Grofit: Fitting biological growth curves

Matthias Kahm[1*], Guido Hasenbrink[2], Hella Lichtenberg-Fraté[2], Jost Ludwig[2], Maik Kschischo[1]
[1]University of Applied Sciences Koblenz, RheinAhrCampus Remagen
[2]University of Bonn, Institute for Cellular and Molecular Botany

Introduction:

Growth experiments are routinely used to analyze basic properties of a given organism or cellular model. Similarly, within the TRANSLUCENT project the characterization of cellular growth and particularly under conditions of high extra cellular potassium or sodium is an important task. Any growth analysis should ideally reveal a relationship between the concentration of a compound/substrate and its effect on a particular growth parameter. In view of the quite labor intensive analysis of hundreds of growth curves - sometimes not revealing ideal relationships - the Grofit-software was developed to support biologists and such approaches. Within the software package, specifically tailored regression and bootstrapping techniques are utilized to statistically estimate the effect of different growth conditions. Grofit was implemented in R, an open source statistical software environment.
A complete description of the Grofit software and information about the implemented methods is given in a manuscript submitted to the Journal of Statistical Software. The software can also be obtained from our homepage:
http://www.rheinahrcampus.de/Research-Group-of-Maik-Kschisc.2452.0.html

AG Molekulare Bioenergetik

Characteristic growth parameters:

To investigate the specific effect of a given experimental set up or condition, characteristic parameters (λ= lag phase, μ = slope and A = maximum growth) of the growth curves are determined (Figure 1). Such parameter values are used to summarize a growth curve for a given condition.

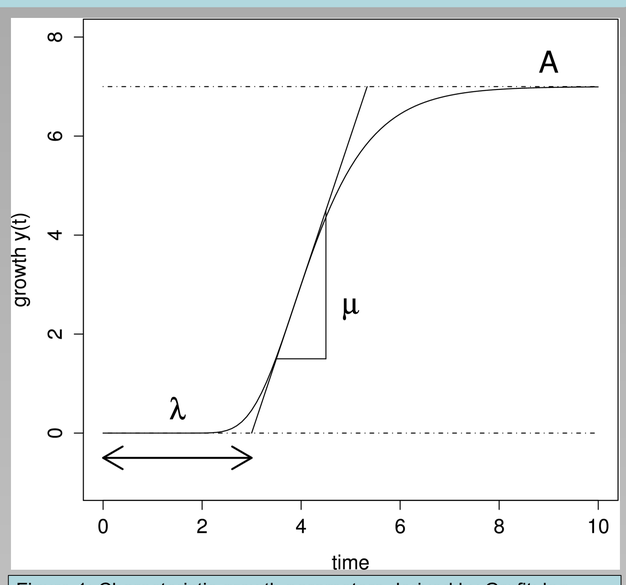

Figure 1: Characteristic growth parameters derived by Grofit: lag phase λ, maximal growth rate μ (maximum slope), maximal growth A and the area under the growth curve.

