

Developing A New Dimension of an Applied Exponential Model: Application in Biological Sciences

W. M. A. W. Ahmad
School of Dental Sciences
Universiti Sains Malaysia, Malaysia

R. A. A. Rohim
Universiti Sains Malaysia
Malaysia

Y. Norhayati
Universiti Sains Malaysia
Malaysia

N. A. Aleng
School of Informatics and Applied Mathematics
Universiti Malaysia Terengganu
Malaysia

Z. Ali
School of Mathematical Sciences
Universiti Sains Malaysia
Malaysia

Abstract—Modeling of exponential growth or decay is a nonlinear regression technique. In the real world, the exponential growth is often used to model population growth while the exponential decay is often used to a model declining population or a decreasing size. In this study, we try to improve the performance of exponential growth by adding bootstrap and fuzzy techniques. This gives us the option to perform analysis even when there is not enough data. The aim of the current work is to develop a new dimension of an applied exponential analysis with improved results. The suggested method was tested and applied to biological data. The gathered data was compared by measuring the average width of the predicted interval using least squares method and fuzzy method. The result shows that the average width of the predicted interval using least squares method was 0.522 while using fuzzy method was 0.082. This indicated the superiority of the fuzzy regression methodology. Besides that, this paper provides the algorithm for the prediction of cell growth and inferences.

Keywords—bootstrap; exponential growth; fuzzy regression; exponent decay; nonlinear regression

I. INTRODUCTION

The exponential function is one of the most important and widely occurring functions in physics and biology [1]. The exponential distribution has many applications in biology. Examples of the exponential application include the bacteria growth time and the decay time of bacterial pathogens with constant failure rates. Exponent distribution of an event is the probability of the event occurring in the next small time interval which does not vary through time, and time between events has a Poisson distribution. The exponential growth or decay follows a function N which changes with time in such a way that the change ΔN in N during a short time interval Δt is proportional to N and to Δt , as $\Delta N = k\Delta t$, rearranging this equation, we obtain:

$$\frac{\Delta N}{\Delta t} = k N \quad (1)$$

Therefore, the constant of proportionality k can be seen by rearranging equation (1). Then, we obtain $k = (\Delta N/N)/\Delta t$. The constant k is then the fractional change $(\Delta N/N)$ in N per unit time Δt . The dimension of k is $(\text{time})^{-1}$ [2].

II. STATISTICAL THEORY AND METHODOLOGY

The linear regression model is a powerful method for modeling and forecasting, especially in conventional regression analysis. For modeling purposes, data should be crisp and should follow a normal assumption, this will lead to better significant results [3]. In the current study, an exponential model was applied after transforming the data to a linear model with mathematical programming by assuming that the dependent variables were crisp while the independent variable was a symmetric fuzzy number. This paper provides an algorithm for the exponential growth model using cell cultured dataset. The basic analysis is to transform the nonlinear equation into a linear form. After transforming into linear regression, we will obtain an equation in the form of $Y = \beta_0 + \beta_i x_i + \varepsilon$. We transformed the nonlinear equation in order to get better results and better significant inferences. The random error term is added to make the model probabilistic rather than deterministic. The value of the coefficient β_i determines the contribution of the independent variables and β_0 is the y -intercept [3, 4]. To be more accurate, a fuzzy regression can be rearranged as

$$Y = Z_0 + Z_1 x_1 \quad (2)$$

where the explanatory variables x_i are assumed to be precise. However, according to (2), response variable Y is not crisp but is instead fuzzy in nature. For the fuzzy approach, Z_i are

assumed symmetric fuzzy numbers which can be presented by the interval. Z_i can be expressed as a fuzzy set given by $Z_i = \langle a_{i_c}, a_{i_w} \rangle$ where a_{i_c} is the center and a_{i_w} is radius or associated vagueness. This reflects the confidence in the regression coefficients around a_{i_c} in terms of symmetric triangular membership function. So, the relationship is also considered to be fuzzy. This $Z_i = \langle a_{i_c}, a_{i_w} \rangle$ can be written as $Z_i = [a_{i_{L}}, a_{i_{R}}]$ with $a_{i_{L}} = a_{i_c} - a_{i_w}$ and $a_{i_{R}} = a_{i_c} + a_{i_w}$. In fuzzy regression methodology, parameters are estimated by minimizing total vagueness in the model. Using $Z_i = \langle a_{i_c}, a_{i_w} \rangle$, we can write the fuzzy regression as

$$y = \langle a_{0_c}, a_{0_w} \rangle + \langle a_{1_c}, a_{1_w} \rangle x \tag{3}$$

