ISSN : 1812-5379 (Print)
ISSN : 1812-5417 (Online)
http://ansijournals.com/ja

JOURNAL OF AGRONOMY

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Pakistan Journal of Agronomy, 2 (4): 223-236, 2003
ISSN 1680-8207

Fitting Nonlinear Gompertz Curve to Tobacco Growth Data

Zuhaimy Ismail, [1]Azme Khamis and [2]Md. Yunus Jaafar
Jabatan Matematik, Fakulti Sains, Universiti Teknologi Malaysia,
81310 UTM Skudai, Johor, Malaysia
[1]Pusat Pengajian Sains, Kolej Universiti Teknologi Tun Hussein Onn,
86400 Parit Raja, Batu Pahat, Johor, Malaysia
[2]Program Statistik Gunaan, MARDI, Peti Surat 12301, GPO, Kuala Lumpur, Malaysia

Abstract: This study is to discuss the application of nonlinear Gompertz curve to measure the growth data. Data used are the growth of leaves, stem and roots of tobacco. The sample are divided into two and measured in kilograms. By using specific starting values, it is found that the nonlinear Gompertz curve is suitable to match the growth of leaves, stem and roots of tobacco plant. It is found that both samples have the sum squares error, which is low and the variance analysis conducted showing that this model is statistically significant. Furthermore, it is supported by the as asymptotic correlation matrix value among the parameter estimated much lower.

Key words: Nonlinear gompertz, growth data, tobacco

Introduction

The analysis of growth data becomes more important in many fields of study. Biologists are interested in the description of biological growth and in trying to understand its underlying biological process. Chemists are interested in the formulation of the product of a reaction of two or more chemical substance. In agriculture there are obvious economic and management advantages in knowing how large things grow, how fast they grow and how does these factors respond to environmental conditions or treatment over time. Social scientists are interested, for example, in the growth of populations, new political parties, the foods supply and energy demand. The same types of model occur when the explanatory variable x is no longer time but increasing intensity of some other factors, such as the growth of smog with increase in solar radiation, weight gains with increased nutrient in diet, changes in crop yield with increase density of planting and so on.

There have been two main approaches to the analysis of growth curve data, namely statistical and biological. Statistical approach is purely empirical and involves fitting polynomial curves to the data using univariate or multivariate models. These polynomial methods can provide useful predictive information and may be the best approach if the growth information has been collected over the limited range of the growth cycle. However, their parameters have no physical interpretation and they do not model subject-matter knowledge of the growth process, for example the fact that the size of leaves of tree tends to stabilize after a certain period.

For biological growth, the mechanistic approach is often termed 'biological'. These approaches search a model with a biological basis and biologically interpretable parameters.

Research methodology: Gompertz growth curve

The Gompertz model, based upon a model given by Gompertz in 1825 for the hazard in life table, then used as growth model by Wright (1926). Its initial formulation was largely empirical, but later Medawar (1940) derived it as a growth model for the heart of a chicken. The Gompertz model is very popular and used in various fields such as population studies and animal growth in situations where growth is not symmetrical about the point of inflection. The growth rate (Seber and Wild 1989) is

$$\frac{df}{dt} = kf\,(\log\alpha - \log f) \text{ for } \kappa > 0 \text{ and } \alpha > 0. \quad (1)$$

and the relative growth rate now declines with log(size). From (1) we obtain the Gompertz function as

$$f(t) = \alpha \exp\{-e^{-\kappa(t-\gamma)}\} \quad (2)$$

The time inflection for (2) is at time t = γ with f = α/e, at which point the maximum growth rate, w_{max} = $\kappa\alpha$/e, occurs. For Gompertz curve log (f) is monomolecular. Also the Gompertz function ha the property that any power of f is also Gompertz function.

Ratkowsky (1983) suggested the nonlinear Gompertz curve as

$$f(t) = ae^{-e^{\beta-\gamma t}} \quad (3)$$

where $\alpha = f_0$ is the minimum or maximum value of growth organism or crop yield.

Parameter estimation

Nonlinear models are more difficult to specify and estimate than linear models. Instead of simply listing exploratory variable, we must write the regression expression, declare parameter names, guess starting values for them and possibly specific derivatives of the model respects to the parameters. Some models are difficult to fit and there is no guarantee that the procedure will be able to fit the model successfully. On the other hand, we have to try and error process, but experience may be useful. By using SAS package, Proc nlin (short form for nonlinear) fits nonlinear regression models by least square. For each nonlinear model to be analyzed, we must specify the following;

- The names and starting values of the parameters to be estimated.
- The model (using a single dependent variable).
- Partial derivative of the model with respect to each parameter (except for Doesn't Use Derivative (DUD) method).
- The second derivatives of the model with respect to each parameter (Only for the Newton method)

Example for SAS programming on nonlinear modeling.
Proc nlin option;

Model dependent=expression;
Parameter|Parms parameter=values……..;
Bound expression ….;
By variables;
Der. parameter [.parameter] = expression;
Id variables;
Output out=SAS dataset keywords=names;

SAS package via Proc nlin, give five alternatives for iteration method such as steepest descent or gradient method (Gradient), Newton method (Newton), Modified Gauss-Newton (Gauss), Marquardt method (Marquardt) and multivariate secant or false position method or Doesn't Use Derivative (DUD).

