

On Parameter Estimation by Nonlinear Least Squares in Some Special Two-Parameter Exponential Type Models

Darija Marković* and Luka Borozan

Department of Mathematics, J.J. Strossmayer University of Osijek, Trg Ljudevita Gaja 6, HR-31 000 Osijek, Croatia

Received: 20 Feb. 2015, Revised: 20 Apr. 2015, Accepted: 21 Apr. 2015

Published online: 1 Nov. 2015

Abstract: Two-parameter growth models of exponential type $f(t; a, b) = g(t)\exp(a + bh(t))$, where a and b are unknown parameters and g and h are some known functions, are frequently employed in many different areas such as biology, finance, statistic, medicine, ect. The unknown parameters must be estimated from the data (w_i, t_i, y_i) , $i = 1, \dots, n$, where t_i denote the values of the independent variable, y_i are respective estimates of regression function f and $w_i > 0$ are some data weights. A very popular and widely used method for parameter estimation is the method of least squares. In practice, to avoid using nonlinear regression, this kind of problems are commonly transformed to linear, which is not statistically justified. In this paper we show that for strictly positive g and strictly monotone h original nonlinear problem has a solution. Generalization in the l_p norm ($1 \leq p < \infty$) and some illustrative examples are also given.

Keywords: two-parameter models, least squares, parameter estimation, existence problem, data fitting

1 Introduction

In this paper we will investigate parameter estimation problem for the models of the type:

$$f(t; a, b) = g(t)e^{a+bh(t)}, \quad (1)$$

where a and b are unknown parameters and g and h are some known functions. This type of models are often used in applied research, such as biology, ecology, political science, psychology, economics and finance (see e.g. [10, 16, 23, 5, 25]).

The structure of the paper is as follows. In Section 2 we briefly describe few models of type (1). In Section 3 we formulate ordinary least squares (OLS) fitting problem for this type of models. In Section 4 we present our main result (Theorem 1) which guarantees the existence of the least squares estimate (LSE), provided the data satisfy natural conditions. Illustrative numerical examples are given in Section 5.

2 Some useful models of type (1)

Now we will give brief descriptions of some models of type (1), which are commonly used in applied research.

Given models range from generally applied to very specifically used.

2.1 Exponential regression

If we assume that the average rate of change of the population (P) over an interval of time is proportional to the size of the population (see [9, 18]), we have the following differential equation model:

$$\frac{dP}{dt} = kP, \quad (2)$$

where (for growth) k is a positive constant.

To solve it, we can rewrite equation (2) and get a following equation

$$\frac{dP}{P} = kdt.$$

Integration of both sides of the last equation yields

$$\ln P = kt + C,$$

or

$$P(t) = e^{kt+C}.$$

* Corresponding author e-mail: darija@mathos.hr

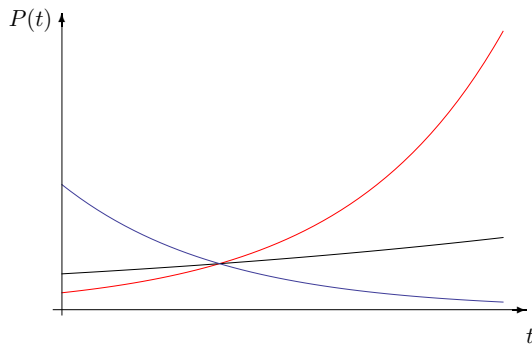


Fig. 1: Plots of the exponential model for some values of k and C

This model is sometimes also called log-level or log-linear model (or regression), since when dependent variable is log transformed relation becomes linear. Depending on the sign of parameter k , it can be increasing or decreasing. Figure 1 gives some examples of the exponential model.

2.2 Power regression

Assumption that percentage increases in independent variable t leads to constant percentage changes in dependant variable y (see e.g. [19, 16]) implies:

$$\frac{dy}{y} = k \frac{dt}{t},$$

and a simple calculation gives

$$y(t) = e^{k \ln t + C}.$$

This model has a property of scale invariance (see e.g. [5, 10, 25]). Model is also known as power law, log-log regression and allometric equation.

