AMSE JOURNALS-AMSE IIETA publication-2017-Series: Advances D; Vol. 22; N°1; pp 1-15

Submitted Oct. 2016; Revised May 1, 2017, Accepted May 30, 2017

Modelling Telecommunication Pathway in Nigeria

*Eno E. E. Akarawak, **Ismaili I. Adeleke

*Department of Mathematics, Faculty of Science, University of Lagos, Akoka, Lagos, Nigeria; eakarawak@unilag.edu.ng

**Department of Actuarial Science & Insurance, Faculty of Business Administration, University of Lagos, Akoka, Lagos, Nigeria; iadeleke@unilag.edu.ng

Abstract

The aim of this work is to model the pathway from caller to recipient of GSM telecommunication in Nigeria, with a view to produce a model that can help reduce the problem of drop calls experienced in the industry. Our dependent variable was total successful calls against 9 explanatory variables. An initial multiple linear regression produced low R^2 of 25.5%. Diagnostics interventions of some transformations with removal of leverage points improved the R^2 to 82%. Two model selection techniques, Mallow's Cp and adjusted R^2 were used to obtain the best parsimonious model, which contained 7 explanatory variables. The results show that the main variables that explain total successful call are Percentage drop calls, Proportion of transmission failure, Call traffic Congestion, Control channel failure, Earlang, P_HR and Availability. We, therefore, advise telecommunication industries in Nigeria to use the model to counteract the problem of drop calls.

Key words*:* Telecommunication pathway, drop calls, Box-Cox transformation, Mallow's Cp, model selection.

1. Introduction

The Global System of Mobile Telecommunication (GSM) in Nigeria came into mainstream in 2001 when the deregulation of the subsector of the economy gave way to private involvement. Until then, the Nigerian Telecommunication (NITEL) was the only operator in the market. The deregulation was generally accepted and brought hope of a better society to Nigerians. The industry has since grown from 500000 NITEL subscribers to about 65 million mobile subscribers (Fadeyibi 2009). The advantages of the wireless mobile telecommunication system over the fixed system are numerous. The benefits include competition in the market, affordability, employment opportunity, revenue generation and increased accessibility even in rural areas.

Despite these benefits subscribers still have ugly stories to tell as mobile telecommunication in Nigeria is characterized by high drop calls, among other problems (Fadeyibi 2009). Drop call is the failure of a call to successfully move from the caller to the receiver. The problem of drop calls cuts across all GSM networks and sometimes makes the use of GSM very frustrating. In this work, our objective is to model the pathway of a GSM call from the caller to the receiver using multiple linear regression and model selection techniques. Our response variable is total successful calls of a GSM network and 9 explanatory variables claimed by mobile network operators to affect the success of a call.

We propose a multiple linear regression model of the form: $Y=X\beta + e$;

while
Y is a vector of the response variable
X is the design matrix
β is the vector of regression parameters
e is the vector of the random noise
With the assumptions that $Y \sim \mathrm{N}(X\beta, \sigma^2 I)$ where σ^2 is the variance of the random noise. Adeleke *et al*. (2007) has illustrated the use of Box-Cox transformation to correct multiple assumption violations in regression analysis of telecommunication data. In this work, we intend to use some model selection techniques to select variables that actually explain total successful calls and build the best parsimonious model for telecommunication call pathway in Nigeria.

The remainder of the article is organized as follows: section 2 discusses relevant literature; in section 3, we dwell on data description; section 4 is on methods used, section 5 discusses the results and conclusion is in section 6.

2. Relevant Literature on Model Selection

Model selection is very important in multiple linear regressions. This is so because the true model is not always known and some of the variables used might not have significant explanatory power on the response variable. However, the chosen model is assumed to be the true model and analysis and inference are done accordingly (Hurvich & Tsai, 1990).