Computational application

For the system of equation represented by the nonlinear model

$$Y = F(X_1, X_2, \ldots, X_n, \beta_0, \beta_1, \ldots, \beta_p,) + \epsilon$$
$$Y_t = F(\beta, X_t) + \epsilon_t \quad (4)$$

where X is a matrix of the independent variables β is a vector of unknown parameters, ϵ is the random error vector and F is a function of the independent variables and the parameters, Y_t is the observed value on the tth experiments (t =1, 2, …,n). The least squares estimate of β^* denoted by minimizes the error of sum squares

$$S(\beta) = \sum_{t=1}^{n} [Y_t - F(X_t, \beta)]^2 \quad (5)$$

It should noted that, nonlinear least squares situation may have several relative minima in addition to the absolute minimum $\hat{\beta}$. We can estimate the value of $\hat{\beta}$ using linear approximation. Let say we noting that in a small neighborhood of β^*, the true value of β. By using linear Taylor expansion

$$f_i(\beta) \approx f_i(\beta^*) + \sum_{r=1}^{k} \frac{\partial f_i}{\partial \beta_r}\Big|_{\beta^*} (\beta_r - \beta_r^*) \quad (6)$$

equation (5) also can rewrite as $f(\beta) \approx f(\beta^*) + F(\beta-\beta^*)$. Hence $S(\beta) = ||z - F.\theta||^2$, where $z = Y_t - f(\beta)$ and $\theta = \beta-\beta^*$. From the properties of the linear model is minimized when $\theta$is given by

$$= (F.^T F.)^{-1} F.^T z, \text{where } F. = \frac{\partial f(\beta}{\partial \beta_r}$$

When n is large enough $\hat{\beta}$, is almost certain to be within a small neighborhood of β^*. Hence $\hat{\beta}-\beta^* \approx \theta$ and $\hat{\beta}-\beta^* \approx (F.^T F.)^{-1} F.^T z$.

In the nonlinear situation, both X and F(β) are functions of β and a closed-form solution generally does not exist. Thus nlin procedure uses an iterative process. A starting value for β chosen and continually improved until the error of sum-squares, $e^T e$ (SSE) is minimized. The iterative techniques involving the matrix X evaluated for the current values of β and e = Y - F(β), the residual evaluated for the current values of β. The iterative process begins at some point (starting/initial value) β_0. Then X and Y are used to compute a δ such that $SSE(\beta_0 + k\delta) < SSE(\beta_0)$.

For most nonlinear models they cannot be solved analytically, so that iteration method is necessary. Let $\beta^{(k)}$ is an approximation to the least squares estimate $\hat{\beta}$ of a nonlinear model. For β close to $\beta^{(k)}$, we use linear Taylor expansion $f(\beta) \approx f(\beta^{(k)} + F_.^{(k)}(\beta-\beta^{(k)})$. If r(β) is residual vector, then $r(\beta) = Y-f(\beta) \approx r(\beta^{(k)}-F^{(k)}(\beta-\beta^{(k)})$. By substituting $S(\beta) = r^T(\beta)r(\beta)$ leads to $S(\beta) \approx r^T(\beta^{(k)})r(\beta^{(k)})-2r^T(\beta^{(k)})F^{(k)}(\beta-\beta^{(k)}) + (\beta-\beta^{(k)})^T F^{(k)T} F^{(k)}(\beta-\beta^{(k)})$. The right side is minimized with respect to β when $\beta-\beta^{(k)}=\{(F^{(k)T}F^k)^{-1}F^{(k)T}r(\beta^{(k)})\}=\delta^{(k)}$. If $\beta^{(k)}$ is a starting value, so the next approximation should be $\beta^{(k+1)} = \beta^{(k)} + \delta^{(k)}$. This procedure provides an iterative scheme for obtaining $\hat{\beta}$. There are four difference methods to determine how the value of δ computed to change the vector of parameters.

- Gradient, $\delta=X^T e$
- Gauss-Newton, $\delta=(X^T X)^{-1}X^T e$
- Newton, $\delta=(G^{-1})X^T e$, G is Moore-Panrose matrix
- Marquardt, $\delta= (X^T X + \lambda diag(X^T X)^{-1}X^T e$.