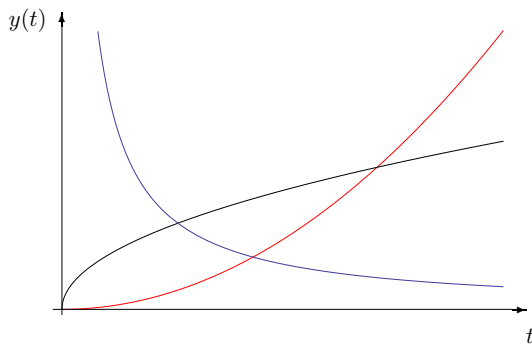


Fig. 2: Plots of the power model for some values of k and C

2.3 Fox surplus-yield model

This model is mainly used in fishery sciences for estimating the maximum sustainable yield. It was proposed by Fox in 1970. (see [8]). The equilibrium harvest or sustainable yield occurs when a fish stock's harvest rate H equals its growth rate G (see [6]), and that is when

$$\frac{dN}{dt} = G(N) - H(E, N) = 0. \tag{3}$$

The Fox model assumes a Gompertz growth of the underlying stock size in the absence of harvest

$$G(N) = rN \ln\left(\frac{K}{N}\right),$$

where r is the intrinsic growth rate and K is the carrying capacity of the environment. Another assumption of the Fox model is that harvest rate is proportional to fishing effort and the biomass of the stock; that is

$$H(E, N) = qEN,$$

where q is the catchability coefficient. Then, from (3), the equilibrium biomass N^* satisfies the relation

$$rN^* \ln\left(\frac{K}{N^*}\right) - qEN^* = 0$$

which implies that the nonzero equilibrium is given by

$$N^*(E) = Ke^{-\frac{q}{r}E}.$$

Therefore, the sustainable yield is

$$Y(E) = qEN^* = qK E e^{-\frac{q}{r}E} = E e^{kE+C}. \tag{4}$$

The highest possible sustainable yield is called the maximum sustainable yield, denoted Y_{MSY} . The maximum sustainable yield occurs when

$$\frac{dY}{dE} = qK \left(1 - \frac{q}{r}E\right) e^{-\frac{q}{r}E} = 0.$$

The corresponding optimal level of effort is given by

$$E_{MSY} = \frac{r}{q}.$$

Substituting this value of effort in (4), the maximum sustainable yield is

$$Y_{MSY} = K r e^{-1} = -\frac{e^{C-1}}{k}.$$

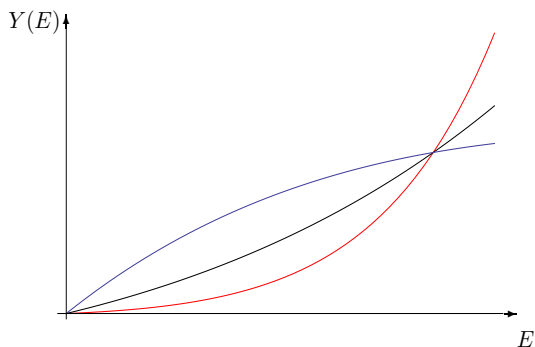


Fig. 3: Plots of the Fox model for some values of k and C

2.4 Schumacher equation

Next model was proposed independently by Terazaki [24], Johnson [14], Schumacher [21] and Michailoff [17]. The model assumes that the relative growth rate increases linearly with the squared inverse of time (see [4]):

$$\frac{1}{y} \frac{dy}{dt} = k \frac{1}{t^2}.$$

This leads us to model

$$y(t) = e^{C-k\frac{1}{t}}.$$

This model is primarily employed for timber growth and yield modeling (see e.g. [4, 13]).

3 Formularization of the problem

Parameters a and b of models of type (1) have to be estimated from the experimental or empirical data (w_i, t_i, y_i) , $i = 1, \dots, n$, $n \geq 3$, where $t_1 < t_2 < \dots < t_n$ denote the values of the independent variable, y_i are the respective measured function values and $w_i > 0$ are the

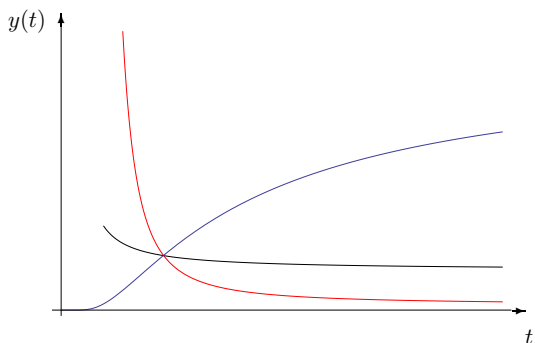


Fig. 4: Plots of the Schumacher model for some values of C and k

data weights which describe the assumed relative accuracy of the data.