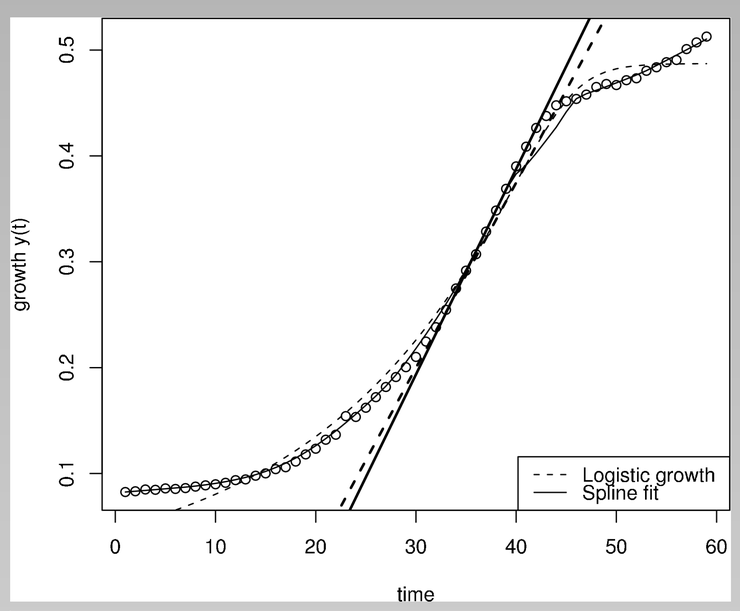

Figure 2: Comparison of parametric and model free spline fits. The growth data (circles) were fitted by a spline fit (solid line). The maximum slope of the spline fit was used as an estimate for the growth rate μ. This estimate is more accurate than the best fitting parametric model (logistic equation, dashed lines), as can be seen from the difference in the slopes of the tangents (straight lines).

Fitting strategies:

Grofit applies two different strategies for fitting any given growth curve: Model-based fits and model-free spline fits. The former requires a mathematical model for the description of cellular growth. To this end, four different known models were implemented in Grofit: 1. Logistic growth, 2. Gompertz growth, 3. modified Gompertz growth and 4. Richards growth. Such parametric growth curves are indeed useful and straight forward to interpret when accurately fitting the data. However, quite often the real data cannot sufficiently be described by a parametric model. As an alternative we implemented a model-free method that applies a smoothed cubic spline that does not assume a functional relationship between time and growth data.
Figure 2 shows the main difference between the two approaches. According to the Akaike criterion, the best fitting parametric model was the logistic equation. In this example, the maximum slope μ (see Figure 1) was regarded as the characteristic growth parameter. It it obvious that the smoothed spline provides a more accurate estimate of μ. We therefore conclude that the derivation of descriptive characteristics from parametric fits may potentially lead to unreliable predictions. A spline fit offers more accurate estimates of the characteristic growth parameters.

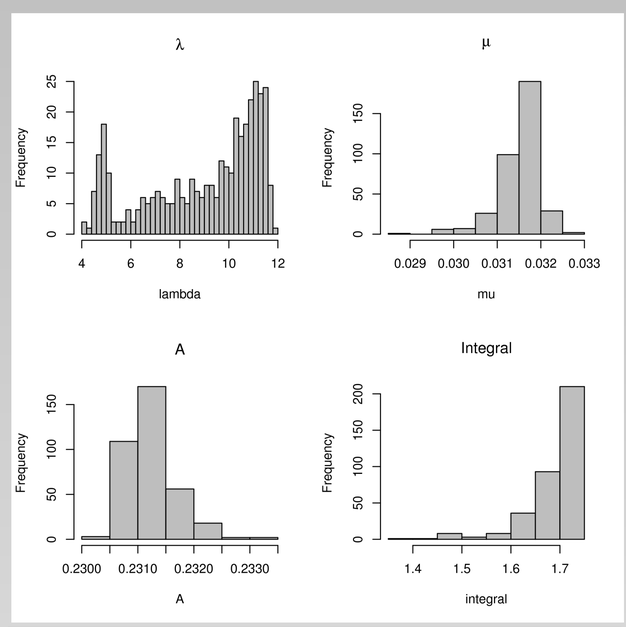

Figure 3: Bootstrap and cross-validation techniques are used for estimating confidence intervals of all derived parameters.

Dose response curves:

Upon availability of a statistical relevant number of growth curves corresponding dose response plots can be computed that enable the determination of characteristic descriptive values such as EC50.
Figure 4 demonstrates how a dose response curve is derived from growth experiments: Yeast cells were treated with different concentrations of a compound (here Hygromycin B). Growth was measured as optical density (Hasenbrink et al., 2006) at different time points. From the fitted growth curves the maximum growth rate μ (see Figure 1) was derived. This response μ was then plotted *versus* the dose in Figure 4(b). In a subsequent step a dose response curve was fitted and the EC50 value determined.

