MODEL SELECTION IN MULTIPLE REGRESSION MODELS USING BUDGETED PROFIT, PRODUCTION AND SALES VARIABLES

Iheagwara Andrew I.
Director, Imo State Bureau of Statistics
No. 35 Chief Executive Quarters, Area "B", New Owerri, Imo State, Nigeria

Okenwe Idochi

Department of Statistics, School of Applied Sciences, Rivers State Polytechnic
PMB 20, Bori, Rivers State Nigeria

Ihekuna Stephen, O.
Department of Statistics, Imo State University
PMB 2000, Owerri Nigeria

ABSTRACT

This study is on model selection in multiple regression models. Data for this study were collected in Nigerian bottling company plc, Owerri plant from 1999 to 2013. The response variable is budgeted profit, while the explanatory variables are budgeted production and budgeted sales. Four regression models; Linear, Lin-Log, Polynomial, and Inverse were examined in this study. The E-views software was used in this study. Four model selection techniques known as; coefficient of determination, Akaike Information Criterion, Schwarz Information Criterion, and Hannan-Quinn Information Criterion were used to select the best model. From the analysis, it can be concluded that the nonlinear models perform better than the linear model. However, in the overall goodness of fit assessment, the study concluded that the polynomial regression model performs far better than the other three regression models used in this study. Therefore, future researchers should look at a similar work by incorporating other nonlinear regression models like Double-Log and Log-Lin Regression models to compare results. It should be noted by future researchers that if Double-Log and Log-Lin Regression models are employed, then Quasi - R^2 is needed instead of R^2 as employed in this study.

Key words: Coefficient of Determination, Akaike Information Criterion, Schwarz Information Criterion, Hannan-Quinn Information Criterion, Regression models

Background to the Study

Fitting of linear and nonlinear models to data is normally employed within all fields of science; pharmaceutical and biochemical assay quantification, even though fitting a linear model to data seldom arises, because most data tend to follow nonlinear models. Nonlinear models exist, and the choice of selecting the right model for the data is a mixture of experience, knowledge about the underlying process and statistical interpretation of the fitting outcome. It is of paramount important in quantifying the validity of a fit by some measure which discriminates a 'good' from a 'bad' fit. Many researchers usually employ a common measure known as the coefficient of determination R^2 used in linear regression when conducting calibration experiments for samples to be quantified (Montgomery et al, 2006). Hence, in the linear perspective, this measure is very intuitive as values between 0 and 1 produce an easy interpretation of how much of the variance in the data is explained by the fit. Even though for some time, it has been established that R^2 is an inadequate measure for nonlinear regression, many scientists and researchers still make use of it in studies dealing with nonlinear data analysis (Nagelkerke, 1991; Magee, 1990). According to Juliano and Williams (1987), several initial and older descriptions for R^2 being of no avail in nonlinear fitting had pointed out this issue but have probably fallen into oblivion. This observation might be due to differences in the mathematical background of trained statisticians and researchers who often employ statistical methods but lack detailed statistical insight (Spiess and Neumeyer, 2010).

Having stated that researchers indiscriminately employ R^2 as a means of assessing the validity of a particular model when dealing with nonlinear data fit, it is stated that R^2 is not an optimal choice in a nonlinear regime as the total sum-of-squares (TSS) is not equal to the regression sum-of-squares (REGSS) plus the residual sum-of-squares (RSS), as is the case in linear regression, and hence it lacks the appropriate interpretation. The rationale behind a high occurrence in solely using R^2 values in the validity of nonlinear models could be as a result of researchers not being aware of this misconception.

Even though the use of only R^2 to access the performance of nonlinear data analysis has been discouraged, this study will employ it together with other three model selection techniques known as; Akaike Information Criterion, Schwarz Information Criterion, and Hannan-Quinn Information Criterion for proper interpretation and conclusion.