There is no unique way to estimate the unknown parameters in a nonlinear regression function and many different methods have been proposed in literature (see e.g. [1, 2, 3, 11, 12, 20, 22]).

If the errors in the measurements of the independent variable are negligible, and the errors in the measurements of the dependent variable are independent random variables following the normal distribution with expectancy zero, i.e. that

$$y_i = f(t_i; a, b) + \varepsilon_i, \quad i = 1, \dots, n,$$

then in practical applications the unknown parameters a and b of model (1) are usually estimated in the sense of the least squares (LS) method by minimizing the functional

$$F_2(a, b) = \sum_{i=1}^n w_i \varepsilon_i^2 = \sum_{i=1}^n w_i [f(t_i; a, b) - y_i]^2$$

on the set $\mathcal{P} (\mathbb{R} \times (-\infty, 0], \mathbb{R} \times [0, -\infty) \text{ or } \mathbb{R}^2$. A point $(a^*, b^*) \in \mathcal{P}$ such that $F_2(a^*, b^*) = \inf_{(a,b) \in \mathcal{P}} F_2(a, b)$ is called the least squares estimate (LS estimate), if it exists.

Special numerical methods have been developed for the purpose of solving nonlinear LS problems (see e.g. [7]). However, prior to minimization itself difficult questions are posed referring to the existence and uniqueness of the LS estimate as well as the problem of determining a good initial approximation. In the next section we will prove existence result.

Models of type (1) belong to the class of models that can be linearized by transforming the dependent variables. This kind of transformation changes the error structure, as well as the influences of the data values. Because of that, it is not clear in which sense resulting parameters would be optimal. However, result obtained from transformed data usually gives a good initial approximation for iterative minimization methods.

Models of type (1) also belong to the family of the quasilinear regression models. Recent result on parameter estimation problem for quasilinear models can be found in [15].

4 The existence result

In this section we consider the existence of the best l_p -norm ($1 \leq p < \infty$) estimator in regression model of the form (1) where a and b are the unknown parameters, with

$$F(a, b) = \sum_{i=1}^n w_i |g(t_i)e^{a+bt_i} - y_i|^p \quad (5)$$

The next lemma will be used in the proof of Theorem 1.

Lemma 1. Suppose we are given data (w_i, t_i, y_i) , $i = 1, \dots, n$, $n \geq 3$, such that $t_1 < t_2 < \dots < t_n$ and $w_i, y_i > 0, i = 1, \dots, n$.

A If h is strictly increasing, then:

- (i) There exists a point in $\mathbb{R} \times (-\infty, 0)$ at which functional F defined by (5) attains a value less than $\sum_{i=2}^n w_i y_i^p$.
- (ii) There exists a point in $\mathbb{R} \times (0, \infty)$ at which functional F defined by (5) attains a value less than $\sum_{i=1}^{n-1} w_i y_i^p$.

B If h is strictly decreasing, then:

- (i) There exists a point in $\mathbb{R} \times (0, \infty)$ at which functional F defined by (5) attains a value less than $\sum_{i=2}^n w_i y_i^p$.
- (ii) There exists a point in $\mathbb{R} \times (-\infty, 0)$ at which functional F defined by (5) attains a value less than $\sum_{i=1}^{n-1} w_i y_i^p$.

Proof. We first prove A(i). Let $a(b) := \ln(\frac{y_1}{g(t_1)}) - bh(t_1)$. It is easy to verify that $f(t_1; a(b), b) = y_1$ and $\lim_{b \rightarrow -\infty} f(t_i; a(b), b) = 0$ for each $i = 2, \dots, n$.

By definition of the limit there exists a point $b_0 \in (-\infty, 0)$ such that for all $b \in (-\infty, b_0)$,

$$0 < f(t_i; a(b), b) < y_i, \quad i = 2, \dots, n.$$

Therefore, it follows that

$$F(a(b), b) = \sum_{i=1}^n w_i |f(t_i; a(b), b) - y_i|^p < \sum_{i=2}^n w_i y_i^p$$

for all $b \in (-\infty, b_0)$, and therefore claim A(i) holds.