By using the nonlinear least squares Ralston and Jennrich (1978) and Seber and Wild (1989) recommended relative change convergence criteria based on changes in S(β) (equation 5) and the parameters in going from the ith to the (i+1)th iteration. That is, if the relative change in the sum of squares at the ith iteration,

$$\frac{(S(\beta^{(i)}) - S(\beta^{(i+1)}))}{S(\beta^{(i)})} \qquad (7)$$

falls in the interval of 0 to ξ_s, where ξ_s is a pre-selected tolerance level such as 10^{-4}, then the reduction in the sum of squares is considered insufficient to continuing and so the computational may be halted. This is usually accompanied by a parameter relative change criterion such as

$$\frac{|\beta_j^{(i+1)} - \beta_j^{(i)}|}{|\beta_j^{(i)}|} < \xi_p, j = 1, 2, \ldots . \qquad (8)$$

so that when every relative parameter change at the ith iteration is less than ξ_p, the parameter increments are too small to warrant continuing and the program terminates.

Gallant (1987) and Seber and Wild (1989), showed that the confidence interval of is given by $\beta_i \pm t_{\alpha/2} \sqrt{S^2 \hat{c}_i}$ where $\hat{c}_i$ is the i-th diagonal element of $\hat{C}c_{ii} = (F^T(\hat{\beta}) F(\hat{\beta}))^{-1}$.

Optimization algorithm and starting value

Gradient method

The steepest descent method is based on the gradient of $\epsilon'\epsilon$:

$$0.5\, \partial\epsilon^T\epsilon/\partial\beta = -XY + XF(\beta) = -X^Te.$$

The quantity $-X^Te$ is the gradient along which $\epsilon^T\epsilon$ increase. Thus $\delta = X^Te$ is the direction of steepest descent. Let $\beta_{i+1} = \beta_i + k\delta$, where the scalar of k chosen such that $SSE(\beta_i + k\delta) < SSE(\beta_i)$. The steepest descent may converge very slowly and is therefore not generally recommended. It is sometimes useful when the initial values are poor.

Newton method

This method uses the second derivatives and solve the equation $\delta=(G^{-1})X^Te$ when G is a matrix of $(X^TX) + \sum H(\beta)e$ and $H(\beta)$ is the hessian of e and $H_{ij}(\beta) = \partial^2 e/\partial\beta_i\, \partial\beta_j$

Gauss newton method

This technique basically uses the Taylor series

$$F(\beta) = F(\beta_0) + X(\beta - \beta_0) + \text{.....}$$

where $X = \partial F/\partial\beta$ is evaluated at $\beta = \beta_0$

By substituting the first two terms of this series into the normal equations

$$X^TF(\beta) = X^T Y$$
$$X^T(F(\beta_0) + X(\beta - \beta_0) = X^TY$$
$$X^TF(\beta_0)^TX(\beta - \beta_0)^TY$$
$$(X^TX)(\beta - \beta_0) = X^TY - X^TF(\beta_0)$$
$$(X^TX)\delta = X^Te \text{ and therefore } \delta = (X^TX)^{-1}X^Te$$

Marquardt method

The Marquardt method is a compromise between Gauss-Newton and steepest descent. The Marquardt updating formula is as follows:

$$\delta = (X^TX + \lambda \text{diag}(X^TX)^{-1}X^Te$$

As $\lambda \to 0$, the direction approaches Gauss-Newton. As $\lambda \to \infty$, the direction approaches steepest descent. Marquardt's studies indicate that the average angle between Gauss-Newton and steepest descent direction is about 90° (SAS, 1989). A choice of λ between 0 and ∞, produces a compromise direction.

By default (SAS), Proc nlin choose $\lambda = 10^{-3}$ to start and compute a δ. If $SSE(\beta_0 + \delta) < SSE(\beta_0)$, then $\lambda = \lambda/10$ for the next iteration. For each time $SSE(\beta_0 + \delta) > SSE(\beta_0)$, then $\lambda = \lambda^*10$. Marquardt's method is equivalent to performing a series of ridge regressions and is useful when the parameter estimates are highly correlated or the objective function is not well approximated by a quadratic.

Secant Method (DUD)

The multivariate secant method (DUD) is like Gauss-Newton, except that the derivatives are estimated from the history of iteration rather than supplied analytically. The method is also called the method of false position or the DUD method for Doesn't Use Derivatives (Ralston and Jennrich 1978). If only one parameter is being estimated, the derivative for iteration i + 1 can be estimated from the previous two iterations:

$$Der_{(i+1)} = \frac{\hat{Y}_i - Y_{i-1}}{b_i - b_{i-1}}$$

When k parameters are to be estimated, the method uses the last k+i iteration to estimate the derivatives.

Starting value

The starting value $\beta^{(1)}$, which is initial guess at the minimum $\hat{\beta}$, can sometimes be suggested by prior information. Sometimes there will be a starting value that tends to works well for a class of problems. Fisher's scoring algorithm for generalized linear models as an iterative re-weighted least square method suggests a uniform starting mechanism for the whole class of models (McCullagh and Nelder, 1983). However, it is very difficult to say anything about producing good starting values in general. Methods that are sometimes suggested include a grid search or a random search over a defined rectangular region of the parameter space. If nor sensible bounds can be suggested for a parameter β_r, a transformed parameter can be used. e.g.

$$\varphi = \frac{e^{\beta}}{1+e^{\beta}} \text{ and } \varphi = \arctan(\beta) \text{ both satisfy } 0 < \varphi < 1.$$

Draper and Smith (1998) and Ratkowsky (1983) gives the detail discussion on starting value for nonlinear model.