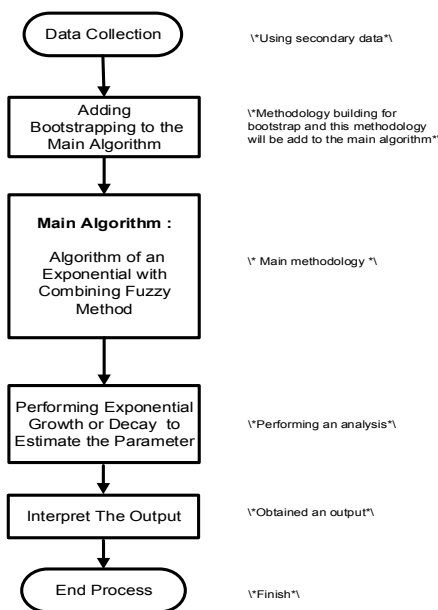


Fig. 1. Flowchart of the integrated exponential calculation.

Thus this can be written as $y_{j_c} = a_{0_c} + a_{1_c} x_{1j}$ then it can be written straightly as $y_{j_w} = a_{0_w} + a_{1_w} |x_{1j}|$. As y_{j_w} represents radius and so cannot be negative, therefore on the right-hand side of the equation $y_{j_w} = a_{0_w} + a_{1_w} |x_{1j}|$, absolute values of x_{ij} are taken. Suppose there are m data points, each comprising of $a(n+1)$ -row vector. The parameters Z_i are estimated by minimizing the quantity, which is the total vagueness of the model-data set combination, subject to the constraint that each data point must fall within the estimated value of response variable. This can be visualized as the following linear programming problem, minimized $\sum_{j=1}^m (a_{0_w} + a_{1_w} |x_{1j}|)$ and subject

$$\text{to: } \left\{ \begin{aligned} (a_{0_c} + a_{1_c} x_{1j}) + (a_{0_w} + a_{1_w} |x_{1j}|) &\geq Y_j \\ (a_{0_c} + a_{1_c} x_{1j}) - (a_{0_w} + a_{1_w} |x_{1j}|) &\leq Y_j \end{aligned} \right.$$

and $a_{i_w} \geq 0$. A simple procedure is commonly used to solve the linear programming problem [5]. The present data is a sample of the results obtained from [6] which characterize the proliferative capacity of mesenchymal stem cells. The data are composed of two variables which are the days of culture (X) and population doubling level (Y) – see Table I.

TABLE I. X AND Y DESCRIPTION

Variable	Description
X	A total number of time (days) between each cell passage.
Y	Population doubling level is the proportion of cells count at 80% confluency over the original number of cells seeded.

A. Exponential Growth Transforming to Linear Form

Exponential growth formula and exponential decay formula are given by $Y = Ae^{bX}$. The procedure to transform the growth and decay formula into a linear form follows. The equation for the linear form for exponential growth is given by (4):

$$\ln Y = \ln(Ae^{bX}) = \ln(A) + \ln(e^{bX}) = \ln(A) + bx \tag{4}$$

B. Calculations of an Exponential Cell Growth using SAS Algorithm

- First step: Data for exponential cell growth should enter in SAS algorithm as follows.

```

Data Cell_Growth;
input x y lny;
datalines;
0.00      38.00      3.64
5.00      39.31      3.67
8.00      39.74      3.68
10.00     40.98      3.71
13.00     43.10      3.76
17.00     45.78      3.82
20.00     59.15      3.89
22.00     49.90      3.91
24.00     53.98      3.99
28.00     57.46      4.05
31.00     61.03      4.11
34.00     63.80      4.16
37.00     65.52      4.18
40.00     68.54      4.23
44.00     72.62      4.29
47.00     75.42      4.32
50.00     79.38      4.37
53.00     83.31      4.42
;
run;
    
```

