Modelling of Input Parameters for Power Generation using Regression Models

Agbondinmwin U.[1] and Ebhojiaye R. S.[2,*]

[1,2]Department of Production Engineering, University of Benin, Benin City, Nigeria

Corresponding Author: *raphael.ebhojiaye @uniben.edu

ABSTRACT

In this study, multiple linear regression models were employed in the correlation of gas supply and power generation using a gas Power Plant in Niger Delta, Nigeria as a Case study. From the analysis based on outlier detection, reliability analysis and test of homogeneity, it was observed that the independent variable data such as ambient temperature, gas pressure and compressed temperature failed normality test. Therefore, the use of any linear model for either analysis or modelling of the data was not acceptable. Data used for reliability analysis of the gas pressure and compressed temperature difference were positively correlated with power generation, having a covariance value of 0.639 and 113.148. The ambient temperature was negatively correlated with power generation, having a covariance value of 14.564. The positive value showed that both dimensions exclusively increased and decreased together with respect to the output while the negative value showed that increment in value of one variable led to decrease in the value of the other, and vice versa.

Keywords: Natural gas, Power generation, Regression models, Power plant, Ambient temperature

1.0. Introduction

Nigeria's natural gas reserves, estimated at about 188 trillion standard cubic feet is the largest in Africa and known to be substantially larger than its oil resources (Nwokeji, 2007; Izuwan, 2017). According to Oyedepo (2012), the largest single consumer of natural gas in Nigeria (before its privatization in November 2013) was the Power Holding Company of Nigeria (PHCN), it accounted for over 70% used in operating electricity-generating gas plant in the country. Given the current reserves and rate of exploitation (about 900mmscfd) for power generation, the expected life span of Nigerian natural gas is over 1000 years, thus making it a good means for power generation. As a result, the gas produced in Nigeria is used mostly in the power sector for power generation and for export as liquefied Petroleum Gas (LPG) (Sambo et al., 2010). Electricity plays a vital role in economic growth and social welfare, thus it is essential to have accessible and reliable electricity at safe conditions (Luis *et al*., 2019).

Rapu *et al*. (2015) put the average generation capacity of electricity in Nigeria to be fluctuating within the range of 2,623.1 MW/hr in 2007 and 3,485.5 MW/hr in 2014 as against the estimated demand of 10,000MW per day. Enete and Alabi (2011) estimated the distribution of household final energy consumption by types in Nigeria to be 4% electricity, 13% kerosene, 1% LPG and 82% wood and others. Currently, over 70% of Nigeria power generation is from natural gas utilization in power plants. There are over 15 power stations in Nigeria with total generation capacity of over 15,000MW that are currently generating about 7000MW (Onochie *et al*., 2015). The problem is attributed to various factors that include gas supply, transmission, grid capacity, plant ambient temperature and operating conditions (such as gas pressure and compressed temperature difference), etc. Orogun (2015) in his work discussed the challenge of gas pipeline vandalisation as one of the major challenge militating against the Federal Government of Nigeria efforts to utilize Nigeria's natural gas sustainably for power generation. This challenge is also compounded due to the inadequacy of natural gas transmission and distribution infrastructure. The power station situated in Benin City, Edo state is

incorporated with a simple cycle gas turbine with over 450MW capacity and has the following coordinates: $6^o24'20"N$ $5^o41'00"E$, with Escravos-Lagos pipeline system as the source of gas supply.

Although some studies such as Oricha and Olarinoye (2012) and Iwuamadi and Dike (2012) have shown that poor plant maintenance, operational policies and power transmission issues are factors that also affect adequate power generation in Nigeria, however, the veracity of these militating factors has been tested on the field assessment in some gas power plants and results show less impact on both plant capacity performance and economic viability when compared to the effect of gas supply to the plants. This paper therefore, attempts to correlate the gas supply and power generation based on statistical regression models.

2.0. Materials and Methods

Questionnaires were designed from the data obtained from the National Integrated Power Project (NIPP) power stations and some relevant literature. Forty-three (43) variables were considered in the questionnaires which was scaled with five (5) point Resits Likert's attitudinal scale and administered to 150 respondents. Respondents' responses were transposed into metric variables. Gas pressure, ambient temperature, compressed temperature difference (CTD) and power generation have been identified as variable (Oyedepo *et al*., 2015) and were used as research parameters in this study. The correlation of gas supply and power generation of gas power plant was done using the multiple linear regression models to assess the data quality, normality of data and the diagnostic analysis of data.