Numerical example

Data and scope of study

Data, which is going to be use, are experiment data which have be conducted by Malaysian Agriculture Research and Development Institute (MARDI). The tobacco plants are being planted in various plots and under a strict surveillance. The surrounding environments are being monitored until it can be assumed that growth can be standardized. When the saplings are week old, they are being uprooted and separate between the leaves, stems and roots and later being measured. The data then will be recorded. The data will be gathered every seven days. For each sample, three stems will be used and each part will be measured their average. About 36 samples will be pick at random. The experiments will last until week thirteen (sample #1) and week twelve (sample #2). Normally, the tobacco will be harvested at the end of week twelve or thirteen.

Analyses and result

Example SAS program for Gompertz growth model;

```
Data tobacco;
Input Y T @;
Cards;
.....data here.....
;
Proc nlin data=tobacco plot method=marquardt/;
    Parms A=A1
          B=B1
          C=C1;
    L=exp(B-CT);
    Model Y=A*exp(-L);
    Der.A=exp(-L);
    Der. B=A*L*exp(-L);
    Der.C=B*T*L*exp(-L)
Output out p=Yhat R=Yresid;
Run;
```

This study found that the growth of tobacco leaves, stem and root are in S-shaped, also called sigmiodal function (refer Fig.1 a, b and c). So Gompertz curve is expected to be a suitable tool in fulfilling the nature of agricultural plant growth.

In applying Gompertz curve to be fitted in the growth of tobacco leaves, stem and root, the starting value must firstly be determined. Since there is no such accurate way to fix the starting value in non-linear modeling (in this case Gompertz curve), 'trial and error' method is considered as a better solution. In this study, several sets of different data are being used. Therefore, several sets of starting values are also needed either the data criteria or data nature. Table 2 shows that the starting value for α, κ and γ for the growth of tobacco leaves sample #1 and sample #2 are 1.10, 0.10 and 25.00. Whereas, the starting value chosen for stem tobacco growth are 0.550, 0.10 and 25.00 in order to estimate the α, κ and γ.

The starting values of α, κ and γ for sample #1 and sample #2 are 0.50, 0.10 and 27.50, The iteration process showed that the starting values chosen are suitable or applicable since the number of iteration needed to achieve global convergence are small between 16 to 18 (refer Table 1 and Table 2). It means that the starting values are approaching to the actual estimated value. The sum square error values decline rapidly fast before reaches the consistent level during the last iteration.

The parameter estimated values could be determined during the last iteration when the global convergences are met. The estimated values of non-linear Gompertz curve for leaves growth sample #1 is α = 1.5311, κ = 0.1112 and γ = 34.4373. The parameter estimated for non-linear Gompertz curve of sample #2 is α = 1.6192, κ = 0.1076 and γ = 34.1598. Estimated parameter for stem growth sample #1 is α = 08292, κ = 0.1083 and γ = 38.9422 and for sample #2 is α = 0.7816, κ = 0.1211 and γ = 43.1989. Other estimated values, such as asymptotic standard error and the 95% confidence interval for each estimated parameters are shown in Table 3 and Table 4).

The Gompertz curve goodness of fit is considered after starting values and estimated parameters are determined. The goodness of fit can be measured by the value of sum square