Many model selection techniques have been proposed in the literature. They include F-test, Akaike Information Criterion (AIC), Mallow's C_p, exhaustive search, stepwise methods, adjusted R^2, forward and backward procedures. Others are cross-validation, Baye's factor, Bayes Information Criterion (BIC), Bayesian model averaging, Hannan-Quinn Criterion (HQC), LASSO, etc. These techniques can be classified as Bayesian or frequentist (information-based) techniques. In comparison, the Bayesian model techniques are generally statistically consistent, but sometimes achieve slower rates of convergence than other methods (Erven, et al 2009). Kwon, et al, 2009 studied the performance of some Bayesian and information-theoretic model selection techniques and concluded that they work well in small samples as well as large samples. This helps to avoid cumbersome computations with large samples.

Applications of the various model selection techniques in different situations have been documented. For example, to explain the effectiveness of advertisement, Lee (2016) used LASSO and AIC to predict attention score based on 23 predictor variables from questionnaire responses; and found that the LASSO method provided simpler and more stable results. Also, Kadane and Lazar (2004) evaluate the various proposed frequentist and Bayesian techniques, including AIC, Bayes Factors, BIC, Mallow's Cp, Model Averaging, Subset Selection, from a decision-theoretic perspective and proposed a unifying conceptual framework which can guide people on when to choose Bayesian or frequentist model selection techniques. In another study involving the drying characteristics of fresh grains, Iwundu and Efezino (2015) examine the adequacy of variable selection techniques using some model selection criteria namely, R^2, R^2adj, PRESS, AIC and Cp-statistic to determine the most suitable model; and in addition, illustrate the use of D-optimality criterion for measuring the goodness and adequacy of regression models. In order to propose a hybrid fuzzy time series model to forecast weather, Agrawal and Qureshi (2014) employed root mean square error, R^2 and moving average.

The task of model choice is critical and challenging. In a study on system identification and model signaling, Pintelon *et al*. (1997) posits that one major problem of model selection techniques is detection of under-modelling. In line with this, Kadane & Lazar (2004) notes that there may be cases when model choice is unobjectionable and other cases when the choice is misleading. Supporting these views, Hurvich & Tsai, 1990 maintains that "model selection stage of linear regression induces some difficulties in the analysis that if not well handled can affect the validity of standard regression procedures". These difficulties includes cases like inflated R^2 (Rencher & Pun, 1980) and biased estimates of mean squared prediction errors (Breiman, 1988). The goal of model selection is to select the best model that explains the given data. With this in mind, Kundu & Murali, 1995 applies model selection techniques like AIC, BIC, etc to linear regression so as to choose the best penalty function.

Not much work, if any, has been done to select the best model that explains successful calls of a GSM network. In this work, we intend to select signaling variables on the pathway of a GSM call that best explains the success of the call.

3. Data Description and Exploration

3.1 Data Description

The data represents 752 observations which were obtained on the response variable Y and 9 explanatory variables. The response variable is the total successful G.S.M calls (TSC), representing the calls that successfully go through the pathway of a G.S.M call from the maker to the recipient. A GSM call moves from the caller to the Base Transceiver Station (BTS) to the Base Station Controller (BSC) to the Mobile Switching Center (MCS), where the switching takes place; then, to BTS2, to BSC2 and finally, to the receiving mobile phone. There are a lot of signaling that go on between all these points and the factors which are claimed to determine a successful call (TSC) are Percentage drop calls (X_1), Power transmission failure (X_2), Call traffic Congestion (X_3), Power control Congestion (X_4), Control channel failure (X_5), Power supply failure (X_6), Earlang or System run time (X_7), P_HR (X_8) and Availability (X_9). A general characteristic of the data is that some observations of Y and the X variables are extremely high or low and these largely corresponded to periods of systems failure or malfunction. Analysis in due course shows that these largely constituted leverage points.

3.2 Exploratory Data Analysis

Exploratory Data Analysis shows that both response and regressor variables showed clear departures from normality, whereas the assumption of the linear model (in matrix form) $Y = X\beta + e$ is that $Y \sim \mathrm{N}(X\beta, \sigma^2 I)$.

