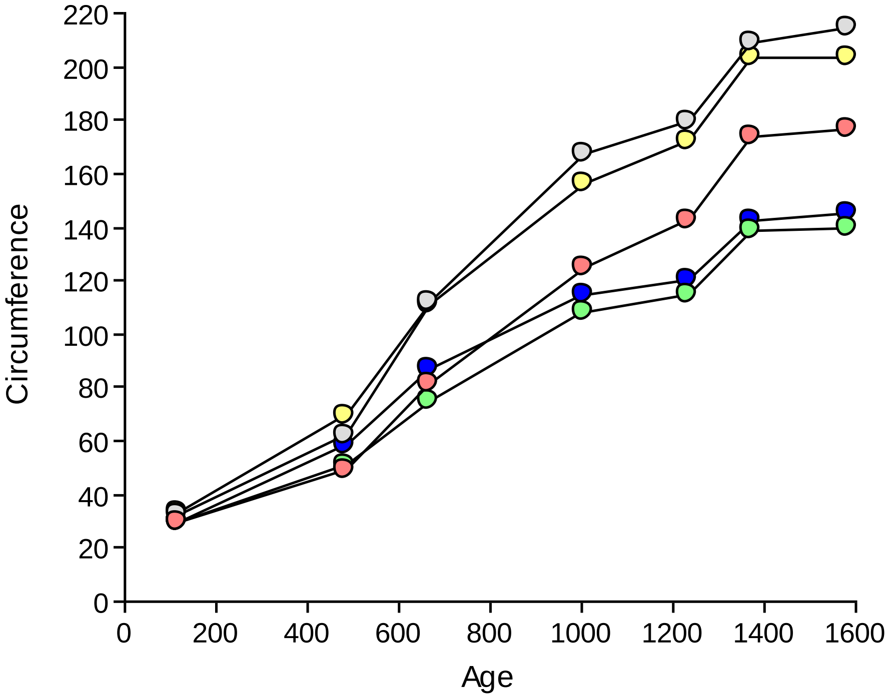

Figure 4. Scatter plot of Circumference versus Age for each tree.

Results

Different models were compared, no random tree effect (AIC=301.59, Figure 3 right), random tree effect on alpha (AIC=251.52, Figure 3 left) and random tree effect on alpha and beta (AIC=253.52). The results for the best model are presented. All fitted models considered first order autoregressive correlation for the error term.

InfoStat output

Fit measurements

N	AIC	BIC	logLik	Sigma
35	251.52	260.31	-119.76	8.22

Smaller AIC and BIC is better

Fixed effects coefficients

	Value	Std.Error	DF	t-value	p-value
alfa	190.99	16.16	28.00	11.82	2.1E-12
beta	722.21	35.38	28.00	20.41	0.00
gamma	344.04	27.31	28.00	12.60	0.00

Correlation between fixed-effects estimates

	alfa	beta	gamma
alfa	1.00	0.38	0.36
beta	0.38	1.00	0.75
gamma	0.36	0.75	1.00

Covariance matrix for random effects

	alfa
alfa	16.01

Random effects matrix (BLUPs)

	alfa
1	-29.39
2	31.53
3	-36.97
4	39.99
5	-5.16

Correlation structure

Correlation model: AR(1)
Formula: ~ 1 | Tree
Model parameters

Parameter	Estim.
Phi	0.01078

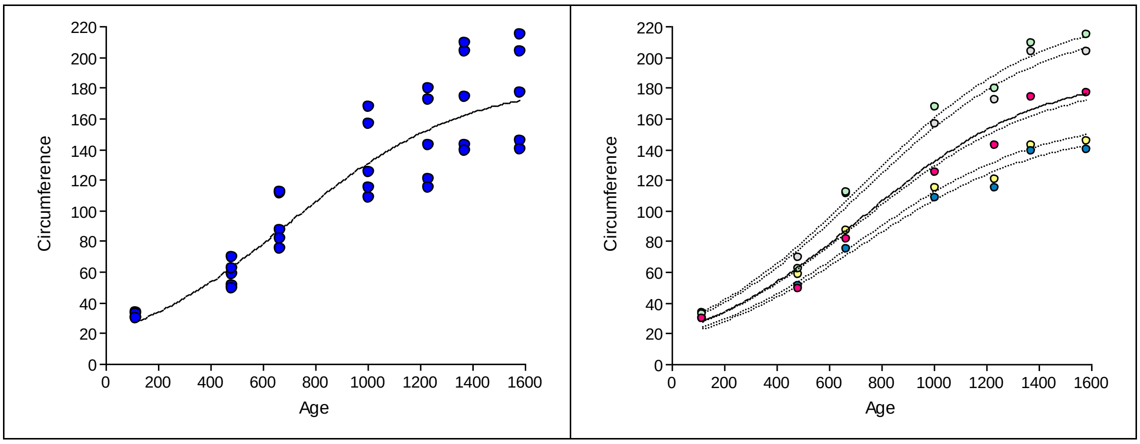

Figure 5. A three-parameter logistic regression model average fit (left) and the individual tree data fit (right)

5. Conclusions

This interface offers a user-friendly tool to fit nonlinear mixed models by maximum likelihood methods, making use of very reliable R functions implemented in the nlme library.
Complex model structures can be easily accommodated both in the fixed and the random parts of the model.
The implementation allows the fitting of nonlinear models including variance and correlation functions, to accommodate heteroscedasticity and lack of independence in the error term, commonly found in growth curves fitting.

References

Di Rienzo J.A., Casanoves F., Balzarini M.G., Gonzalez L., Tablada M., Robledo C.W. InfoStat versión 2016. InfoStat Group, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina. URL http://www.infostat.com.ar

Di Rienzo J.A., Macchiavelli, R.E., Casanoves F. 2011. Linear Mixed Models: Applications in InfoStat. InfoStat Group, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina. 244 p.

Di Rienzo J.A., Macchiavelli, R.E., Casanoves F. 2017. Generalized Linear Mixed Models: Applications in InfoStat. InfoStat Group, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina. 107 p.

Draper N.R. and Smith H. 1998. Applied Regression Analysis, 3rd ed. Wiley, New York.

Pinheiro J.C. and Bates D.M. 2004. Mixed-Effects Models in S and S-PLUS. Springer, New York.