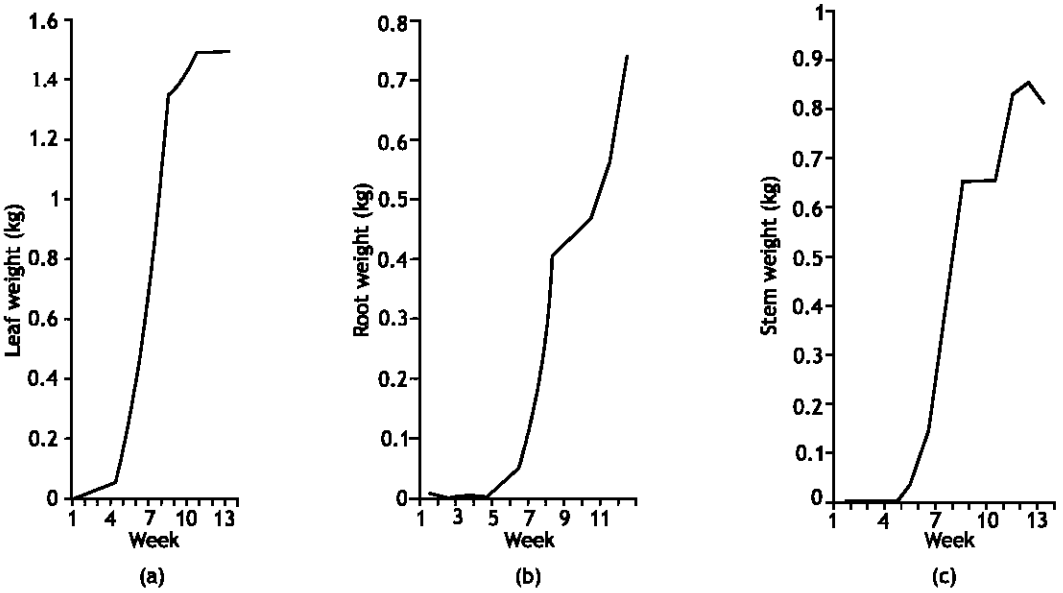

Fig. 1: Plot original data for leaves (a), stem (b) and root (c) for sample #1

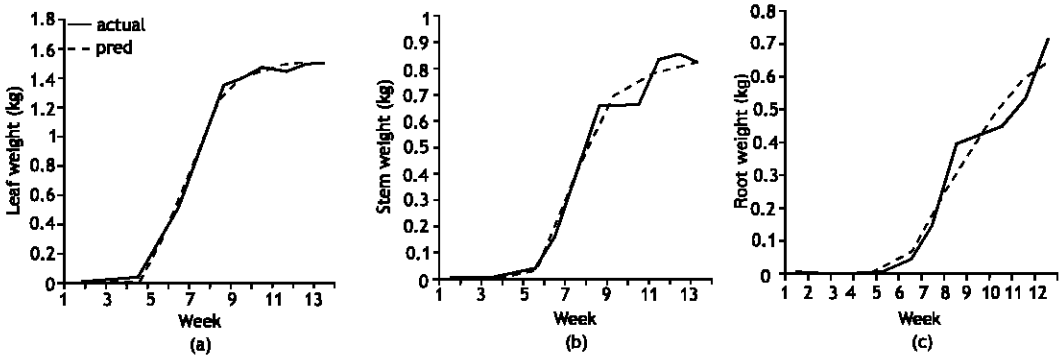

Fig. 2: Predicted and observed value for sample # 1. (a) leaves, (b) stem and (c) root

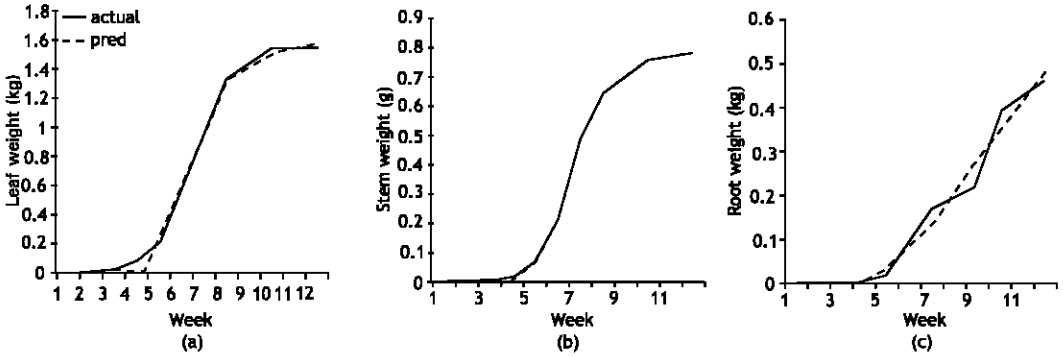

Fig. 3: Predicted and observed value for sample # 2, (a) leaves, (b) stem and (c) root