4. Methodology

4.1 Multiple Linear Regression Model

4.1.1 Overview

Multiple Linear Regression Model is a type of model in which a dependent variable y is determined by two or more explanatory variables $X_1, X_2, \ldots, X_k$ (Omotosho, 2000). The model is of the form $Y = X\beta + e$. β is estimable by $\hat{\beta} = (X'X)^{-1}X'Y$, with variance $V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$. We determine the scalar σ^2 using the estimator, $S^2 = \dfrac{Y'Y - \hat{\beta}'X'X\hat{\beta}}{n-k}$, where n is the number of observations and k is the number of parameters. Accordingly, a t-test given by $t^* = \dfrac{\hat{\beta}_i - 0}{S\sqrt{V_i}}$, where V_i is the ith diagonal element of $(X'X)^{-1}$ is used to test the null hypothesis H_0: $\beta_i = 0$. H_0 is rejected if t^* is greater than critical t at a given significance level, say $\alpha = 0.05$ (i.e, $p < 0.05$).

4.1.2 Transformations

Since the exploratory data analysis revealed departures from the normality assumption, deviations from other important assumptions like homoscedasticity and independence of the random noise were suspected. Therefore, there was need for a transformation that could correct all departures at the same time. Box-Cox transformation offers real opportunity. Before Box-Cox transformation on the response variable, a power transformation on the X variables was done to induce normality.

4.1.3 Transformations on the Explanatory Variables

We carried out various transformations of the X variables using simple power transformations. The method was computer intensive. For each X_i we applied:

$$X_i^1 = X_i^\lambda, \tag{1}$$

where λ is such that X_i^1 had the best Q-Q plot, that is, X_i^1 was closest to a normal distribution. The λ that could do this best was determined by dividing the interval (-3, 3) into 250 sub-intervals. Then, λ was that value in (-3, 3) which gave the best Q-Q plot. With the X's transformed, we then ran another regression and R^2 improved to 35%. The Q-Q plot of each X_i^1 and the respective λ are shown Fig 1.