2.1. Assessment of data quality

To assess the quality of the data, three important tests were conducted which includes: outlier detection; data fitness using reliability analysis; and test of homogeneity.

2.1.1. Detection of outliers using the labelling rule
In this study, the labelling rule method was employed to detect the presence of outliers. The labelling rule is the statistical method of detecting the presence of outliers in data sets using the 25th percentile (lower bound) and the 75th percentile (upper bound). The underlying mathematical equation based on the lower and the upper bound is presented as follows:

$$\text{Lower Bound } Q_1 - [2.2 \times (Q_3 - Q_1)] \tag{1}$$
$$\text{Upper Bound } Q_3 + [2.2 \times (Q_3 - Q_1)] \tag{2}$$

At 0.05 degree of freedom, any data lower than Q_1 or greater than Q_3 was considered an outlier and need to be removed before analysis (Levi *et al*., 2009).

2.1.2. Data fitness using reliability analysis
Reliability analysis of the data was done to ascertain the fitness of the data for the selected analysis. Descriptive analysis of the reliability test was based on the data scale (measured in terms of weight and order of distribution). The summary statistics was done to compute the data means, variance, covariance and correlations using the intra class correlation coefficient.

2.1.3. Test of homogeneity
Homogeneity test was carried out to establish the fact that the data used (i.e. gas pressure, ambient temperature, CTD and power generation) for the analysis were from the same power plant (same population). Homogeneity test is based on the cumulative deviation from the mean as expressed using the mathematical equation below (Raes *et al*., 2006).

$$S_k = \sum_{i=1}^{k}\left(X_i - \bar{X} \right) \quad k = 1, \dots n \tag{3}$$

where, X_i = the record for the series X_1 X_2, ... X_n,

$\bar{X}$ = the mean,

S_{ks} = the residual mass curve.

For a homogeneous record, one may expect that the S_{ks} fluctuate around the zero-centre line in the residual mass curve since there is no systematic pattern in the deviation X_i's from the average values $\bar{X}$. To perform the homogeneity test, a software package for analysing time series data known as Rainbow (Raes *et al*., 2006) was used.

2.2 Assessment of normality

In the study the Jarque-Bera (JB) test for normality was employed because the sample size was large (i.e. >1000). Mathematically, the JB test is defined (Bowman and Shenton, 1975) as follows:

$$\text{JB} = n[(\sqrt{b_1})^2 / 6 + (b_2 - 3)^2 / 24] \tag{4}$$

where, n = sample size,
$\sqrt{b_1}$ = sample skewness, and
b_2 = kurtosis coefficient.

The hypothesis for the JB test is:
H_0 = Data follows a normally distribution
H_1 = Data do not follow a normal distribution

In general, a large JB value indicates that the residuals are not normally distributed. A value of JB greater than 10 means that the null hypothesis has been rejected at the 5% significance level. In other words, the data do not come from a normal distribution. JB value of between (0-10) indicates that the data is normally distributed (Das and Imon, 2016).

2.3 Diagnostic analysis of data

Diagnostic statistics were conducted to verify the statistical properties of the overall regression model. The selected diagnostic statistics include:
  i.   Heteroskedasticity test using Breusch-Pagan Godfrey
  ii. Serial Correlation test using Breusch Godfrey
  iii. Variance Inflation Factor (VIF)

3.0. Results and Discussion

3.1 Data quality assessment results

The results of outlier detection test; data fitness test using reliability analysis; and test of homogeneity are discussed below.

3.1.1 Data fitness test result using the labelling rule
Results of the computed percentiles for both the dependent and independent variable are presented in Table 1.