Table 1: Nonlinear least squares iterative phase for sample #1

		Parameter estimate			
	# iteration	α	κ	γ	Sum square error
Leaves	0	1.1000	0.1000	25.0000	1.0613
	1	1.5653	0.0469	37.2897	0.6419
	2	1.6081	0.0568	38.2828	0.3918
	3	1.4122	0.1119	30.6893	0.1326
	4	1.5617	0.1029	34.7583	0.0265
	5	1.5601	0.1024	34.4389	0.0262
	6	1.5324	0.1088	34.1983	0.0239
	7	1.5300	0.1100	34.3712	0.0235
	8	1.5308	0.1106	34.3635	0.0234
	9	1.5306	0.1111	34.4084	0.0234
	10	1.5318	0.1109	34.4207	0.0234
	11	1.5312	0.1111	34.4321	0.0234
	12	1.5312	0.1112	34.4307	0.0234
	13	1.5311	0.1112	34.4357	0.0234
	14	1.5312	0.1112	34.4361	0.0234
	15	1.5311	0.1112	34.4370	0.0234
	16	1.5311	0.1112	34.4369	0.0234
	17	1.5311	0.1113	34.4372	0.0234
	18	1.5311	0.1113	34.4374	0.0234
Stem	0	0.5500	0.1000	25.0000	0.4432
	1	0.6976	0.0397	33.7764	0.4086
	2	0.6629	0.0560	31.7247	0.2977
	3	0.6637	0.0558	31.7597	0.2977
	4	0.6597	0.0569	31.5034	0.2970
	5	0.7891	0.0626	37.2254	0.1055
	6	0.8792	0.0813	40.9291	0.0338
	7	0.8406	0.0937	38.5353	0.0249
	8	0.8371	0.0966	38.3165	0.0247
	9	0.8451	0.1082	39.2027	0.0240
	10	0.8418	0.1094	39.1267	0.0238
	11	0.8302	0.1083	39.0482	0.0231
	12	0.8295	0.1079	38.9026	0.0230
	13	0.8291	0.1083	38.9411	0.0230
	14	0.8292	0.1083	38.9422	0.0230
	15	0.8292	0.1083	38.9422	0.0230
Root	0	0.5000	0.1000	27.5000	0.2275
	1	0.5156	0.0813	28.9812	0.1856
	2	0.6242	0.0472	38.8828	0.0989
	3	0.6530	0.0681	40.0691	0.0540
	4	0.7044	0.0567	44.5009	0.0401
	5	0.7113	0.0557	45.0736	0.0399
	6	0.7776	0.0666	49.6236	0.0313
	7	0.7755	0.0644	49.5326	0.0311
	8	0.7601	0.0676	47.4970	0.0265

Table 1: Continue

		Parameter estimate			
	# iteration	α	κ	γ	Sum square error
	9	0.7626	0.0671	47.6129	0.0265
	10	0.7886	0.0598	47.8356	0.0263
	11	0.7919	0.0605	47.9556	0.0261
	12	0.8010	0.0602	48.2036	0.0260
	13	0.8031	0.0601	48.2746	0.0261
	14	0.8038	0.0600	48.3582	0.0261
	15	0.8051	0.0598	48.3959	0.0260
	16	0.8094	0.0592	48.4970	0.0260
	17	0.8106	0.0591	48.5290	0.0260
	18	0.8106	0.0591	48.5298	0.0260

Table 2: Nonlinear least squares iterative phase for sample #2

		Parameter estimate			
	# iteration	α	κ	γ	Sum square error
Leaves	0	1.1000	0.1000	25.0000	1.1630
	1	1.6607	0.0392	37.5438	0.9509
	2	1.9303	0.0687	41.1549	0.2931
	3	1.4659	0.1200	28.5944	0.2382
	4	1.5859	0.1173	32.0347	0.0582
	5	1.6651	0.1021	35.0271	0.0221
	6	1.6442	0.1026	34.6808	0.0179
	7	1.6299	0.1046	34.1282	0.0157
	8	1.6301	0.1048	34.1395	0.0157
	9	1.6193	0.1071	34.1371	0.0154
	10	1.6189	0.1074	34.1345	0.0154
	11	1.6189	0.1074	34.1365	0.0154
	12	1.6195	0.1075	34.1588	0.0154
	13	1.6196	0.1075	34.1587	0.0154
	14	1.6192	0.1076	34.1586	0.0154
	15	1.6192	0.1076	34.1581	0.0154
	16	1.6192	0.1076	34.1592	0.0154
	17	1.6192	0.1076	34.1599	0.0154
	18	1.6192	0.1076	34.1599	0.0154
Stem	0	0.5500	0.1000	25.0000	0.4095
	1	0.5807	0.0782	27.2896	0.3214
	2	0.6134	0.0644	29.4823	0.2708
	3	0.6054	0.0744	29.0435	0.2702
	4	0.6284	0.0628	30.8324	0.2436
	5	0.6597	0.0555	32.7525	0.2209
	6	0.9213	0.0603	52.7930	0.1567
	7	0.9266	0.0711	52.6913	0.1438
	8	0.7650	0.1323	42.5678	0.0033
	9	0.7654	0.1334	42.7512	0.0030
	10	07778	0.1295	43.2636	0.0022

Table 2: Continue

		Parameter estimate			
	# iteration	α	κ	γ	Sum square error
	11	0.7857	0.1241	43.3906	0.0022
	12	0.7822	0.1227	43.2452	0.0020
	13	0.7804	0.1217	43.2367	0.0020
	14	0.7809	0.1218	43.2128	0.0020
	15	0.7819	0.1210	43.2153	0.0019
	16	0.7818	0.1211	43.2043	0.0019
	17	0.7817	0.1211	43.1989	0.0019
	18	0.7817	0.1211	43.1989	0.0019
Root	0	0.5000	0.1000	27.5000	0.4324
	1	0.4449	0.0327	31.8375	0.1149
	2	0.5769	0.0183	49.4612	0.1098
	3	0.7978	0.0154	72.2860	0.0864
	4	0.7929	0.0155	71.7891	0.0865
	5	1.0273	0.0228	75.5878	0.0173
	6	1.1422	0.0264	76.4153	0.0092
	7	0.9319	0.0318	68.4466	0.0083
	8	0.7966	0.0356	63.5151	0.0078
	9	0.7745	0.0361	62.8872	0.0077
	10	0.6968	0.0414	59.8539	0.0075
	11	0.7056	0.0411	60.2855	0.0074
	12	0.7051	0.0412	60.2526	0.0074
	13	0.7051	0.0412	60.2542	0.0074
	14	0.7047	0.0412	60.2375	0.0074
	15	0.7048	0.0412	60.2407	0.0074
	16	0.7048	0.0412	60.2407	0.0074