A(ii) The proof is similar to that of part A(i). Let $a(b) := \ln(\frac{y_n}{g(t_n)}) - bh(t_n)$. It is easy to check that $f(t_n; a(b), b) = y_n$. Since $\lim_{b \rightarrow \infty} f(t_i; a(b), b) = 0$ for each $i = 1, \dots, n - 1$, there exists a point $b_0 \in (0, \infty)$ such that for all $b \in (b_0, \infty)$,

$$0 < f(t_i; a(b), b) < y_i, \quad i = 1, \dots, n - 1.$$

Therefore, for every $b \in (b_0, \infty)$ we have that $F(a(b), b) = \sum_{i=1}^n w_i |f(t_i; a(b), b) - y_i|^p < \sum_{i=1}^{n-1} w_i y_i^p$.

The rest of the proof is similar, so it is omitted. \square

Theorem 1. Let \mathcal{P} be one of the sets $\mathbb{R} \times (-\infty, 0]$, $\mathbb{R} \times [0, \infty)$ and \mathbb{R}^2 . If the data (w_i, t_i, y_i) , $i = 1, \dots, n$, $n \geq 3$, are

such that $t_1 < t_2 < \dots < t_n$ and $w_i, y_i > 0$, $i = 1, \dots, n$, then there exists a point $(a^*, b^*) \in \mathcal{P}$ such that

$$F(a^*, b^*) = \inf_{(a,b) \in \mathcal{P}} F(a, b).$$

Proof. The proof will be carried out for the case when h is strictly increasing. The proof for the second case (h is strictly decreasing) is essentially identical, so it will be omitted here.

Since functional F is nonnegative, there exists $F^* := \inf_{(a,b) \in \mathcal{P}} F(a, b)$. It should be shown that there exists a point $(a^*, b^*) \in \mathcal{P}$ such that $F(a^*, b^*) = F^*$.

By continuing the proof, let us note that Lemma 1. implies that

$$F^* < \min \left\{ \sum_{i=2}^n w_i y_i^p, \sum_{i=1}^{n-1} w_i y_i^p \right\}. \quad (6)$$

Let (a_k, b_k) be a sequence in \mathcal{P} , such that

$$F^* = \lim_{k \rightarrow \infty} F(a_k, b_k) = \lim_{k \rightarrow \infty} \sum_{i=1}^n w_i |g(t_i) e^{a_k + b_k h(t_i)} - y_i|^p. \quad (7)$$

Without loss of generality, in further consideration we may assume that sequences (a_k) and (b_k) are monotone. This is possible because the sequence (a_k, b_k) has a subsequence (a_{l_k}, b_{l_k}) , such that all its component sequences (a_{l_k}) and (b_{l_k}) are monotone; and since $\lim_{k \rightarrow \infty} F(a_{l_k}, b_{l_k}) = \lim_{k \rightarrow \infty} F(a_k, b_k) = F^*$.

Since each monotone sequence of real numbers converges in the extended real number system $\overline{\mathbb{R}}$, define

$$a^* := \lim_{k \rightarrow \infty} a_k, \quad b^* := \lim_{k \rightarrow \infty} b_k.$$

Note that $-\infty \leq a^*, b^* \leq \infty$.

To complete the proof it is enough to show that $(a^*, b^*) \in \mathcal{P}$, i.e. that $-\infty < a^* < \infty$ and $-\infty < b^* < \infty$. Indeed, the continuity of functional F will then imply that $F^* = \lim_{k \rightarrow \infty} F(a_k, b_k) = F(a^*, b^*)$.

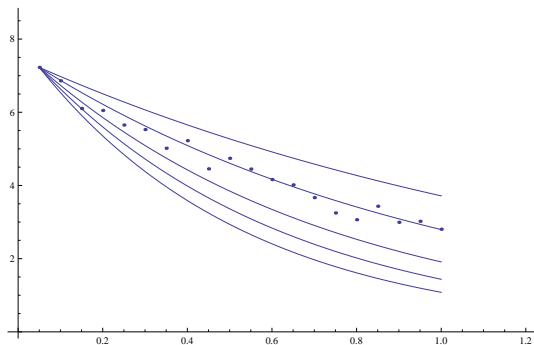


Fig. 5: Plots of $f(t; a(b), b)$ used in the proof of claim A(i) in Lemma 1.