4.1.4 The Box-Cox Transformation

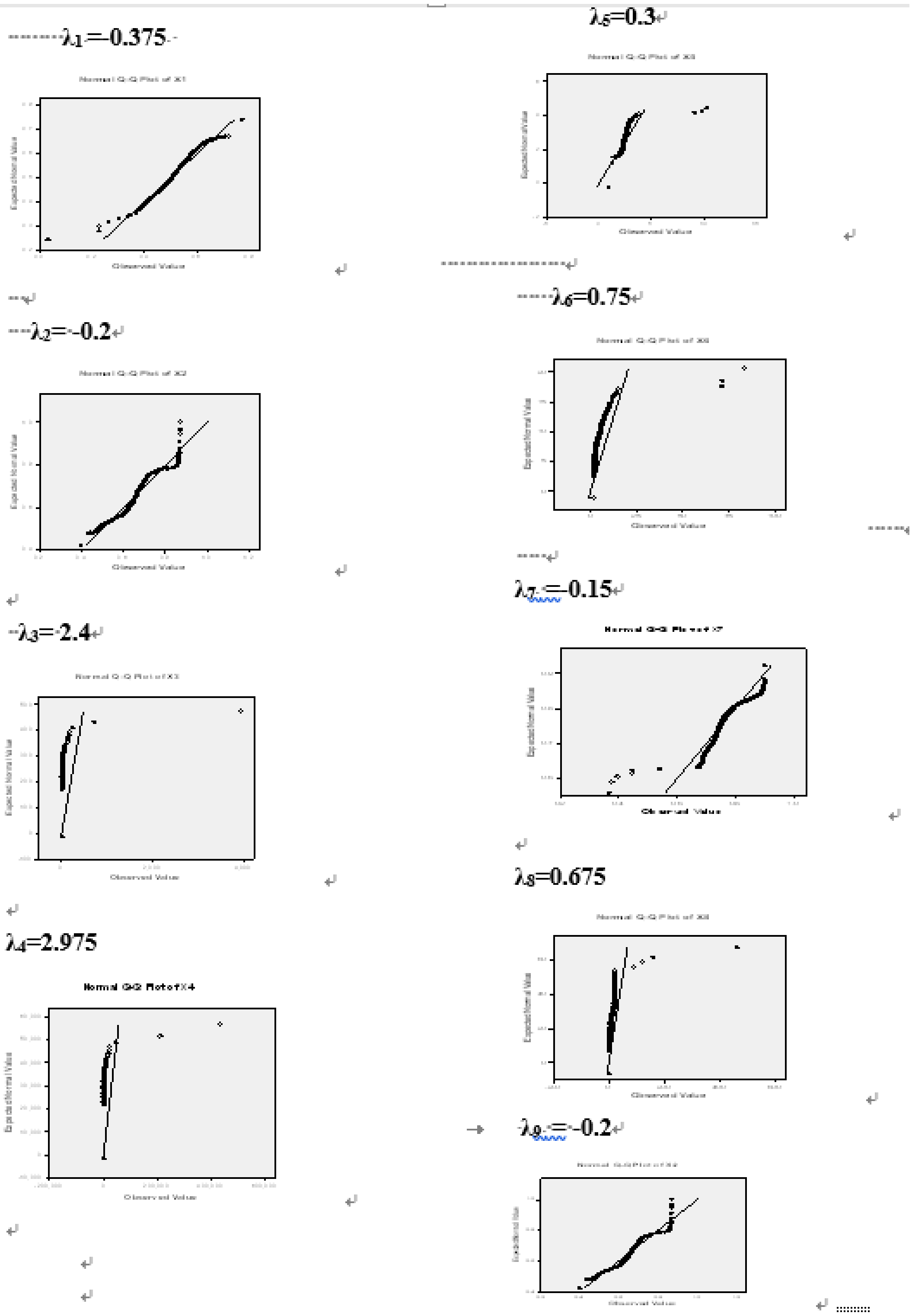

Fig 1. Normal Q-Q plots of the transformed *X* variables

When the residual vector of a regression analysis shows heteroscedastic error variances, non-normality and non-independence, “it is natural to seek a transformation that, so far as possible, both satisfies and combines information from the three desirata simultaneously” (Box, Hunter & Hunter, 1978). This can be achieved through the method of maximum likelihood where the individual values of y (observations) are transformed into $y^{(\lambda)}$ such that:

$$y^{(\lambda)} = \frac{y^{\lambda} - 1}{\lambda \dot{y}^{\lambda - 1}}, \qquad (2)$$

The transformation in (2) is a Box-Cox transformation; where $\dot{y}$, the geometric mean of y, is the scaling factor and λ, the Box-Cox parameter, is the value at which the transformed data attains the smallest standard deviation or produces the smallest mean square error when a linear regression is fitted to the data. Many modifications have been proposed since the work of Box and Cox (Li, 2005). However, we choose the modification in (2) because of its relative simplicity.

For various values of λ, we perform standard regression analysis on $y^{(\lambda)}$. "The Maximum likelihood values is that for which the residual sum of squares (S_λ, say) from the fitted model is minimized" (Box, Hunter & Hunter, 1978). Here a computer generated system chooses a λ, and on this basis transforms y_i into $y_i^{(\lambda)} = \frac{y_i^{\lambda} - 1}{\lambda \dot{y}_i^{\lambda - 1}}$

Then a regression of $Y_i^{(\lambda)} = \beta_o + \beta_1 X^1{}_{1i} + \ldots + \beta_9 X^1{}_{9i} + \varepsilon$ is run and the error sum of squares calculated. The λ which minimizes the error sum of squares is calculated. The λ for this Box-Cox transformation was 0.52 (see Fig 2). A regression after Box-Cox transformation improved R^2 to 75.5%.