Literature Review

Scarneciu et al (2017) worked on Comparison of Linear and Non-linear Regression Analysis to determine pulmonary pressure in hyperthyroidism. The study aimed at assessing the incidence of pulmonary hypertension (PH) at newly diagnosed hyperthyroid patients and at finding a simple model showing the complex functional relation between pulmonary hypertension in hyperthyroidism and the factors causing it. The 53 hyperthyroid patients (H-group) were evaluated mainly by using an echocardiographical method and compared with 35 euthyroid (E-group) and 25 healthy people (C-group). In order to identify the factors causing pulmonary hypertension, the statistical method of comparing the values of arithmetical means was employed. By applying the linear regression method described by a first-degree equation the line of regression (linear model) was determined; by applying the non-linear regression method described by a second degree equation, a parabola-type curve of regression (non-linear or polynomial model) was determined. The study made the comparison and the validation of these two models by calculating the determination coefficient (criterion 1), the comparison of residuals (criterion 2), application of AIC criterion (criterion 3) and use of F-test (criterion 4). The result of the study revealed that from the H-group, 47% have pulmonary hypertension completely reversible when obtaining euthyroidism. The factors causing pulmonary hypertension were identified: previously known- level of free thyroxin, pulmonary vascular resistance, cardiac output; new factors identified in the study- pre-treatment period, age, systolic blood pressure. According to the four criteria and to the clinical judgment, the study considered that the polynomial model (graphically parabola- type) was better than the linear one. The study thereby concluded that the better model showing the functional relation between the pulmonary hypertension in hyperthyroidism and the factors identified in the study was given by a polynomial equation of second degree where the parabola was its graphical representation.

Hamidian et al (2008) researched on comparison of linear and nonlinear models for estimating brain deformation using finite element method. The study presented finite element computation for brain deformation during craniotomy. The results were used to illustrate the comparison between two mechanical models: linear solid-mechanic model, and non linear finite element model. To this end, the study employed a test sphere as a model of the brain, tetrahedral finite element mesh, two models that described the material property of the brain tissue, and function optimization that optimized the model's parameters by minimizing distance between the resulting deformation and the assumed deformation. Linear and nonlinear model assumed finite and large deformation of the brain after opening the skull respectively. By using the accuracy of the optimization process, the study concluded that the accuracy of nonlinear model was higher but its execution time was six time of the linear model.

Aristizábal-Giraldo et al (2016) carried out a study on a comparison of linear and nonlinear model performance of shia_landslide: a forecasting model for rainfall-induced landslides. The study explained that landslides are one of the main causes of global human and economic losses. The study compared the forecasting performance of linear and nonlinear SHIA_Landslide model. The results obtained for the La Arenosa Catchment during the September 21, 1990 rainstorm showed that the nonlinear SHIA_Landslide replicated more accurately landslides triggered by rainfall features.

Hunt and Maurer (2016) did a work on comparison of linear and nonlinear feedback control of heart rate for treadmill running. The purpose of the study was to compare linear (L) and nonlinear (NL) controllers using quantitative performance measures. Sixteen healthy male subjects participated in the experimental L vs. NL comparison. The linear controller was calculated using a direct analytical design that employed an existing approximate plant model. The nonlinear controller had the same linear component, but it was augmented using static plant-nonlinearity compensation. At moderate-to-vigorous intensities, no significant differences were found between the linear and nonlinear controllers in mean RMS tracking error (2.34 vs. 2.25 bpm [L vs. NL], p=0.26) and average control signal power (51.7 vs. 60.8×10^{-4} m^2/s^2, p=0.16), but dispersion of the latter was substantially higher for NL (range 45.2 to 56.8 vs. 30.7 to 108.7×10^{-4} m^2/s^2, L vs. NL). At low speed, RMS tracking errors were similar, but average control signal power was substantially and significantly higher for NL (28.1 vs. 138.7×10^{-4} m^2/s^2 [L vs. NL], $p<0.001$). The performance outcomes for linear and nonlinear control were not significantly different for moderate-to-vigorous intensities, but NL control was overly sensitive at low running speed. Accurate, stable and robust overall performance was achieved for all 16 subjects with the linear controller.

Spiess and Neumeyer (2010) worked on an evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. The intensive simulation approach undermined previous observations and emphasized the extremely low performance of R^2 as a basis for model validity and performance when applied to pharmacological/biochemical nonlinear data. With the 'true' model having up to 500 times more strength of evidence based on Akaike weights, this was only reflected in the third to fifth decimal place of R^2. In addition, even the bias-corrected R^2_{adj} exhibited an extreme bias to higher parameterized models. The bias-corrected AIC and also BIC performed significantly better in this respect. The study concluded that researchers and reviewers should be aware that R^2 is inappropriate when used for demonstrating the performance or validity of a certain nonlinear model. It should ideally be expunged from scientific literature dealing with nonlinear model fitting or at least be supplemented with other methods such as AIC or BIC or used in context to other models in question.