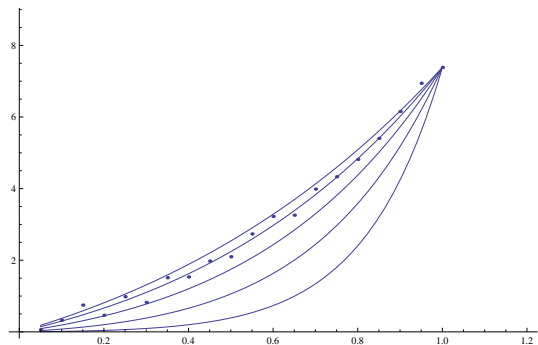


Fig. 6: Plots of $f(t; a(b), b)$ used in the proof of claim A(ii) in Lemma 1.

It remains to show that a^* and b^* are real numbers. To do this, let us denote

$$l_i^* := \lim_{k \rightarrow \infty} (a_k + b_k h(t_i)), \quad i = 1, \dots, n.$$

Note that $l_i^* \neq \infty$ for each $i = 1, \dots, n$. Indeed, if $l_i^* = \infty$ for some i , then it would follow from (7) that $F^* = \infty$, which is impossible. Also note that $l_i^* \neq -\infty$ for at least one index i because otherwise it would follow from (7) that $F^* = \sum_{i=1}^n w_i y_i^p$, which contradicts (6). Let index i_0 be such that $l_{i_0} \in \mathbb{R}$. Then only one of the following three cases can occur: (i) $i_0 = 1$, (ii) $i_0 = n$, or (iii) $1 < i_0 < n$. Now we are going to show that a^* and b^* are real numbers in each of these three cases, and so complete the proof of the theorem. Since $l_{i_0} = \lim_{k \rightarrow \infty} (a_k + b_k h(t_{i_0})) \in \mathbb{R}$, note that it will be enough to show that b^* is real. To do this, we will use the following identities

$$a_k + b_k h(t_i) = a_k + b_k h(t_{i_0}) + b_k (h(t_i) - h(t_{i_0})), \quad i = 1, \dots, n. \tag{8}$$

Case (i): $i_0 = 1$. If $b^* = \infty$, then it would follow from (8) that $l_i^* = \infty, i = 2, \dots, n$, which is impossible. If $b^* = -\infty$, then it would follow from (8) that $l_i^* = -\infty, i = 2, \dots, n$, and consequently $F^* \geq \sum_{i=2}^n w_i y_i^p$, which contradicts (6).

Case (ii): $i_0 = n$. If $b^* = \infty$, then it would follow from (8) that $l_i^* = \infty, i = 1, \dots, n-1$, and consequently $F^* \geq \sum_{i=1}^{n-1} w_i y_i^p$, which contradicts (6). If $b^* = -\infty$, then it would follow from (8) that $l_i^* = -\infty, i = 1, \dots, n-1$, which is impossible.

Case (iii): $1 < i_0 < n$. If $b^* = \infty$, then it would follow from (8) that $l_{i_0+1}^* = \infty$, which is impossible. If $b^* = -\infty$, then it would follow from (8) that $l_{i_0-1}^* = -\infty$, which is impossible.

Herewith we completed the proof. \square

5 Numerical examples

Further in the text, we will give few illustrative examples.

Example 1.

Let us consider data set given in the Table 1.

This data represent observed average height of Cryptomeria Japonica of quality I in the Main Island and Kiushu (source [24]). To fit this date, we will use following two-parameter model:

$$y(t) = ae^{\frac{b}{t}}.$$

Optimal parameters will be obtain in the sense of (unweighted) least squares, i. e. by minimizing functional

$$F(a, b) = \sum_{i=1}^{29} (ae^{\frac{b}{t_i}} - y_i)^2.$$

As a initial approximation, we use the result of linearized model

$$(a^0, b^0) = (26.77, -23.46),$$

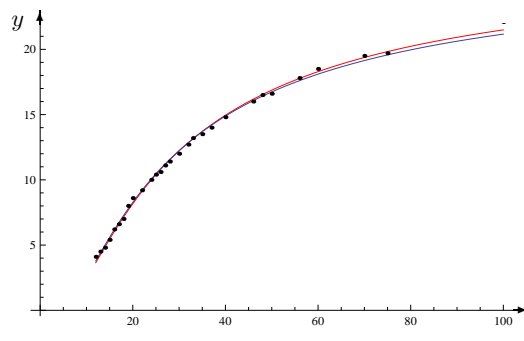


Fig. 7: The data, initial fit (blue), and optimal fit (red)

with $F(a^0, b^0) = 1.966131$.

