The Use and Application of Weilbull Model in Growth Analysis

Dagogo J.
Dept. of Statistics
Federal Polytechnic Ukana,
Akwa Ibom, State,
Nigeria.
goodluck4all2003@yahoo.com

Nduka E. C. & Biu O. E.
Dept. of Mathematics and Statistics
University of Port Harcourt,
Rivers State, Nigeria.
emmanuelbiu@yahoo.com

Abstract
This study used Weilbull growth model to analyse the amount of transmitted voltage in the signal detection readings of voltage against time. The method of nonlinear least square estimation, using a modified version of the Levenberg-Marquardt was applied; by taking the derivative with respect to the parameters $(\beta_0, \beta_1, \beta_2)$. *The Gretl statistical software was used, then the data set was added and the initial values for the parameters based on the mathematical properties of the Weibull growth derived and the second-order polynomial (quadratic) model obtained. The results of the sum of squared residual, R-squared and Log-likelihood showed that the identified Weibull growth model is adequate and can be used for forecasting.*

1. Introduction

Weibull model is a nonlinear regression model that can be used to determine the relationships between the dependent and independent where their parameters are not linear. In statistics, nonlinear regression is a form of regression analysis in which observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. The data are fitted by a method of successive approximations. Also nonlinear regression is a regression in which the dependent or criterion variables are modeled as a non-linear function of model parameters and one or more independent variables (i.e. nonlinear in the parameters).

There are several common models, such as Asymptotic Regression/Growth Model, Gompertz, Weibull, Hill, Richards, Logistic, S-shaped Curves etc. These models (or curves) referred to as Sigmoidal Growth Models (Sigmoidal curves) which arise in various application including bioassay, signal detection theory, agriculture, engineering field, tree diameter, height distribution in forestry, fire size, high-cycle fatigue strength predication, seismological data analysis for earthquakes and economics [Tjorve and Tjorve, (2017); Panik, (2014); Seber and Wild, (1989); Meade and Islam, (1995); Bethea, *et al.*, (1985) etc].

Various types of growth data often conform to sigmoidal curves. The seminar work only considered Weilbull model in growth analysis and its application. The data used in this work are results from an experiment in which several time (in minutes) were used to measure the amount of transmitted voltage to determine the relationship between the signal detection readings of voltage against time.

In some situation, it is possible to transform a nonlinear regression function using appropriate transformation of the response variable Y_i, the predictor variable, the parameters, or any combination of these, such that the transformed function is linear in the unknown parameters. If the transformed variables satisfy assumption for simple or multiple linear regressions, then the result can be applied to the transformed problem. Using the results for the transformed, we can often obtain results for the original problem.

The Weibull model with four parameters is expressed as

$$Y_i = \beta_0 x^{\beta_1 - 1} e^{-\beta_2 X_i^{\beta_3}} \tag{3.1}$$

where

e represents Euler number (e = 2.71828)

X represents time

$\beta_0, \beta_1, \beta_2, \beta_3$ are the parameters

β_0 represents upper asymptote when time approaches positive infinite (or maximum growth response or scale parameter)

β_1 represents the shape parameter related to initial time

β_2 represents growth range (or intrinsic growth range)

β_3 represents growth rate (or shape parameter)

Y_i is the ith observation at time

The most Weibull model used is the three parameters model that shifts the curve horizontally without changing its shape [Tjorve and Tjorve, (2017) and Raji *et al.* (2014)]. A simple rewrite of the four parameter Weibull model can be expressed as three parameters, that is

$$Y_i = \beta_0 \, e^{-\beta_1 X_i^{\beta_2}} \tag{3.2}$$

where

e represents Euler number (e = 2.71828)

X represents time

$\beta_0, \beta_1, \beta_2$ are the parameters

β_0 represents upper asymptote when time approaches positive infinite

β_1 represents growth range

β_2 represents growth rate

Y_i is the ith observation at time

By applying the nonlinear least squares method (or performing nonlinear least square (NLS) estimation using a modified version of the Levenberg-Marquardt algorithm.