Table 1: Computed percentile for both dependent and independent variables

		Percentiles						
		5	10	25	50	75	90	95
Weighted Average (Definition 1)	Gas Pressure (FPG2)	20.300	20.300	20.600	21.400	21.500	21.500	21.700
	Ambient Temp. (INLET)	25.300	25.900	27.000	28.200	29.300	30.700	31.600
	Compressed Temp. Diff. (CTD)	328.400	331.400	335.650	340.100	350.750	359.700	364.530
	Power Generation (MWATTS)	76.200	81.500	86.400	91.900	101.150	107.700	109.430
Tukey's Hinges	Gas Pressure (FPG2)			20.600	21.400	21.500		
	Ambient Temp. (INLET)			27.000	28.200	29.300		
	Compressed Temp. Diff. (CTD)			335.700	340.100	350.700		
	Power Generation (MWATTS)			86.400	91.900	101.100		

Using the weighted average shown in Table 1, the 25th percentile (Q_1) for gas pressure was observed to be 20.600 while the 75th percentile (Q_3) was observed to be 21.500. Substituting into Eqn. (1) and Eqn. (2), the lower and upper bound statistics were computed to be 18.62 and 23.58 respectively.

Table 2 shows the highest gas pressure extreme value statistic to be 33.2 and 29.4. These values are higher than the calculated upper bound value of 23.58. Likewise the lowest gas pressure are seen to be 19.4 and 19.5, these are greater than the calculated lower bound of 18.62. It was observed that gas pressure values indicated by case number 587, 87, 226, 327 and 351 contained values that are greater than the calculated upper bound, hence, they were declared outliers and were removed (Levi *et al.*, 2009).

From Table 1, the 25th percentile (Q_1) for ambient temperature was observed to be 27.000 while the 75th percentile (Q_3) was observed to be 29.300. Using eqns. (1) and (2) the lower and upper bound statistics were computed as 21.94 and 34.36 respectively. Table 3 shows the extreme value statistics of ambient temperature for the highest and lowest ambient temperature. The highest ambient temperature values are 344.3, 330.5 and 185.8, these are higher than the calculated upper bound of 34.36. The lowest ambient temperature values are 21.0, 21.5, 22.3, 22.8 and 23.5 some of which were greater than the calculated lower bound of 21.94. In this case, it was observed that ambient temperature values indicated by case number 242, 542, 151, 470 and 1113 contained values that are greater than the calculated upper bound, and case number 669 and 739 contained values that are lower than the calculated lower bound, hence, they are considered as outliers and removed.

Table 2: Extreme value statistics for gas pressure

			Case Number	Value
Gas Pressure (FPG2)	Highest	1	587	33.2
		2	87	29.4
		3	226	29.4
		4	327	29.4
		5	351	29.4ᵃ
	Lowest	1	954	19.4
		2	264	19.4
		3	496	19.5
		4	486	19.5
		5	455	19.5ᵇ

Table 3: Extreme value statistics for ambient temperature

Ambient Temp. (INLET)	Highest	1	242	344.3
		2	542	344.3
		3	151	330.5
		4	470	330.5
		5	1113	185.8
	Lowest	1	669	21.0
		2	739	21.5
		3	652	22.3
		4	44	22.8
		5	631	23.5

The extreme value statistics of CTD is shown in Table 4. The 25th percentile (Q_1) for CTD was seen to be 335.650 while the 75th percentile (Q_3) was seen to be 350.750. The lower and upper bound statistics were computed as 302.43 and 383.97 using eqns. (1) and (2) respectively. From Table 4, the highest CTD values were 625.1, 582.2 and 470.6. These values are higher than the calculated upper bound of 383.97. The lowest CTD values were observed to be 21.4, 97.5, 100.7 and 123.4. These are lower than the calculated lower bound of 302.43. It was observed that CTD values indicated by case number 300, 986, 511, 378, 1055 contained values that are greater than the calculated upper bound and case number 949, 259, 2, 21 and 15 contained values that are lower than the calculated lower bound hence, they were declared outliers and were removed.

Table 5 shows the extreme value statistics of power generation. From Table 1, the 25th percentile (Q_1) for power generation was observed to be 86.40 while the 75th percentile (Q_3) was observed to be 101.15. Using Eqns. (1) and (2), the lower and upper bound statistics were computed as 53.95 and 133.60 respectively. From Table 5, the highest power generation values are 161.7, 159.6, 153.8, 152.5 and 138.9. These are higher than the calculated upper bound of 133.60. The lowest power generation values are 21.4 and 28.7. These values are lower than the calculated lower bound of 53.95. It was observed that power generation values indicated by case numbers 622, 658, 46, 604 and 275 contained power generation values that are greater than the calculated upper bound. Also, case numbers 949, 259, 948, 257 and 196 contained power generation values that are lower than the calculated lower bound, hence, removed as outliers.