Methodology

Regression Models

Linear: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ … (1)

Lin-Log: $Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 X_2$ … (2)

Polynomial: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 + \beta_3 X_2$ … (3)

Inverse: $Y = \beta_0 + \beta_1 (1/X_1) + \beta_2 X_2$ … (4)

Regression Analysis

Regression analysis is a statistical technique that express mathematically the relationship between two or more quantitative variables such that one variable (the dependent variable) can be predicted from the other or others (independent variables). It is very useful in predicting or forecasting. It can also be used to examine the effects that some variables exert on others. It may be simple linear, multiple linear or non linear.

However, the study concentrates on the multiple regression due to the nature of our data.

Multiple Linear Regression

If a regression model involves more than one explanatory variable, it is called a multiple regression model and is of the form

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \quad \ldots \quad (5)$$

Considering the case of only two independent variables as it agrees with the nature of the data for this study, Equation (6) is given:

$$Y_i = \beta_0 + \beta_i X_{1i} + \beta_2 X_{2i} + u_i \quad \ldots \quad (6)$$

where Y is the response variable, X_1 and X_2 are the explanatory variables (or regressors), u_i is the stochastic disturbance term, and i is the ith observation; in case the data are time series, the subscript t will denote the tth observation. Since the data are the form t (t= 1,2,3,…., 15, the number years under study), then Equation (6) will now be written as

$$Y_t = \beta_0 + \beta_i X_{1t} + \beta_2 X_{2t} + u_t \quad \ldots \quad (7)$$

when Y, X_1, X_2 are in deviation forms, then Equation (7) becomes

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \qquad \ldots \qquad (8)$$

where

$$y_t = Y_t - \overline{Y},\ x_{2t} = X_{2t} - \overline{X}_{2t},\ x_{1t} = X_{1t} - \overline{X}_{1t}$$

$$\hat{y}_t = \hat{\beta}_1 x_{it} + \hat{\beta}_2 x_{2t}$$

$$e_t = y_t - \hat{y}_t = y_t - \hat{\beta}_1 x_{1t} - \hat{\beta}_2 x_{2t} \qquad \ldots \qquad (9)$$

Hence, the error sum of squares is given by;

$$\Sigma e_t^2 = \Sigma(y_t - \hat{\beta}_1 x_{1t} - \hat{\beta}_2 x_{2t})^2 \qquad \ldots \qquad (10)$$

Using the OLS technique, the estimate of the parameters is as follows:

$$\hat{\beta}_1 = \frac{\Sigma x_1 y \Sigma x_2^2 - \Sigma x_1 x_2 \Sigma x_2 y}{\Sigma x_1^2 \Sigma x_2^2 - (\Sigma x_1 x_2)^2} \qquad \ldots \qquad (11)$$

$$\hat{\beta}_2 = \frac{\Sigma x_1^2 \Sigma x_2 y - \Sigma x_1 x_2 \Sigma x_1 y}{\Sigma x_1^2 \Sigma x_2^2 - (\Sigma x_1 x_2)^2} \qquad \ldots \qquad (12)$$

and $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}_1 - \hat{\beta}_2 \overline{X}_2 \qquad \ldots \qquad (13)$

However, given the sample size n, the totals $\sum x_1$, $\sum x_2$, $\sum y$, the sums of squares Σx_1^2, Σx_2^2 and cross products $\sum x_1 x_2$, $\sum x_1 y$ and $\sum x_2 y$, their respective sum of squares and cross products adjusted for means are obtained using the formula:

$$\Sigma x_1^2 = \Sigma(X_1 - \overline{X}_1)^2 = \Sigma X_1^2 - \frac{(\Sigma X_1)^2}{n} \qquad \ldots \qquad (14)$$

$$\Sigma y^2 = \Sigma(Y - \overline{Y})^2 = \Sigma Y^2 - \frac{(\Sigma Y)^2}{n} \qquad \ldots \qquad (15)$$

$$\Sigma x_2^2 = \Sigma(X_2 - \overline{X})^2 = \Sigma X_2^2 - \frac{(\Sigma X_2)^2}{n} \qquad \ldots \qquad (16)$$

$$\Sigma x_1 x_2 = \Sigma(X_1 - \overline{X})(X_2 - \overline{X}) = \Sigma X_1 X_2 - \frac{\Sigma X_1 \Sigma X_2}{n} \qquad \ldots \qquad (17)$$

$$\Sigma x_1 y = \Sigma(X_1 - \overline{X}_1)(Y - \overline{Y}) = \Sigma X_1 Y - \frac{\Sigma X_1 \Sigma Y}{n} \qquad \ldots \qquad (18)$$

$$\Sigma x_2 y = \Sigma(X_2 - \overline{X}_2)(Y - \overline{Y}) = \Sigma X_2 Y - \frac{\Sigma X_2 \Sigma Y}{n} \qquad \ldots \qquad (19)$$