Let $\beta = \left(\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}\right)$ be the initial parameters. This can be obtained by fitting the log-linear model [or by Logarithm transformation in Equation (3.2)]. We have

$$\ln(Y_i) = \ln(\beta_0) - \beta_1 X_i^{\beta_2} \tag{3.3}$$

The NLS estimation using a modified version of the Levenberg-Marquardt, we take the derivative with respect to the parameters $(\beta_0, \beta_1, \beta_2)$. Then, arbitrarily value for $\left(\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}\right)$ as the initial guess vector for iteration process. The Gretl statistical software was used, and then adds the data set and the initial values for the parameters basis of the mathematical properties of the Weibull growth (or curve).

To illustrate the mathematical properties of the Weibull model (or curve), the table below has been prepared.

Table 1: Weibull model property

PROPERTY	**Weibull**
Equation	$Y_i = \beta_0 \, e^{-\beta_1 X_i^{\beta 2}}$
Number of parameters	3
Asymptotes	$\begin{cases} Y = 0, \; if \;\; \beta_0 = 0 \\ Y = \beta_0 \, , \; if \;\; \beta_1 = 0 \; or \;\; X = 0 \end{cases}$
Inflection	$\begin{cases} X = \left(\frac{1}{\beta_1}\right)^{\frac{1}{\beta_2}} \\ Y = \frac{\beta_0}{e} \end{cases}$
Straight line form of equation	$\ln \ln\left(\frac{\beta_0}{Y}\right) = \ln(\beta_1) + \beta_2 \ln(X)$
Symmetry	Asymmetrical
Growth rate	$\frac{dY}{dx} = -\beta_1 \beta_2 X^{\beta_2 - 1} Y$
Maximum growth rate	$\frac{-\beta_0 \, \beta_2 \left(\frac{1}{\beta_1}\right)^{-\frac{1}{\beta_2}}}{e}$
Relative growth rate as function of time	$\frac{1}{Y}\frac{dY}{dx} = -\beta_1 \beta_2 X^{\beta_2 - 1}$
Relative growth rate as function of size	$\frac{1}{Y}\frac{dY}{dx} = -\beta_1 \beta_2 \left[\frac{1}{\beta_1} \ln\left(\frac{\beta_0}{Y}\right)\right]^{\beta_2 - 1}$

The mathematical properties of the Weibull growth derived in Table 1 are as follow

Asymptotes: $if \;\; \beta_1 = 0 \;\; , \; Y = \beta_0 \, e^{-0 X^{\beta_2}} = \beta_0$

Inflection: if $X = \left(\frac{1}{\beta_1}\right)^{\frac{1}{\beta_2}}$, $Y = \beta_0 e^{-\beta_1 \left(\left(\frac{1}{\beta_1}\right)^{\frac{1}{\beta_2}}\right)^{\beta_2}} = \beta_0 e^{-1} = \frac{\beta_0}{e}$

To obtain the linear form of the equation:

$$Y = \beta_0 e^{-\beta_1 X^{\beta_2}}$$

Take ln

$$\ln(Y) = \ln(\beta_0) - \beta_1 X^{\beta_2}$$

$$\ln(\beta_0) - \ln(Y) = \beta_1 X^{\beta_2}$$

$$\ln(\frac{\beta_0}{Y}) = \beta_1 X^{\beta_2}$$

Take ln

$$\ln\ln(\frac{\beta_0}{Y}) = \ln(\beta_1) + \beta_2 \ln X$$

Growth rate: differential Equation (3.2) with respect to X

$$Y = \beta_0 e^{-p} \ \ and \ \ p = \beta_1 X^{\beta_2}$$

$$\frac{dY}{dp} = -\beta_0 e^{-p} \ \ and \ \ \frac{dp}{dX} = \beta_1 \beta_2 X^{\beta_2 - 1}$$

$$\frac{dY}{dX} = -\beta_0 \beta_1 \beta_2 X^{\beta_2 - 1} e^{-\beta_1 X^{\beta_2}} = -\beta_1 \beta_2 X^{\beta_2 - 1} Y$$