- Second step: Adding bootstrapping to the calculation. The following algorithm calculates the data using a bootstrap method and prints out the data.

```

%MACRO bootstrap(data=_last_,booted=booted,boots=2,
seed=1234);
DATA &booted;
pickobs = INT(RANUNI(&seed)*n)+1;
SET &data POINT = pickobs NOBS = n;
REPLICATE=int(i/n)+1;
    
```

```
i+1;
IF i > n*&boots THEN STOP;
RUN;
%MEND bootstrap;
ods rtf file='abc.rtf' style=journal;
%bootstrap(data= Cell_Growth, boots=2);
run;
proc print data=booted;
run;
```

- Third step: estimating the parameters of an exponential equation. The algorithm below is used to estimate parameters according to the data.

```
Title "Exponential Equation";
ods graphics / imagename = "Exponential Equation";
proc nlin data=booted plots=fit;
parameters A=1 b=-1;
model y = A*exp(b*x);
ods output EstSummary=summExp;
run;
```

- Fourth step: Estimating the linear form of regression according to the transform bootstrap data.

```
Proc reg data=booted;
model lny=x;
run;
ods rtf
close;
```

C. Parameter Estimation for Exponential Cell Growth

The SAS output for the parameter estimation follows. The summary of the output follows.

TABLE II. PARAMETER ESTIMATE OF AN EXPONENTIAL EQUATION

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
A	35.9723	0.2566	35.4509	36.4937
b	0.0160	0.000219	0.0155	0.0164

The obtained equation is given by (5):

$$Y = 35.97e^{0.016X} \tag{5}$$

Exponential growth formula is given by $Y = Ae^{bX}$. After parameter estimation, we obtain (6). Using the equation of exponential growth equation we can estimate the growth of cell at the certain point $\hat{Y} = 35.97e^{0.016X}$. Taking a logarithm, we obtain

$$\ln y = \ln(35.9735e^{0.016X}) = \ln(35.9735) + \ln(e^{0.016X}) \Rightarrow$$

$$\ln y = 3.58278 + 0.0160x$$

So the parameter estimation and standard error are given as

$$\ln y = 3.58278 + 0.0160x$$

$$Std\ Error = (0.2566)(0.000219) \tag{6}$$

The upper limits of prediction interval for the exponential model are computed using (7):

$$\ln y = (3.58278 + 0.2566) + (0.0160 + 0.000219)x \tag{7}$$

$$\ln y = (3.83938) + (0.016219)x$$

and the lower limits of prediction interval for the exponential model are computed using (8):