The results of nonlinear LS are given in Table 2

Table 2: Least squares parameter estimates and the corresponding value of functional F

a^*	b^*	$F(a^*, b^*)$
27.37	-24.18	1.443361

The data and graphs of inital and optimal fit are shown in Figure 7.

Example 2.

In this example we will illustrate influence of weights to resulting fit. We will consider the population growth of the United States of America in the period between 2000 and 2013. We will use the data from the period between 2000 and 2010 to build an exponential growth model (exponential regression) using few different sets of weights. The data from 2011 to 2013 will be used to test how well obtained model fit the data few years in advance. Let us present the data in Table 3.

Besides unweighted least squares where all weights are equal 1 (w^0), we will use following sets of (normalized) weights:

- linear (w^1): $w_i = \frac{x_i}{\sum_{k=1}^{11} x_k}$,
- square (w^2): $w_i = \frac{x_i^2}{\sum_{k=1}^{11} x_k^2}$,
- exponential (w^e): $w_i = \frac{e^{x_i/2}}{\sum_{k=1}^{11} e^{x_k/2}}$,

Table 1: The observed average height

Age (t_i)	12	13	14	15	16	17	18	19	20	22	24	25	26	27	28
Height (y_i)	4.1	4.5	4.8	5.4	6.2	6.6	7	8	8.6	9.2	10	10.4	10.6	11.1	11.4
Age (t_i)	30	32	33	35	37	40	46	48	50	56	60	70	75	100	
Height (y_i)	12	12.7	13.2	13.5	14	14.8	16	16.5	16.6	17.8	18.5	19.5	19.7	22.1	

Table 3: US population (Source: United States Bureau of the Census)

Year (t_i)	2000	2001	2002	2003	2004	2005	2006
Population (y_i)	282162411	284968955	287625193	290107933	292805298	295516599	298379912
Year (t_i)	2007	2008	2009	2010	2011	2012	2013
Population (y_i)	301231207	304093966	306771529	309326295	311582564	313873685	316128839

–modified exponential (w^{me}): $w_i = \frac{e^{x_i^2}}{\sum_{k=1}^{11} e^{x_k^2}}$,

for $i = 1, \dots, 11$.

The results are given in Table 4, where prediction error (PE) is calculated as

$$PE = \sum_{k=12}^{14} (e^{a^*+b^*t_i} - y_i)^2,$$

for all sets of weights.

We can conclude that for this particular data, weights w^{me} give the best prediction. That could be interpreted in a way that the more recent data have a bigger influence for the data in the near future.

The data and graphs of initial and optimal fit with weights w^{me} are shown in Figure 8.

Acknowledgement

The authors acknowledge the financial support for the research by J. J. Strossmayer University of Osijek (project “Parameter estimation problem in some two-parameter monotonic mathematical model”).

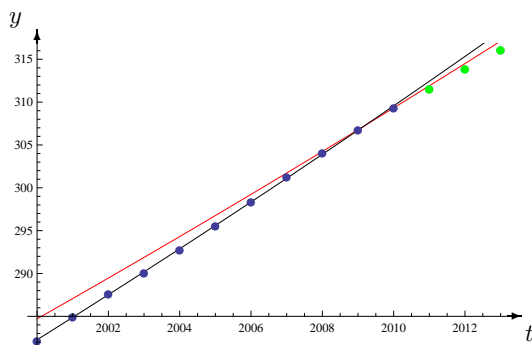


Fig. 8: Plots of the data (blue dots), test data (green dots), initial fit (blue) and optimal fit (red)

Table 4: Least squares parameter estimates and the corresponding prediction errors

	a^*	b^*	PE
w^0	-12.8227	0.00923279	7.37169
w^1	-12.8226	0.00923274	7.37054
w^2	-12.8225	0.00923268	7.36939
w^e	-12.5298	0.00908687	5.48948
w^{me}	-10.9354	0.00829343	1.47465