Maximum growth rate: substitute $X = \left(\frac{1}{\beta_1}\right)^{\frac{1}{\beta_2}}$ into $\frac{dY}{dX}$

$$\frac{dY}{dX} = -\beta_0 \beta_1 \beta_2 X^{\beta_2 - 1} e^{-\beta_1 X^{\beta_2}} = -\beta_0 \beta_1 \beta_2 \left(\left(\frac{1}{\beta_1}\right)^{\frac{1}{\beta_2}}\right)^{\beta_2 - 1} e^{-\beta_1\left(\left(\frac{1}{\beta_1}\right)^{\frac{1}{\beta_2}}\right)^{\beta_2}} = \frac{-\beta_0 \beta_2 \left(\frac{1}{\beta_1}\right)^{-\frac{1}{\beta_2}}}{e}$$

Relative growth rate as function of size: substitute $X^{\beta_2} = \frac{1}{\beta_1} \ln\left(\frac{\beta_0}{Y}\right)$ into $\frac{dY}{dX}$, then

$$\frac{dY}{dX} = -\beta_0 \beta_2 \left(\frac{1}{\beta_1} \ln\left(\frac{\beta_0}{Y}\right)\right)^{\beta_2 - 1}$$

2. Aim and objectives

The aim of this seminar work is to show the use and application of Weilbull model in growth analysis. The objectives are

1. To derive the Weibull model properties
2. To illustrate the nonlinear least square (NLS) estimation using a modified version of the Levenberg-Marquardt algorithm.
3. To estimate the parameters of the Weibull model, using real data and Gretl statistical software.

3. Method

Nonlinear regression modelling is similar to linear regression modelling in that both seek to graphically track a particular response from a set of variables. Nonlinear models are more complicated than linear models to develop because the function is created through a series of approximations (iterations) that may stem from trial-and-error. Mathematicians use several established methods, such as the Gauss-Newton method and the Levenberg-Marquardt method. The seminar used the modified version of the Levenberg-Marquardt method. That is

1) Obtain partial derivative of the model with respect to the three parameters $(\beta_0, \beta_1, \beta_2)$.
2) Then, arbitrarily value for $\left(\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}\right)$ as the initial guess values for iteration process.
3) Input the data and initial guess values on the program develop in the Gretl statistical software. Then, run the iteration to obtain the results.

In Equation (3.3), let $\ln(Y_i) = Z_i$, $\ln(\beta_0) = \hat{\beta}_0$, $-\beta_1 = \hat{\beta}_1$ *and* $\beta_2 = \hat{\beta}_2$

$$Z_i = \hat{\beta}_0 + \hat{\beta}_1 X_i^{\hat{\beta}_2} \tag{3.4}$$

By take partial derivative of Equation (3.4) with respect to $\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right)$, we have

$$\frac{\partial Z_i}{\partial \hat{\beta}_0} = 1 \tag{3.5}$$

$$\frac{\partial Z_i}{\partial \hat{\beta}_1} = X_i^{\hat{\beta}_2} \tag{3.6}$$

$$\frac{\partial Z_i}{\partial \hat{\beta}_2} = \beta_1 \times \left(X_i^{\hat{\beta}_2}\right) \times \ln\left(X_i\right) \tag{3.7}$$

The initial guess values $\left(\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}\right)$ were estimated by fitting a second-order polynomial (quadratic model), using Minitab 17 software.

$$Z_i = \hat{\beta}_0^{(0)} + \hat{\beta}_1^{(0)} X_i + \hat{\beta}_2^{(0)} X_i^{\,2} \tag{3.8}$$

4 Analysis

The data used in this study is an experiment used to determine the amount of transmitted voltage against time collected from the Department of Electrical/Electronic Engineering, University of Port-Harcourt (Table 2) to analyze the results of Weibull growth model building and parameter estimates.

- Program

The program developed in the Gretl software using Equation (3.3), (3.4), (3.5), (3.7) and inputting the initial values fitted by second-order polynomial (quadratic model), using Minitab 17 software, see Appendix [$\left(\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}\right)$= (3.59784, -0.00374, 0.000011)].

```
    genr beta zero = 3.59784
    genr beta one = -0.00374
    genr beta two = 0.000011
nls Z = beta zero + beta one * X^ beta two
    deriv beta zero  = 1
    deriv beta one = X^ beta two
    deriv beta two = beta one * X^ beta two * log(X)
    end nls –vcv
```

- **Results output**

The result output in appendix is summary as follow in Table 3.