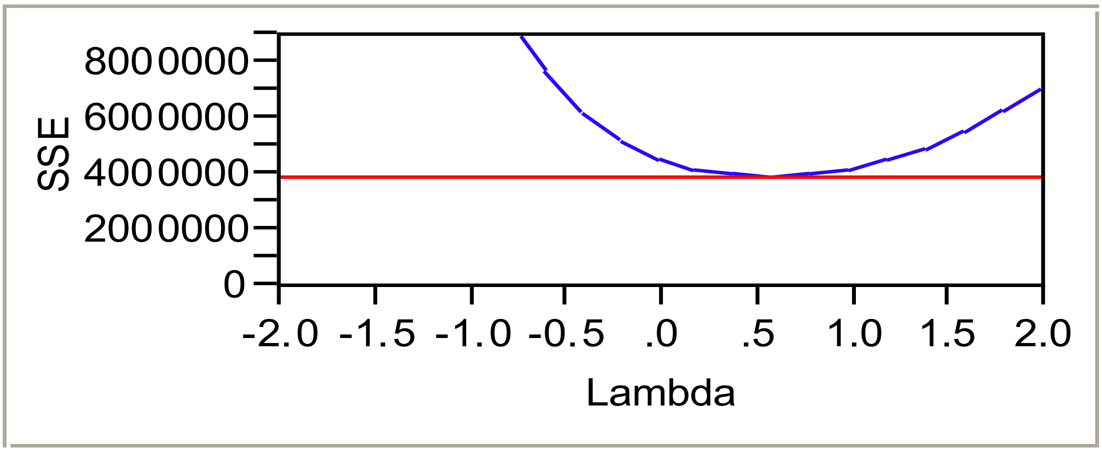

Fig 2. Plot of SSE against Lambda for Box-Cox Transformations

4.1.5 Removal of Leverage Points

Generally, leverage points are those points in a system where a small change in one thing can result in much change in every other thing. In regression analysis, a leverage point is a value of an independent variable *X*, which is far away from other values of that independent variable and can exert undue influence on the regression of the response variable on *X*.

For this study, the very nature of the data itself made leverage points inevitable. For instance, successful call rates during system breakdown would be far from the norm. The residual versus the predicted plot showed this (see Fig 3).

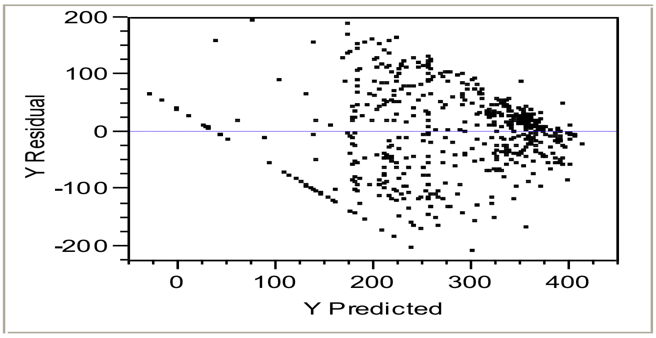

Fig 3. Residual by Predicted Plot

The leverage points are more vivid when one examines the separate leverage plots for the *X* variables. First of all, identification of a leverage point is based on the value of the diagonal element (m_{ii}) of the hat matrix (H)

$$H = [I - X(X^1X)^{-1}X^1], \tag{3}$$

where X is the design matrix and I is the identity matrix of order n. Thus, m_{ii} shows the amount of influence which y_i has on the regression. If m_{ii} is more than 3 times the average of the m_{ii}'s then, it was removed (Hoaghin & Welsch 1978). In theory, not all leverage points are removable but in the instant data, there is a clear case for removal as most of these points correspond to the abnormal period mentioned earlier. With the removal of leverage points, R^2 improved to 82%.

4.2 Model Selection

Two known criterion-based methods of model selection were employed, the adjusted R^2 and Mallow's C_p statistics.

4.2.1 The Adjusted R^2 Criterion

R^2 is defined as the ratio of the sum of squares of regression to the total sum of squares, i.e.,

$$R^2 = \frac{SSR}{SST} \text{ X } 100\% = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y - \bar{y})^2} \text{ X}100\%, \tag{4}$$

The problem with using R^2 as a criterion, for comparing models of different sizes, is that the sum of squares for regression, and hence R^2 itself increases with increasing variables in the model. Therefore, the adjusted R^2, which takes into consideration the number of parameters in the model is usually used instead (Kadane & Lazar, 2004). Adjusted R^2 is defined as:

$$R^2_{adj} = 1 - \frac{n-1}{(n-p)(1-R^2)} \tag{5}$$

where n is the sample size and p is the number of parameters in the model. The best model will correspond to the largest adjusted R^2.