$$\ln y = (3.58278 - 0.2566) + (0.0160 - 0.000219)x \tag{8}$$

$$\ln y = (3.32618) + (0.015781)x$$

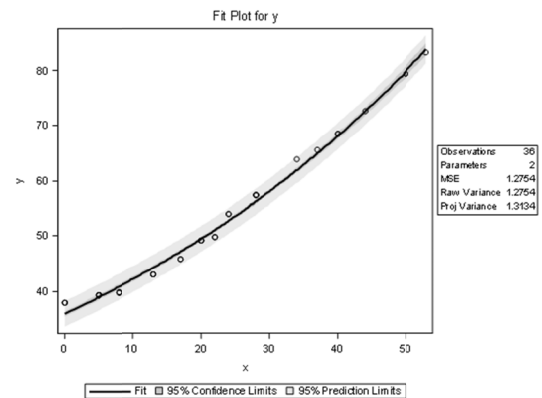


Fig. 2. The plot of an exponential

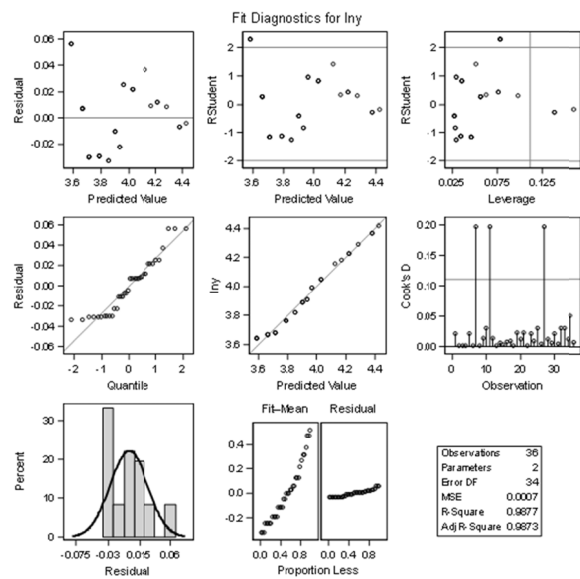


Fig. 3. Fit diagnostic for lny.

D. Calculation of Fuzzy Least Squares(FLS) For Exponential Growth

The algorithm below calculates the fuzzy least squares (FLS) for exponential growth. The full set of calculations can be visualized by the following programming. Table III summarizes the optimization results for the parameter estimates.

```
ods rtf file='abc.rtf' style=journal;
```

```

proc nlp;
min Y;
decvar a0c a0w a1c a1w;
bounds a0w>=0, a1w>=0;
lincon a0c+13*a1c-a0w-13*a1w<=3.76;
lincon a0c+5*a1c-a0w-5*a1w<=3.67;
lincon a0c+20*a1c-a0w-20*a1w<=3.89;
lincon a0c+5*a1c-a0w-5*a1w<=3.67;
lincon a0c+13*a1c-a0w-13*a1w<=3.76;
lincon a0c+5*a1c-a0w-5*a1w<=3.67;
lincon a0c+0*a1c-a0w-0*a1w<=3.64;
lincon a0c+5*a1c-a0w-5*a1w<=3.67;
lincon a0c+24*a1c-a0w-24*a1w<=3.99;
lincon a0c+8*a1c-a0w-8*a1w<=3.68;
lincon a0c+0*a1c-a0w-0*a1w<=3.64;
lincon a0c+24*a1c-a0w-24*a1w<=3.99;
lincon a0c+5*a1c-a0w-5*a1w<=3.67;
lincon a0c+50*a1c-a0w-50*a1w<=4.37;
lincon a0c+53*a1c-a0w-53*a1w<=4.42;
lincon a0c+40*a1c-a0w-40*a1w<=4.23;
lincon a0c+22*a1c-a0w-22*a1w<=3.91;
lincon a0c+20*a1c-a0w-20*a1w<=3.89;
lincon a0c+17*a1c-a0w-17*a1w<=3.82;
lincon a0c+28*a1c-a0w-28*a1w<=4.05;
lincon a0c+17*a1c-a0w-17*a1w<=3.82;
lincon a0c+20*a1c-a0w-20*a1w<=3.89;
lincon a0c+13*a1c-a0w-13*a1w<=3.76;
lincon a0c+22*a1c-a0w-22*a1w<=3.91;
lincon a0c+8*a1c-a0w-8*a1w<=3.68;
lincon a0c+37*a1c-a0w-37*a1w<=4.18;
lincon a0c+0*a1c-a0w-0*a1w<=3.64;
lincon a0c+28*a1c-a0w-28*a1w<=4.05;
lincon a0c+50*a1c-a0w-50*a1w<=4.37;
lincon a0c+13*a1c-a0w-13*a1w<=3.76;
lincon a0c+44*a1c-a0w-44*a1w<=4.29;
lincon a0c+8*a1c-a0w-8*a1w<=3.68;
lincon a0c+8*a1c-a0w-8*a1w<=3.68;
lincon a0c+28*a1c-a0w-28*a1w<=4.05;
lincon a0c+34*a1c-a0w-34*a1w<=4.16;
lincon a0c+40*a1c-a0w-40*a1w<=4.23;
lincon a0c+13*a1c+a0w+13*a1w>=3.76;
lincon a0c+5*a1c+a0w+5*a1w>=3.67;
lincon a0c+20*a1c+a0w+20*a1w>=3.89;
lincon a0c+5*a1c+a0w+5*a1w>=3.67;
lincon a0c+13*a1c+a0w+13*a1w>=3.76;
lincon a0c+5*a1c+a0w+5*a1w>=3.67;
lincon a0c+0*a1c+a0w+0*a1w>=3.64;
lincon a0c+5*a1c+a0w+5*a1w>=3.67;
lincon a0c+24*a1c+a0w+24*a1w>=3.99;
lincon a0c+8*a1c+a0w+8*a1w>=3.68;
lincon a0c+0*a1c+a0w+0*a1w>=3.64;
lincon a0c+24*a1c+a0w+24*a1w>=3.99;
lincon a0c+5*a1c+a0w+5*a1w>=3.67;
lincon a0c+50*a1c+a0w+50*a1w>=4.37;
lincon a0c+53*a1c+a0w+53*a1w>=4.42;
lincon a0c+40*a1c+a0w+40*a1w>=4.23;
lincon a0c+22*a1c+a0w+22*a1w>=3.91;
lincon a0c+20*a1c+a0w+20*a1w>=3.89;
lincon a0c+17*a1c+a0w+17*a1w>=3.82;
lincon a0c+28*a1c+a0w+28*a1w>=4.05;