Table 3: Parameter estimates of the Weibull Model

Parameter	*Estimate (p-value)*	Remark
$\hat{\beta}_0$	3.8877 (0.000)***	Sig.
$\hat{\beta}_1$	-0.1929 (0.206)	Not Sig.
$\hat{\beta}_2$	0.2202 (0.025)**	Sig.

Footnote: Sig. at * 0.10, **0.05, ***0.01

To obtain the actual parameter values, we substitute into $\ln(\beta_0) = \hat{\beta}_0, -\beta_1 = \hat{\beta}_1$ *and* $\beta_2 = \hat{\beta}_2$. Then,

$$\beta_0 = e^{\hat{\beta}_0} = e^{3.8877} = 49.7982$$

$$\beta_1 = \hat{\beta}_1 = -(-0.1929) = 0.1929$$

$$\beta_2 = \hat{\beta}_2 = 0.2202$$

Hence, the fitted Weibull growth model is

$$\mathrm{Y}_i = 49.7982\, e^{-0.1929 \mathrm{X}_i^{0.2202}}$$

The adequacy of the model was also checked by the use of Mean error (-5.09×10^{-11}), R^2 (97.7%) and Log-likelihood (-15.9). It is obvious that the Weibull growth model is adequate and can be used for forecasting.

4. Conclusion

The study was able to show the use of Weibull growth model using real data set. The modified version of the Levenberg-Marquardt method for solving NLE regression model was used. A suitable Weibull growth model was determined using Mean squared error, R^2 and Log-likelihood show that the model is adequate and can be used for forecasting. The results of the sum of squared residual, R-squared and Log-likelihood showed that the identified Weibull growth model is adequate and can be used for forecasting.

5. References

Bethea, R. M.; Duran, B. S.; Boullion, T. L. (1985). *Statistical Methods for Engineers and Scientists. New York: Marcel Dekker.*

Meade, N.; Islam, T. (1995). *"Prediction Intervals for Growth Curve Forecasts". Journal of Forecasting. 14 (5): 413–430.*

Panik, M. (2014). Growth Curve Modeling. Hoboken, New Jersey: Wiley

Raji, A. O., Mbap, S. T. and Aliyu, J. (2014). *Comparison of different Models to Describe Growth of The Japanese Quail* (COTURNIX JAPONICA), Trakia Journal of Sciences, Vol 12, No. 2 pp. 182-188.

Schittkowski, K. (2002*). Data Fitting in Dynamical Systems. Boston: Kluwer.*

Seber, G. A. F., & Wild, C. J. (2003). *Nonlinear regression*. New York: John Wiley & Sons.

Seber, G. A. F.; Wild, C. J. (1989). *Nonlinear Regression. New York: John Wiley and Sons.*

Sheu, C. F. and Heathcote, A. (2001). *A nonlinear regression approach to estimating signal detection models for rating data*. Behavior Research Methods, Instruments and Computers, *33*(2), 108-114.

Tjorve, K. and Tjorve, E. (2017). *The use of Gompertz Models in Growth Analysis, and New Gompertz-Models Approach:* An Addition to The Unified-Richards Family. PLOS 1, 12 (6).

Verboon, P. (1993). *Robust nonlinear regression analysis.* British Journal of Mathematical and Statistical Psychology, *46*(1), 77-94.

Wang, J. (1995). *A symptotic normality of L-sub-1-estimators in nonlinear regression.* Journal of Multivariate Analysis, 54(2), 227-238.