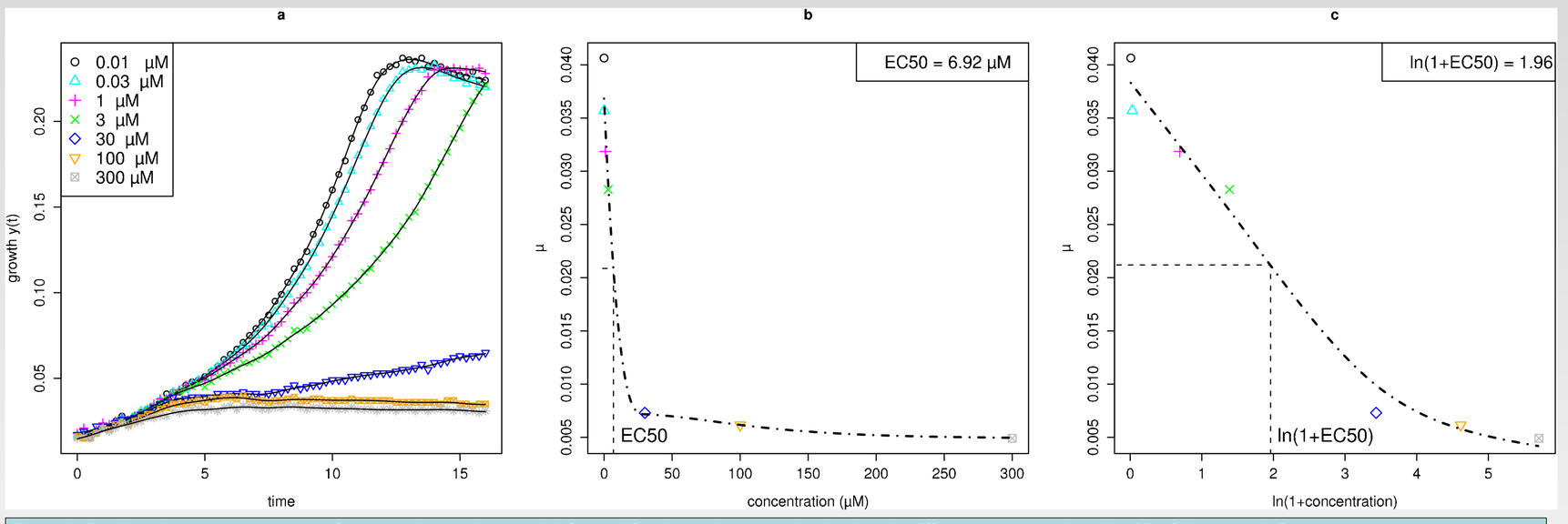

Figure 3: Deriving dose response curves from grwoth experiments: (a) Several fitted growth curves obtained under different concentrations (in μM) of Hygromycin B.
(b) The maximum slope corresponding to the growth rate μ of each curve in (a) is calculated and plotted *vs.* the corresponding concentration. From these data points a dose response curve is estimated by fitting a smoothed spline. Consequently, the EC50 value 6.92 μM can be estimated. (c) In order to obtain a more uniform distribution of the data points a logarithmic transformation to the concentration axis can be applied.

Selected References:
- Kahm M. *et al.*, submitted to the Journal of Statistical Software
- Birch, C., (1999) Annals of Botany, 83, 713-723
- Hasenbrink, G. *et al.,* (2006) Applied and Microbiology Biotechnology, 73, 1212-1221
- Zwietering, M. H., (1990) Applied and Environmental Microbiology, 56, 1875-1881
- Cleveland, W.S., (1979) Journal of American Statistical Association, 74, 829-836
- Efron, B. and Tibshirani, R. J., (1993) Nonlinear Regression Analysis and its Application, Chapman & Hall

Funding:
SysMo Initiative
BMBF grant 031 3982 C

*Matthias Kahm,
RheinAhrCampus, Department Mathematics and Technology,
Südallee 2, 53424 Remagen/ Germany,
kahm@rheinahrcampus.de