References

- [1] A. Atieg, G.A. Watson, Use of l_p norms in fitting curves and surfaces to data, ANZIAM J. **45**(E) C187-C200 (2004)
- [2] D.M. Bates, D.G. Watts, Nonlinear Regression Analysis and its Applications, Wiley, New York, 1988.
- [3] A. Björck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, PA, 1996.
- [4] H. E. Burkhart, M. Tomé, Modeling Forest Trees and Stands, Springer, Dordrecht 2012
- [5] N. Chater, G.D.A. Brown, Scale-invariance as a unifying psychological principle, Cognition **69** B17-B24 (1999)
- [6] C. W. Clark, Mathematical bioeconomics: the optimal management of renewable resources, Wiley, New York, 1990
- [7] J.E. Dennis, R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM, Philadelphia, PA, 1996.
- [8] W.W. Fox, An exponential surplus yield model for optimising exploited fish populations, Transactions of the American Fisheries Society **99** 80-88 (1970)
- [9] F.R. Giordano, W.P. Fox, S.B. Horton, A First Course in Mathematical Modeling, Brooks/Cole Publishers, Boston, 2014.
- [10] Z. Frica, M. Konvicka, Dispersal kernels of butterflies: Power-law functions are invariant to marking frequency, Basic and Applied Ecology **8** 377-386 (2007)
- [11] P.E. Gill, W. Murray, M.H. Wright, Practical Optimization, Academic Press, London, 1981.

- [12] R. Gonin, A.H. Money, *Nonlinear Lp-Norm Estimation*, Marcel Dekker, New York, 1989.
- [13] J. Imaña-Encinas, O.A. Santana, C.R. Imaña, Volumetric and economic optimal rotations for firewood production of *Eucalyptus urophylla* in Ipameri, state of Goias, Floresta, Curitiba **41** 905-912 (2011)
- [14] N.O. Johnson, A trend line for growth series, *J. Am. Stat. Assoc.* **30** 717 (1935)
- [15] D. Jukić, A simple proof of the existence of the best estimator in a quasilinear regression model, *Journal of optimization theory and applications* **162** 293-302 (2014)
- [16] P.A. Marquet et al., Scaling and power-laws in ecological systems, *The Journal of Experimental Biology* **208** 1749-1769 (2005)
- [17] I. Michailoff, Zahlenmäßiges Verfahren für die Ausführung der Bestandeshöhenkurven, *Forstw. Cbl. U. Thar. Forstl. Jarhrb.* **6** 273-279 (1943)
- [18] A.R. Overman, R.V. Scholtz, *Mathematical Models of Crop Growth and Yield*, CRC Press, 2002.
- [19] M.J. Panik, *Growth Curve Modeling: Theory and Applications*, John Wiley, New Jersey, 2014.
- [20] G.J.S. Ross, *Nonlinear Estimation*, Springer, New York, 1990.
- [21] F.X. Schumacher, A new growth curve and its application to timber-yield studies. *Journal of Forestry* **37** 819-820 (1939).
- [22] G.A.F. Seber, C.J. Wild, *Nonlinear Regression*, John Wiley, New York, 1989.
- [23] A. Spirling, The next big thing: scale invariance in political science, <http://www.people.fas.harvard.edu/~spirling/documents/powerlawSend.pdf>
- [24] W. Terazaki, Notes on the Analytical Interpretation of Growth Curves for Single Tree and Stands, and on Application for the Construction of Yield Table for *Cryptomeria Japonica*, *Bulletin of the Forest Experiment Station*, 151-202 (1915)
- [25] D. Zajdenweber, Scale Invariance in Economics and in Finance, *Scale Invariance and Beyond* **7** 185-194 (1997)



Darija Marković holds the B.Sc. degree in mathematics and computer science from University of Osijek, as well as the M.Sc. in mathematics from the University of Zagreb. In 2009 she obtained a Ph.D. in mathematics from the University of Zagreb.

Presently she is an assistant professor at the Department of Mathematics, University of Osijek. Her research interests cover applied and numerical mathematics, specifically parameter estimation, mathematical modelling, nonlinear least squares and smoothing methods. She has co-authored several papers in journals and conference proceedings on these subjects.



Luka Borozan obtained his BSc degree at the J.J. Strossmayer University of Osijek, Department of Mathematics. Currently he is a MSc student at the same university. His research interests are numerical computation and software engineering.