Table 2: Voltage (volts) and Time (minute)

TIME(min)	VOLTAGE(volts)
5	36.36
10	35.24
15	34.72
20	33.90
25	33.34
30	32.77
35	32.26
40	31.77
45	32.26
50	31.30
55	30.58
60	30.22
70	29.51
80	28.90
90	28.50
100	28.15
110	27.84
120	27.56
130	27.34
140	27.18
150	27.11
160	27.05
170	27.02
180	27.01
190	27.00
200	27.00

Appendix
MINITAB 17 STATISTICAL SOFTWARE output

Regression Analysis: LnVOLTAGEvolts versus TIME (min), TT
The regression equation is
LnVOLTAGEvolts = 3.60 - 0.00374 TIME (min) + 0.000011 TT

Predictor	Coef	SE Coef	T	P
Constant	3.59784	0.00402	895.32	0.000
TIME (min)	-0.0037440	0.0001023	-36.61	0.000
TT	0.00001144	0.00000050	22.80	0.000

S = 0.00764348 R-Sq = 99.4% R-Sq(adj) = 99.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	0.23734	0.11867	2031.20	0.000
Residual Error	23	0.00134	0.00006		
Total	25	0.23868			

Source	DF	Seq SS
TIME (min)	1	0.20696
TT	1	0.03037

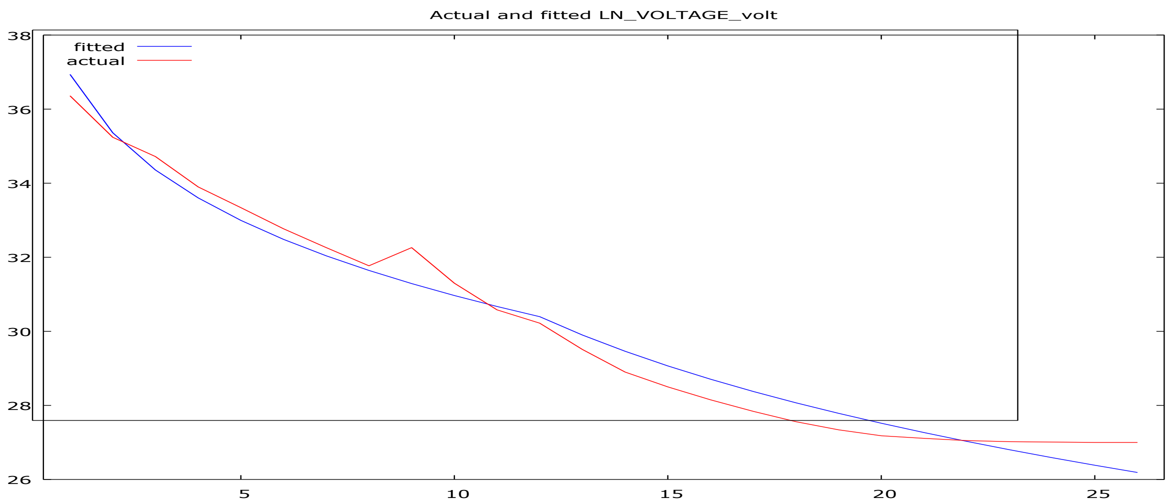
GRETL STATISTICAL SOFTWARE output
Using analytical derivatives
iteration 1: SSR = 1.2287176
iteration 2: SSR = 1.228717
iteration 3: SSR = 1.676486e+130
iteration 4: SSR = 1.2505135e+125
iteration 5: SSR = 6.5225556e+122
iteration 6: SSR = 4.3026778e+099
iteration 7: SSR = 25425.087
iteration 8: SSR = 1.1371526
iteration 9: SSR = 0.98902442
iteration 10: SSR = 0.72509055
iteration 11: SSR = 0.32471799
iteration 12: SSR = 0.035014239
iteration 13: SSR = 0.011171227
iteration 14: SSR = 0.046831
iteration 15: SSR = 0.0095771083
iteration 16: SSR = 0.0091399549
iteration 17: SSR = 0.0092402088
iteration 18: SSR = 0.0086792102
iteration 19: SSR = 0.0085309664
iteration 20: SSR = 0.0077077612
iteration 21: SSR = 0.0075183498
iteration 22: SSR = 0.0070715864
iteration 23: SSR = 0.0067487361