```

```

lincon a0c+17*a1c+a0w+17*a1w>=3.82;
lincon a0c+20*a1c+a0w+20*a1w>=3.89;
lincon a0c+13*a1c+a0w+13*a1w>=3.76;
lincon a0c+22*a1c+a0w+22*a1w>=3.91;
lincon a0c+8*a1c+a0w+8*a1w>=3.68;
lincon a0c+37*a1c+a0w+37*a1w>=4.18;
lincon a0c+0*a1c+a0w+0*a1w>=3.64;
lincon a0c+28*a1c+a0w+28*a1w>=4.05;
lincon a0c+50*a1c+a0w+50*a1w>=4.37;
lincon a0c+13*a1c+a0w+13*a1w>=3.76;
lincon a0c+44*a1c+a0w+44*a1w>=4.29;
lincon a0c+8*a1c+a0w+8*a1w>=3.68;
lincon a0c+8*a1c+a0w+8*a1w>=3.68;
lincon a0c+28*a1c+a0w+28*a1w>=4.05;
lincon a0c+34*a1c+a0w+34*a1w>=4.16;
lincon a0c+40*a1c+a0w+40*a1w>=4.23;
Y= a0w*36+727*a1w;
run;
ods rtf close;

```

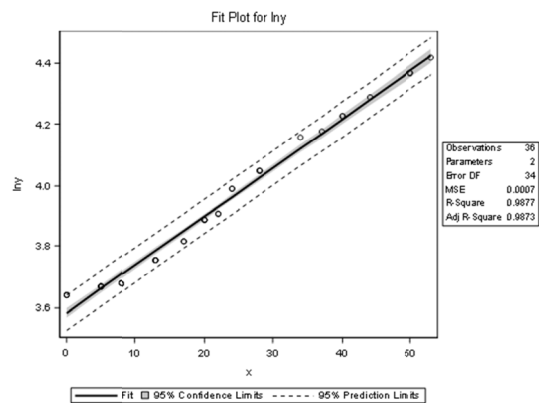


Fig. 4. Plot lny vs x

E. Parameter Estimation for Fuzzy Least Square Exponential Cell Growth

Parameter estimates are given: $a0c=3.598824$, $a0w=0.041176$, $a1c=0.015294$ and $a1w=0.000000$. The fuzzy upper prediction limit for exponential model is computed using the equation as follows:

$$\begin{aligned}
 \ln y &= (3.598824 + 0.041176) \\
 &+ (0.015294 + 0.000000)x \tag{9} \\
 \ln y &= 3.64000 + 0.015294x
 \end{aligned}$$

and the lower limit of prediction interval for the exponential model is computed using the equation:

$$\begin{aligned}
 \ln y &= (3.598824 - 0.041176) \\
 &+ (0.015294 - 0.000000)x \tag{10} \\
 \ln y &= 3.557648 + 0.015294x
 \end{aligned}$$

Table IV shows that fuzzy regression methodology is capable of handling situations in which predictor variables are highly correlated. From this Table, the average width for LS was found to be 0.522, while the one for FR was only 0.082,

indicating thereby the superiority of fuzzy regression methodology.