Coefficient of Determination

The (multiple) coefficient of determination is given by

$$R^2 = \frac{\hat{\beta}_1 \Sigma x_1 y + \hat{\beta}_2 \Sigma x_2 y}{\Sigma y^2} \qquad \ldots \qquad (20)$$

where x_1, x_2, y are in deviation form. The adjusted R^2 written as $\overline{R}^2$ is defined by

$$\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} \qquad \ldots \qquad (21)$$

Table 1: **ANOVA Table**

Source of variation	Df	SS	MS
Regression	2	$\Sigma\hat{y}_t^2$	$\frac{\Sigma\hat{y}_t^2}{2}$
Error	$n-3$	$\Sigma y_t^2 - \Sigma\hat{y}_t^2$	$\Sigma y_t^2 - \frac{\Sigma\hat{y}_t^2}{n-3}$
Total	$n-1$	Σy_t^2	

$$F_{calculated} = \frac{\frac{\Sigma\hat{y}_t^2}{2}}{\frac{\Sigma y_t^2 - \Sigma\hat{y}_t^2}{n-3}} \quad \ldots \quad (22)$$

$$= \frac{RMS}{EMS} \quad \ldots \quad (23)$$

$$TSS = \Sigma y_t^2 \quad \ldots \quad (24)$$

$$RSS = \Sigma\hat{y}_i^2 = \hat{\beta}_1\Sigma x_1 y + \hat{\beta}_2\Sigma x_2 y \quad \ldots \quad (25)$$

$$ESS = TSS - RSS = \Sigma y_t^2 - \Sigma\hat{y}_t^2 \quad \ldots \quad (26)$$

The nonlinear regression models; Lin-Log, Polynomial, and Inverse can follow the same procedure as employed in the linear model.

Akaike Information Criterion (AIC)

The Akaike's information criterion AIC (Akaike, 1974) is a measure of the goodness of fit of an estimated statistical model and can also be used for model selection. Thus, the AIC is defined as;

$$AIC = e^{\frac{2k}{n}} \frac{\sum \hat{u}_i^2}{n} = e^{\frac{2k}{n}} \frac{RSS}{n} \quad \ldots \quad (27)$$

where k is the number of regressors (including the intercept) and n is the number of observations. For mathematical convenience, Equation (27) is written as;

$$\ln(AIC) = \left(\frac{2k}{n}\right) + \ln\left(\frac{RSS}{n}\right) \quad \ldots \quad (28)$$

where ln (AIC) = natural log of AIC and $\frac{2k}{n}$ = penalty factor.

Schwarz Information Criterion (SIC)

Schwarz Information Criterion SIC (Schwarz, 1978) is a measure of the goodness of fit of an estimated statistical model and can also be used for model selection. It is defined as

$$SIC = n^{\frac{k}{n}} \frac{\sum \hat{u}_i^2}{n} = n^{\frac{k}{n}} \frac{RSS}{n} \qquad \ldots \qquad (29)$$

Transforming Equation (29) in natural logarithm form, it becomes (See Equation (30));

$$\ln(SIC) = \frac{k}{n}\ln(n) + \ln\left(\frac{RSS}{n}\right) \qquad \ldots \qquad (30)$$

where $\frac{\mathrm{k}}{\mathrm{n}}\ln(\mathrm{n})$ is the penalty factor.

Hannan-Quinn Information Criterion (HQIC)

The Hannan-Quinn Information Criterion HQIC (Hannan and Quinn, 1979) is a measure of the goodness of fit of an estimated statistical model and is often employed as a criterion for model selection. It is defined as

$$HQIC = n\ln\frac{RSS}{n} + 2k\ln(\ln n) \qquad \ldots \qquad (31)$$

Where n is the number of observations, k is the number of model parameters. RSS is the residual sum of squares that result from the statistical model.

For model comparison, the model with the lowest AIC, SIC score is preferred.