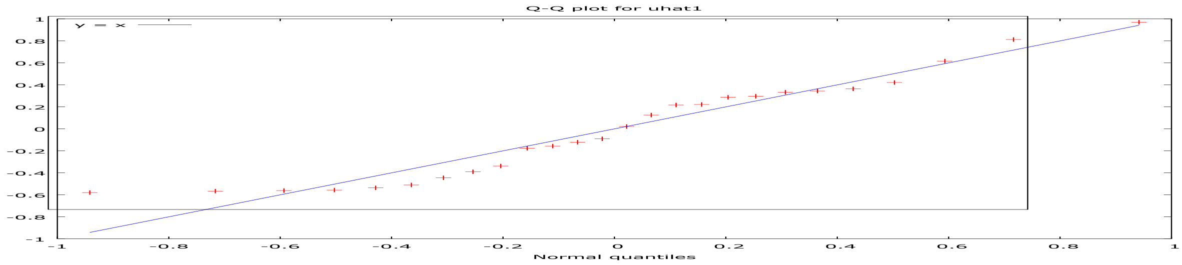
iteration 24: SSR = 0.0065043449
iteration 25: SSR = 0.0063162073
iteration 26: SSR = 0.0061698709
iteration 27: SSR = 0.0064301222
iteration 28: SSR = 0.0060704364
iteration 29: SSR = 0.0060095617
iteration 30: SSR = 0.0059217925
iteration 31: SSR = 0.0058570024
iteration 32: SSR = 0.0058087745
iteration 33: SSR = 0.0057738018
iteration 34: SSR = 0.0057495302
iteration 35: SSR = 0.0057339718
iteration 36: SSR = 0.0057255567
iteration 37: SSR = 0.0057193462
iteration 38: SSR = 0.0057161068
iteration 39: SSR = 0.0057161047
iteration 40: SSR = 0.0057161047
iteration 41: SSR = 0.0057161047
iteration 42: SSR = 0.0057161047
iteration 43: SSR = 0.0057161047
Tolerance = 1.81899e-012
Convergence achieved after 43 iterations
Model 3: NLS, using observations 1-26

Model 3: NLS, using observations 1-26
LN_VOLTAGE_volt = alpha + beta * TIME_min_^gamma
HAC standard errors, bandwidth 2 (Bartlett kernel)

	Estimate	*Std. Error*	*t-ratio*	*p-value*	
Alpha	3.88767	0.182997	21.2444	<0.00001	***
Beta	-0.192918	0.148381	-1.3002	0.20643	
Gamma	0.220167	0.091746	2.3997	0.02490	**

Mean dependent var	3.398963	S.D. dependent var	0.097710
Sum squared resid	0.005716	S.E. of regression	0.015765
R-squared	0.976051	Adjusted R-squared	0.973969
Log-likelihood	72.60093	Akaike criterion	-139.2019
Schwarz criterion	-135.4276	Hannan-Quinn	-138.1150
Rho	0.825989	Durbin-Watson	0.370008





Model estimation range: 1 - 26
Standard error of residuals = 0.474345

	LN_VOLTAGE_volt	fitted	Residual
1	36.3600	36.9395	-0.579494
2	35.2400	35.3617	-0.121703
3	34.7200	34.3564	0.363586
4	33.9000	33.6035	0.296452
5	33.3400	32.9958	0.344163
6	32.7700	32.4833	0.286661
7	32.2600	32.0385	0.221502
8	31.7700	31.6444	0.125592
9	32.2600	31.2899	0.970090
10	31.3000	30.9672	0.332772
11	30.5800	30.6707	-0.0907170
12	30.2200	30.3961	-0.176144
13	29.5100	29.9005	-0.390451
14	28.9000	29.4613	-0.561311
15	28.5000	29.0663	-0.566288
16	28.1500	28.7067	-0.556718
17	27.8400	28.3763	-0.536312
18	27.5600	28.0704	-0.510351
19	27.3400	27.7852	-0.445199
20	27.1800	27.5180	-0.337993
21	27.1100	27.2664	-0.156435
22	27.0500	27.0287	0.0213476
23	27.0200	26.8031	0.216903

24	27.0100	26.5885	0.421528
25	27.0000	26.3837	0.616316
26	27.0000	26.1878	0.812201

Forecast evaluation statistics

Mean Error	-5.0917e-011
Mean Squared Error	0.19904
Root Mean Squared Error	0.44614
Mean Absolute Error	0.38685
Mean Percentage Error	-0.021606
Mean Absolute Percentage Error	1.3073
Theil's U	0.95671