4.2.2 Mallow's C_p Statistic:

Another criterion used is the C_p statistic. It is defined as

$$C_p = \frac{RSS_p}{\hat{\sigma}^2 + (2p - n)} \tag{6}$$

where RSS_p represents the residual sum of squares for a model with p terms and $\hat{\sigma}^2$ the random noise variance estimate based on the full model (Mallow, 1973). If the model is good, it is expected that C_p is unbiased estimator of p and therefore, should itself be approximately equal to p. For a full model with say m parameters, this is exactly true, ie, C_m = m (Kadane & Lazar, 2004). This feature makes the criterion useful to compare models of the same size and its purpose is to guide the researcher in the process of model selection (George 2000). Models with smallest C_p are desirable.

4.2.3 Application to data

Out of 651 complete cases, 512 models were fit for our data. Various models to describe the relationship between Y and 9 predictor variables were fit containing combinations of from 0 to 9 variables. The statistics tabulated include the mean squared error (MSE), the adjusted R^2 and Mallow's C_p values.

To determine which model is best according to the two criterion discussed, the result was summarized according to models with the largest adjusted R^2 and models with the smallest C_p. Based on these the best model was selected to contain 7 independent Variables: X_1, X_2, X_3, X_5, X_7, X_8 and X_9.

It is important to note that both criteria resulted in the same model. This is in line with Kennard (1971), who maintains that the C_p statistic is closely related to the adjusted R^2.

5. Results and Discussion

5.1 Results of Initial Regression

Table 5.1. Results of Initial Regression

*R*Square	0.255387
*R*Square Adj	0.2458
Root Mean Square Error	244.8889
Mean of Response	251.4034
Observations (or Sum Wgts)	709

Parameter Estimates

Term	**Estimate**	**Std Error**	**t Ratio**	**Prob>|t|**
Intercept	-97.77922	58.99176	-1.66	0.0979
% Drop (*X*1)	-0.032577	0.024526	-1.33	0.1845
PtFail (*X*2)	0.8088859	0.532861	1.52	0.1295
Congestion (*X*3)	12.626796	6.874496	1.84	0.0667
pcCong (*X*4)	-2.374931	2.071086	-1.15	0.2519
Control Channel (*X*5)	0.0554714	0.055827	0.99	0.3207
PsFail (*X*6)	-0.343025	0.389226	-0.88	0.3785
Erlang (*X*7)	0.2332903	0.167157	1.4	0.1633
P_HR (*X*8)	-0.000423	0.027041	-0.02	0.9875
Availability (*X*9)	3.9559168	0.48325	8.19	<.0001

An initial regression of the data shows a low R^2 of 25.5% (see Table below). This is not a good fit, yet it is claimed that the 9 factors determine successful calls. The *t*-test shows that only availability is significant.

5.1.2 Regression after Transformations on Response and Regressor Variables

R Square	0.755234
R Square Adj.	0.752122
Root Mean Square Error	66.95807
Mean of Response	242.3974
Observations (or Sum Wgts)	718

After a power transformation on the regressor and response variables, another regression ran shows an improvement over the first one using R^2 (R^2 = 75.52%).

5.1.3 Regression after Removal of Leverage Points

R square	0.820425
R Square Adj	0.817903
Root Mean Square Error	57.74261
Mean of Response	243.863
Observations (or Sum Wgts)	651

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1179.2274	52.32196	22.54	<.0001
*X*1	-127.8084	33.09964	-3.86	0.0001
*X*2	-145.8531	27.23642	-5.36	<.0001
*X*3	-0.797684	0.413363	-1.93	0.0541
*X*4	0.0000844	0.001222	0.07	0.9449
*X*5	-16.08319	9.072534	-1.77	0.0767
*X*6	-0.461916	1.822434	-0.25	0.8
*X*7	-1052.972	57.89197	-18.19	<.0001
*X*8	-4.236453	0.572721	-7.4	<.0001
*X*9	2.0939898	0.077348	27.07	<.0001

The results show that R^2 improved greatly to about 82% with the removal of leverages. This implies that the leverage values actually had undue influence on the results. However, only 5 explanatory variables are statistically significant at 95% significance level. This calls for model selection to arrive at a reduced model.