Data Analysis

Data used for this study is secondary obtained from Nigeria bottling company plc, Owerri Plant Annual Report from 1999-2013. The regression models; Linear, Lin-Log, Polynomial, and Inverse were analyzed via e-view software. The data for the 15 selected years of budgeted profit, budgeted production and budgeted sales are shown in Table 2.

Table 2: Budgeted Profit, Budgeted Production and Budgeted Sales

Year	**Y (₦'000)**	**X_1 (₦'000)**	**X_2 (₦'000)**
1999	500000	269100	2200000
2000	550000	306900	1950000
2001	600000	346500	1800000
2002	790000	430500	1580000

2003	800000	507000	120000
2004	880000	508635	1090000
2005	900000	548895	920000
2006	950000	5875520	918000
2007	1000000	542800	851000
2008	1100000	664900	801000
2009	1500000	828000	780000
2010	2000000	1074400	678000
2011	2100000	1116600	550000
2012	2300000	1248000	495000
2013	2350000	1474000	448500

Source: Nigerian bottling company plc, Annual Report (1999 – 2013)

Y = Budgeted Profit (BPF)
X_1 = Budgeted Production (BPD)
X_2 = Budgeted Sales (BS)

Table 3: Summary of Regression for Linear Model

Dependent Variable: BPF
Method: Least Squares
Date: 08/26/18 Time: 13:07
Sample: 1999 2013
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1955066.	317683.5	6.154133	0.0000
BPD	-0.001061	0.104093	-0.010197	0.9920
BS	-0.723860	0.239100	-3.027438	0.0105

R-squared	0.445785	Mean dependent var	1221333.
Adjusted R-squared	0.353416	S.D. dependent var	653079.2
S.E. of regression	525143.8	Akaike info criterion	29.35759
Sum squared resid	3.31E+12	Schwarz criterion	29.49920
Log likelihood	-217.1819	Hannan-Quinn criter.	29.35608
F-statistic	4.826129	Durbin-Watson stat	0.635389
Prob(F-statistic)	0.028978		

Source*: *E-view software

Table 4: Summary of Regression for Lin-Log Model

Dependent Variable: BPF
Method: Least Squares
Date: 08/26/18 Time: 13:21
Sample: 1999 2013
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1157112.	2986576.	-0.387438	0.7052
LOG(BPD)	218708.9	209196.4	1.045472	0.3164
BS	-0.563449	0.270291	-2.084602	0.0591

R-squared	0.492047	Mean dependent var	1221333.
Adjusted R-squared	0.407388	S.D. dependent var	653079.2
S.E. of regression	502748.7	Akaike info criterion	29.27043
Sum squared resid	3.03E+12	Schwarz criterion	29.41204
Log likelihood	-216.5282	Hannan-Quinn criter.	29.26892
F-statistic	5.812117	Durbin-Watson stat	0.636254
Prob(F-statistic)	0.017177		

Source*: *E-view software

Table 5: Summary of Regression for Polynomial Model

Dependent Variable: BPF
Method: Least Squares
Date: 08/26/18 Time: 13:23
Sample: 1999 2013
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-404745.6	117515.1	-3.444202	0.0055
BPD	2.535798	0.115847	21.88914	0.0000
BPD^2	-3.95E-07	1.79E-08	-22.11360	0.0000
BS	0.106143	0.052733	2.012838	0.0693

R-squared	0.987808	Mean dependent var	1221333.
Adjusted R-squared	0.984482	S.D. dependent var	653079.2
S.E. of regression	81354.02	Akaike info criterion	25.67419
Sum squared resid	7.28E+10	Schwarz criterion	25.86300
Log likelihood	-188.5564	Hannan-Quinn criter.	25.67218
F-statistic	297.0658	Durbin-Watson stat	1.686317
Prob(F-statistic)	0.000000		

Source*: *E-view software

Table 6: Summary of Regression for Inverse Model

Dependent Variable: BPF
Method: Least Squares
Date: 08/26/18 Time: 13:24
Sample: 1999 2013
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2098921.	249953.6	8.397239	0.0000
1/BPD	-3.96E+11	2.07E+11	-1.910304	0.0803
BS	-0.189683	0.345823	-0.548498	0.5934

R-squared	0.575019	Mean dependent var	1221333.
Adjusted R-squared	0.504189	S.D. dependent var	653079.2
S.E. of regression	459858.1	Akaike info criterion	29.09208
Sum squared resid	2.54E+12	Schwarz criterion	29.23369
Log likelihood	-215.1906	Hannan-Quinn criter.	29.09057
F-statistic	8.118287	Durbin-Watson stat	0.547628
Prob(F-statistic)	0.005891		

Source*: *E-view software

Discussion of Results

Having carried out the analysis based on the linear and nonlinear regression models, the results are summarized in Table 7.