Table 3: Parameter estimate, asymptotic standard error and the 95% confidence interval Gompertz model for sample # 1

			95% Confidence interval	
Parameter	Estimate	Asymptotic standard error	Lower limit	Upper limit
Leaves				
α	1.5.311	0.02876	1.4671	1.5953
κ	0.1112	0.0099	0.0891	0.1334
γ	34.4373	0.5654	34.1775	35.6972
Stem				
α	0.8292	0.0326	0.7567	0.9018
κ	0.1083	0.0018	0.0672	0.1494
γ	38.9422	1.0738	36.5495	41.3349
Root				
α	0.8107	0.1565	0.4564	1.1649
κ	0.0591	0.0197	0.0144	0.1039
γ	48.5297	4.3130	38.7729	58.2865

Table 4: Parameter estimate, asymptotic standard error and 95% confidence interval Gompertz model for sample # 2

			95% Confidence interval	
Parameter	Estimate	Asymptotic standard error	Lower limit	Upper limit
Leaves				
α	1.6192	0.0301	1.5511	1.6873
κ	0.1076	0.0083	0.0888	0.1264
γ	34.1598	0.4819	33.0697	35.2500
Stem				
α	0.7817	0.0104	0.7581	0.8052
κ	0.1211	0.0071	0.1051	0.1371
γ	43.1988	0.3326	42.4464	43.9514
Root				
α	0.7047	0.1747	0.3096	1.0999
κ	0.0412	0.0119	0.0142	0.0682
γ	60.2401	7.1013	44.1764	76.3051

Table 5: Non-linear least squares summary statistics for leaves, stem and root (sample #1)

		Leaves		Stem		Root	
Source	df	SS	MS	SS	MS	SS	MS
Regression	3	13.9636	4.6545	3.6269	1.2089	1.4269	47.5634
Residual	10	0.0234	0.0023	0.0230	0.002304	0.0260	0.2604
Uncorrected total	13	13.9870		3.6499		1.4529	
Corrected total	12	5.3794		1.6136		0.7803	

Table 6: Non-linear least squares summary statistics for leaves, stem and root (sample #2)

		Leaves		Stem		Root	
Source	df	SS	MS	SS	MS	SS	MS
Regression	3	12.9819	4.3273	2.9579	0.9859	0.7060	0.2353
Residual	9	0.0154	0.0017	0.0019	0.000222	0.007449	0.000827
Uncorrected total	12	12.9973		29.599		0.7135	
Corrected total	11	5.3058		1.3049		0.3574	

Table 7: Fitting Gompertz curve and SSE value for sample #1 and sample #2

	Sampel #1	SSE	Sampel #2	SSE
Leaves	$1.5311e^{-e^{0.1112(T-34.4373)}}$	0.0234	$0.1692e^{-e^{0.1076(T-34.1598)}}$	0.0154
Stem pokok	$0.8292e^{-e^{0.1083(T-38.9422)}}$	0.0230	$0.7816e^{-e^{0.1211(T-43.1989)}}$	0.0019
Root	$0.8107e^{-e^{0.0591(T-48.5297)}}$	0.0260	$0.7047e^{-e^{0.0412(T-60.2407)}}$	0.0074

Table 8: Asymptotic correlation matrix for sample #1and #2

Variable	Sample #1	Sample #2
Leaves	$(\alpha, \kappa) = -0.6290$	$(\alpha, \kappa) = -0.6949$
	$(\alpha, \gamma) = 0.3531$	$(\alpha, \gamma) = 0.4300$
	$(\kappa, \gamma) = 0.0092$	$(\kappa, \gamma) = -0.0933$
Stem	$(\alpha, \kappa) = -0.6747$	$(\alpha, \kappa) = -0.6649$
	$(\alpha, \gamma) = 0.4018$	$(\alpha, \gamma) = 0.3932$
	$(\kappa, \gamma) = -0.0568$	$(\kappa, \gamma) = -0.0422$
Root	$(\alpha, \kappa) = -0.9212$	$(\alpha, \kappa) = -0.9565$
	$(\alpha, \gamma) = 0.9361$	$(\alpha, \gamma) = 0.9867$
	$(\kappa, \gamma) = -0.8296$	$(\kappa, \gamma) = -0.9414$