5.1.4 Multiple Regression with 7 independent Variables after Model Selection

Dependent Variable: Y

Parameter	Estimate	Standard Error	T Statistic	p-value
CONSTANT	1179.54	51.8063	22.7682	0.0000
$X1$	-125.636	32.0358	-3.92172	0.0001
$X2$	-142.992	24.6156	-5.80899	0.0000
$X3$	-0.802294	0.411683	-1.94881	0.0518
$X5$	-16.6165	8.63682	-1.92392	0.0548
$X7$	-1056.6	55.683	-18.9752	0.0000
$X8$	-4.25359	0.550803	-7.72252	0.0000
$X9$	2.08945	0.0753184	27.7416	0.0000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	9.76E+06	7	1.39E+06	419.61	0.0000
Residual	2.14E+06	643	3324.21		
Total (Corr.)	1.19E+07	650			

The model selection procedure actually yielded 7 independent variables which significantly explain successful GSM calls.

R-squared = 82.0405 percent

R-squared (adjusted for d.f.) = 81.845 percent

Standard Error of Est. = 57.6559

Mean absolute error = 39.5314

The above is the result of Multiple Linear regression to describe the relationship between Y and 7 explanatory variables. Two of the independent variables, X_3 and X_5 showed p-values greater than 0.05 which is not significant at 95% significance level. However, we decided not to drop the variables because the R^2 appears to be stable (no further increase) at this point. Also, by not removing the variable we tried to avoid under modeling, which is a common problem in model selection. We, therefore, show 95% confidence intervals for estimates in the table below. Looking at the data and the random noise, we believe the confidence intervals for the coefficients are reasonably good.

95% Confidence Intervals for coefficient estimates

Parameter	Estimate	Standard Error	Lower Limit	Upper Limit
CONSTANT	1179.54	51.8063	1077.81	1281.27
$X1$	-125.636	32.0358	-188.543	-62.7279
$X2$	-142.992	24.6156	-191.328	-94.655

$X3$	-0.802294	0.411683	-1.6107	0.0061135
$X5$	-16.6165	8.63682	-33.5763	0.343294
$X7$	-1056.6	55.683	-1165.94	-947.253
$X8$	-4.25359	0.550803	-5.33518	-3.17199
$X9$	2.08945	0.0753184	1.94155	2.23735

6. Conclusion

In this work, we applied the diagnostic interventions of Box-Cox and other power transformations and leverage removal to turn around a regression gone awry. The coefficient of determination R^2 increased greatly from an initial value of about 25% to about 82%. However, in line with the principle of parsimony we used two model selection techniques, Mallow's C_p and adjusted R^2 to arrive at a reduced model adjudged to be the best.

The best model contains 7 explanatory variables: X_1 (percentage drop calls), X_2 (Power transmission failure), X_3 (Call traffic Congestion), X_5 (Control channel failure), X_7 (Earlang or system run time), X_8 (P_HR) and X_9 (Availability). Therefore, the best parsimonious model for GSM telecommunication pathway is:

$$\hat{y} = 1179.54 - 125.64X_1 - 142.99X_2 - 0.80X_3 - 16.62X_5 - 1056.60X_7 - 4.25X_8 + 2.09X_9.$$

We therefore advise that telecommunication industries in Nigeria should adopt this model to restructure their call pathway as this might reduce the problem of drop calls experienced by subscribers in the country. If Nigerian telecommunications operators can solve the problem of drop calls by adapting the proposed model, among other things they need to do, it will greatly enhance their services and confidence of the subscribers.

References

[1] I.A. Adeleke, E.E.E. Akarawak, A.I. Opara, E.O. Esan. "Box-Cox Transformation in Regression Analysis: An Illustration with Telecommunication Data" Proceedings of the Nigerian Statistical Association, Asaba, Nigeria; 2007.