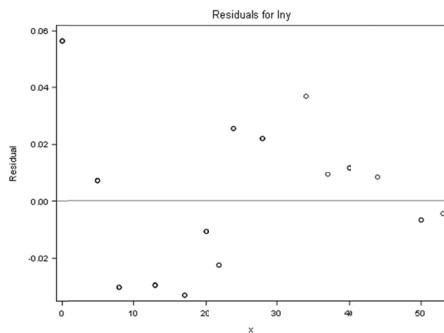


Fig. 5. Residual for ln y

TABLE III. OPTIMIZATION RESULTS

Parameter	Estimate	Gradient Objective Function	Active Bound Constraint
a0c	3.598824	0	
a0w	0.041176	36.000000	
a1c	0.015294	0	Lower BC
a1w	0.000000	727.000000	

Value of objective function=1.4823529412

TABLE IV. AVERAGE WIDTH FOR FITTED REGRESSION MODELS

Method of Least Squares (LS)		Method of Fuzzy Regression (FR)	
$(LP)_i - (UP)_i$	$\sum_{i=1}^{36} (Width)_i / 36$	$(LP)_i - (UP)_i$	$\sum_{i=1}^{36} (Width)_i / 36$
$i = 1, 2, \dots, 36$	= 0.522	$i = 1, 2, \dots, 36$	= 0.082

LP: Lower of width prediction, UP Upper of width prediction

III. SUMMARY AND DISCUSSION

This paper gives an explanation for the alternative programming method of bootstrap approach to exponential growth cell modeling nonlinear regression procedure using SAS software. The aim of the algorithm is to provide the researcher with an alternative programming of data analysis with good accuracy prediction result. Fuzzy model and crisp data developed a better accuracy model compared to the conventional one. The difference on the result can be seen in Table IV. In [1], authors discussed Zadeh's extension principle to classical crisp estimated to the least squares method. They found that modified linear estimates often lead to better performance of an estimated model. They concluded that fuzzy linear regression model had better performance compared to LS method.

ACKNOWLEDGMENT

Authors would like to express their gratitude to Universiti Sains Malaysia for providing the research funding (Grant no.1001/PPSG/8012278, School of Dental Sciences, University Sains Malaysia).

REFERENCES

[1] R. J. Tallarida, R. B. Murray, Manual of Pharmacologic Calculations with Computer Programs, Springer, 1981

[2] A. A Bartlett, "The exponential function-Part I", The Physics Teacher Vol. 14, No. 7, pp. 393-401, 1976

[3] W. M. A. W. Ahmad, A. Nor Azlida, D.Yosza, A. H. Nurfadhina, H. Ruhaya, A. Zalila., B. Adam, Z. Syerrina, "Applied exponential growth regression modeling using SAS: An Alternative Method of Biostatistics", International Journal of Applied Engineering Research, Vol. 12, No. 18, pp. 7853-7856, 2017

[4] T. H. D. Ngo, C. A. La Puente, C. A. "The Steps to Follow in a Multiple Regression Analysis", SAS Global Forum 2012: Statistics and Data Analysis, Orlando, USA, Paper 333-2012, April 22-25, 2012

[5] H. Ghosh, S. Wadhwa, Application of fuzzy regression methodology in agriculture using SAS, Indian Agricultural Statistics Research Institute (IASRI), 2001

[6] N. Yusop, P. Battersby, A. Alraies, A. J. Sloan, R. Moseley, R. J. Waddington. "Isolation and characterization of mesenchymal stem cells from rat bone marrow and the endosteal niche: a comparative study", Stem Cells International, Vol. 2018, Article ID 6869128, 2018

[7] R. Korner, W. Nather, "Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates", Information Sciences, Vol. 109, No. 1-4, pp. 95-118, 1998