Table 9: The individual growth rates for sample #1 and sample #2

Variable	Sample #1	Sample #2
Leaves	$\frac{\partial L}{\partial t} = 0.112L[0.1850-\log(L)]$	$\frac{\partial L}{\partial t} = 0.1076L[0.2093-\log(L)]$
Stem	$\frac{\partial B}{\partial t} = 0.1083B[-0.0813-\log(B)]$	$\frac{\partial B}{\partial t} = 0.1211B[-0.1070-\log(B)]$
Root	$\frac{\partial R}{\partial t} = 0.0591R[-0.0912-\log(R)]$	$\frac{\partial B}{\partial t} = 0.0412R[-0.1519-\log(R)]$

Table 10: Relative growth rate of leaves, stem and root for sample #1 and sample #2

Variable	Sample #1 (week)	Sample #2 (week)
Leaves	0.0592; [6→7]	0.0564 [5→6]
	0.0573 [5→6]	0.0596 [5→6]
Stem	0.0375 [7→8]	0.0387 [6→7]
	0.0322 [6→7]	0.0329 [6→7]
Root	0.0358 [6→7]	0.0243 [9→10]
	0.0172 [6→7]	0.0106 [8→9]

error and the analyses of variance (anova). Table 1 and Table 2 show the values of sum square error are small enough and these values represent the Gompertz curve that has been well fitted to the growth of tobacco leaves, stem and root.

Table 5 and Table 6 (Table analysis of variance regression), have proved that the Gompertz curve is certainly suitable for the tobacco growth. The result of the variance residual test also supported the findings of this study. The normal probability plots for residual which was also performed, shows that the variance residual is normally distributed (not shown).

The next step is to study the confidence interval of each estimated parameter. Table 3 and Table 4, show that every estimated parameter, which is 95% confidence interval for α, κ and γ do not consist zero. This can be defined that dependent variable has significant relationship to independent variable, that is time (day).

The non-linear Gompertz curve equation, which has been fitted to tobacco leaves, stem and root growth are shown in Table 7. For example, Gompertz curve equation for the growth of the tobacco leaves sample #1 is $f(t) = 1.5311e^{-e^{0.1112(T-34.4373}}$ and $f(t) = 1.8292e^{-e^{0.1083(T-38.942}}$ is the equation for stem growth (another equation refer Table 7).

The plotted actual data and forecast data by Gompertz curve are shown in Fig. 2 (a, b and c) and Fig. 3 (a, b and c). The graph shows that the tobacco growth data can be suited very well by on-linear Gompertz curve.

Seber and Wild (1989), have proved that the growth rate of organism was given in equation (1). The growth equation for each variable can be derived by replacing the estimated values (Table 8). It is clearly shows that the growth rate for sample #1 and sample #2 are not significantly different.

Md. Yunus (1999) was defined relative growth rate for growth of cocoa leaves (RGR) as

$$RGR=\frac{f_{i+1}-f_i}{t_{i+1}-t_i}$$

and the result are shown at Table 10 (the estimated RGR by italic) for each variable of sample #1 and sample #2. Table 10 also shows the maximum changes was occurring and found that the model was estimate early than the actual maximum changes.

Non-linear Gompertz curve is found to be very well suited to forecast tobacco growth. This study also finds that the growth rate of tobacco leaves and stem are more rapid than the growth rate of the root. For the future research on tobacco growth, we suggest to consider another factors that also influenced the growth rate of tobacco tree, such as nutrient level, temperature, rainfall, environment etc.

References

Draper, N.R. and H. Smith, 1998. Applied regression analysis. 3rd. Ed. Wiley, NY.

Gallant, A.R., 1987. Nonlinear Statistical Models. Wiley: New York.

Jennrich, R.I., 1995. An introduction to computational statistics: Regression analysis. Prentice Hall Inc. New Jersey.

Md. Yunus Jaafar, 1999. Pensuaian model tak linear Gompertz terhadap pertumbesaran pokok koko. Matematika, Jilid 15, Bil., 1: 1-20.

McCullagh, P. and J.A. Nelder, 1983. Generalized Linear Models. Chapman and Hall: London.

Medawar, P.B., 1940. Growth, growth energy and ageing of the chicken's heart. Proc. Roy. Soc, London, 129: 332-355.

Ralston, M.L. and R.I. Jennrich, 1978. DUD, A Derivative-Free Algorithm for Nonlinear Least Squares. Technometrics, 20: 7-14.

Ratkowsky, D.A., 1983. Nonlinear Regression Modeling. Marcel Dekker: New York.

SAS/STAT User's Guide, Release 6.03 Ed. SAS Institute, Cary, NC.

Seber, G.A.F. and C.J. Wild, 1989, Nonlinear Regression. Wiley: New York.

Wright, D.E., 1986. A Note on the Construction of Highest Posterior Density Intervals. Appl. Stat., 35: 49-53.