[2] A. Agrawal, M. F. Qureshi, "Indian Weather Forecasting using ANFIS and ARIMA based Interval Type-2 Fuzzy Logic Model"; AMSE Journals, Series Advances D, Vol. 19, no 1; pp 52-70, 2014.

[3] G.E.P. Box, W.G. Hunter, J.S. Hunter. Statistics for Experimenters, John Wiley & Sons, 1978.

[4] L. Breiman. “Submodel Selection and Evaluation in Regression – The Conditional Case and Little Bootstrap”, Technical Report 169, Department of Statistics, University of California, 1988.

[5] T. Erven. Grunwald, P. & Rooij, S.; “Catching Up Faster in Bayesian Model Selection and Model Averaging”, 2009;. http://books.nips.cc/papers/files/nips20/NIPS2007_0756.pdf. Assess date September 2009.

[6] I. Fadeyibi. “Industry Analysis, Nigerian Mobile Telco, 2009; PRLog – Global Press Release Distribution”. www.prlog.org/10242511- industry-analysis-nigeria-mobile-telco.pdf. Assess date September 2009.

[7] E.I. George. “The Variable Selection Problem”; Journal of the American Statistical Association, vol 95, No 452, pp 1304-1308, 2000.

[8] F.A. Graybil. Theory and application of Linear Model; Belmat CA: Weadsworth, 1976.

[9] David C. Hoaghin & Roy E. Welsch; “The Hat Matrix in Regression and ANOVA”; The American Statistician, vol 32, no 1, pp 17-22, 1978.

[10] Hurvich, C.M & Tsai, C.; “The Impact of Model Selection on Inference in Linear Regression”; The American Statistician. Vol 44, no 3, pp 214-217, 1990.

[11] Iwundu, M.P. and Efezino, O.P.; “On the Adequacy of Variable Selection Techniques on Model Building”, Asian Journal of Mathematics & Statistics, vol 8: pp 19-34, 2015.

[12] Kadane, J.B. & Lazar, N.A.; “Methods and Criteria for Model Selection” Journal of the American Statistical Association. Vol 99, no 465. Review Article, 2004.

[13] Kennard, R.W.; “A Note on the C_p Statistic”; Technometrics, vol 13, no 4, pp 899-900 , 1971.

[14] Kundu, D. & Murali, G.; “Model Selection in Linear Regression”, Elsevier: Computational Statistics and Data Analysis. vol. 22, pp 461-469, 1996.

[15] Kwon, Y.; Bozdogan, H. & Bensmail, H.; “Performance of Model Selection Criteria in Bayesian Threhhold VAR (TVAR) Models”; Econometric Reviews. Vol. 28, Issue 1-3, pp 83-101, 2009.

[16] Lee, M. A.; “Application of model selection techniques and measures of agreement to advertising data”, an Unpublished Masters Dissertation, Dept. of Mathematics and Statistics, University of New Mexico; http://repository.unm.edu/handle/1928/32963; 2016.

[17] Mallows, C.L.; “Some Comments on C_p”; Technometrics, vol 15, no 4, pp 661-675; 1973.

[18] M.Y. Omotoso. Econometrics: A Practical Approach; Yosode Book Publishers, Ibadan; 2000.

[19] P.F. Li. “Box-Cox Transformation: An Overview”, Department of Statistics, University of Connecticut, www.stat.ucom.edu/~studentjournal/index; 2005. Assess date August 2007.

[20] R.P. Schoukens, G. Vandersteen, “Model Selection through a Statistical Analysis of Global Minimum of a Weighted Nonlinear Least Squares Cost Function”; IEEE Transactions on Signal Processing, vol. 45, no 3, pp 686-693; 1997.

[21] A.C. Rencher, F.C. Pun, “Inflation of R^2 in Best Subset Regression”; Technometrics, 22, no 1, pp 49-53, 1980.

[22] S. Weisberg. Applied Linear Regression (2nd ed); Wiley, New York, 1985.