Table 7: Summary Result of the Linear and Nonlinear Regression Models

Model Form	AIC	SIC	HQIC	R^2
Linear	29.35759	29.49920	29.35608	0.445785
Lin-Log	29.27043	29.41204	29.26892	0.492047
Polynomial	25.67419	25.86300	25.67218	0.987808
Inverse	29.09208	29.23369	29.09057	0.575019

Looking at the summarized results in Table 7, it can be observed that the polynomial regression model has the highest coefficient of determination (0.987808) with the lowest AIC (25.67419), SIC (25.86300), and HQIC (25.67218), which makes it the best model with respect to the data used in this study. The next to polynomial regression model is inverse model which has a coefficient of determination of 0.575019 with the AIC (29.09208), SIC (29.23369), and HQIC (29.09057). It is clear from the result that the linear regression model is the least performed model.

Conclusion

From the analysis, it can be concluded that the nonlinear models perform better than the linear model. However, in the overall goodness of fit assessment, the study concluded that the polynomial regression model performs far better than the other three regression models used in this study. Therefore, future researchers should look at a similar work by incorporating other nonlinear regression models like Double-Log and Log-Lin Regression models to compare results. It should be noted by future researchers that if Double-Log and Log-Lin Regression models are employed, then Quasi - R^2 is needed instead of R^2 as employed in this study.

REFERENCES

Akaike, H. (1974), *"A new look at the statistical model identification"* (PDF), IEEE Transactions on Automatic Control 19 (6): 716–723, *doi:10.1109/TAC.1974.1100705, MR 042371*

Aristizábal-Giraldo, E. V., Vélez-Upegui, J. I., and Martínez-Carvaja, H. E. (2016). A comparison of linear and nonlinear model performance of shia_landslide: a forecasting model for rainfall-induced landslides. Revista Facultad de Ingeniería, No. 80, pp. 74-88, 2016

Hamidian, H., Soltanian-Zadeh, H., Akhondi-Asl, A., Faraji-Dana, R. (2008). Comparison of Linear and Nonlinear Models for Estimating Brain Deformation Using Finite Element Method. In: Sarbazi-Azad H., Parhami B., Miremadi SG., Hessabi S. (eds) Advances in Computer Science and Engineering. CSICC 2008. Communications in Computer and Information Science, vol 6. Springer, Berlin, Heidelberg

Hannan, E. J., and Quinn, B. G. (1979). "The Determination of the order of an autoregression", *Journal of the Royal Statistical Society, Series B*, 41: 190–195.

Hunt, K. J. and Maurer, R.R. (2016). Comparison of linear and nonlinear feedback control of heart rate for treadmill running. Systems Science & Control Engineering, 4:1, 87-98, DOI: 10.1080/21642583.2016.1179139

Juliano, S. A., Williams, F.M. (1987). A comparison of methods for estimating the functional response parameters of the random predator equation. J Anim Ecol. 1987;56:641–653. doi: 10.2307/5074.

Magee L. (1990). R^2 measures based on Wald and likelihood ratio joint significance tests. Amer Stat. 1990; 44:250–253. doi: 10.2307/2685352.

Montgomery, D.C., Peck, E.A., Vining, G.G. (2006). Introduction to Linear Regression Analysis. Wiley & Sons, Hoboken; 2006.

Nagelkerke, N.J.D. (1991). A note on a general definition of the coefficient of determination. Biometrika. 1991;78:691–692. doi: 10.1093/biomet/78.3.691.

Scarneciu, C.C., Sangeorzan, L., Rus, H., Scarneciu, V.D., Varciu, M.S. Andreescu, O. and Scarneciu, I. (2017). Comparison of Linear and Non-linear Regression Analysis to Determine Pulmonary Pressure in Hyperthyroidism. Pak J Med Sci. 2017 Jan-Feb; 33(1): 111–120.

Schwarz, G. (1978). Estimating the dimension of a model. The Annals of Statistics 6, 461-464

Spiess, A. and Neumeyer, N. (2010). An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. BMC Pharmacol. V.10; 2010