Table 4: Extreme value statistics for CTD

Compressed Temp. Diff. (CTD)	Highest	1		300	625.1
		2		986	625.1
		3		511	582.2
		4		378	470.6
		5		1055	470.6
	Lowest	1		949	21.4
		2		259	21.4
		3		2	97.5
		4		21	100.7
		5		15	123.4

Table 5: Extreme value statistics for power generation

Power Generation (MWATTS)	Highest	1		622	161.7
		2		658	159.6
		3		46	153.8
		4		604	152.5
		5		275	138.9
	Lowest	1		949	21.4
		2		259	21.4
		3		948	28.7
		4		257	28.7
		5		196	28.7ᶜ

3.1.2. Results of data reliability analysis

Table 6 shows the summary statistics of data means, variance, covariance and correlations using the intra-class correlation coefficient.

Table 6: Result of summary item statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	121.436	21.250	341.688	320.438	16.079	2.260E4	4
Item Variances	349.277	1.054	886.707	885.653	840.947	1.501E5	4
Inter-Item Covariances	15.937	-14.564	113.148	127.711	-7.769	2.101E3	4
Inter-Item Correlations	.068	-.063	.311	.374	-4.952	.016	4

The two-way mixed model with confidence interval of 95% (i.e. p-value of 0.05) and initial test value of zero was used. The reliability hypothesis was as follows:

H_0: Data are reliable
H_1: Data are not reliable

The Fisher's probability test (F-test) was used for the analysis and result obtained are presented in Table 7.

Table 7: Inter-item correlation and covariance statistics

	Gas Pressure (FPG2)	Ambient Temp. (INLET)	Compressed Temp. Diff. (CTD)	Power Generation (MWATTS)
Gas Pressure (FPG2)	1.000	.011	.112	.051
Ambient Temp. (INLET)	.011	1.000	-.013	-.063
Compressed Temp. Diff. (CTD)	.112	-.013	1.000	.311
Power Generation (MWATTS)	.051	-.063	.311	1.000

Inter-Item Covariance Matrix

	Gas Pressure (FPG2)	Ambient Temp. (INLET)	Compressed Temp. Diff. (CTD)	Power Generation (MWATTS)
Gas Pressure (FPG2)	1.054	.220	3.435	.639
Ambient Temp. (INLET)	.220	360.238	-7.257	-14.564
Compressed Temp. Diff. (CTD)	3.435	-7.257	886.707	113.148
Power Generation (MWATTS)	.639	-14.564	113.148	149.109

Table 7 shows that the gas pressure and compressed temperature difference were positively correlated with power generation and have a covariance value of 0.639 and 113.148 respectively. The large covariance value of compressed temperature difference indicates that the variable has an overriding influence on power generation compare to gas pressure and ambient temperature. These are seen to be negatively correlated with power generation. The computed coefficient of correlations of 0.051 for gas pressure, -0.063 for ambient temperature, and 0.311 for compressed temperature difference were observed to be relatively weak, which is indicative of the absence of co-linearity problem in the regression variables. The highest coefficient of (+0.311) which is between compressed temperature

difference and power generation still did not pose any challenge of multi-collinearity. Hence, we can conclude that there is no issue of multicollinearity and that the regression variables are clearly correlated with the dependent variable. This is evident in the intra-class correlation coefficient presented in Table 8.

Table 8: Computed intra-class correlation coefficients

	Intraclass Correlation[a]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.046[b]	.021	.072	1.191	1152	3456	.000
Average Measures	.161[c]	.079	.237	1.191	1152	3456	.000

Again, we observed from the result of Table 8 that the single and average measure intra-class correlation coefficients are relatively weak (0.046 and 0.161) which is indicative of the absence of multicollinearity. To ascertain the reliability of the data, one-way analysis of variance (ANOVA) was generated and presented in Table 9. At 0.05 degree of freedom (df), with a computed p-value of 0.000 as observed in Table 9, the null hypothesis was accepted and it was concluded that the data are good and can be employed for further analysis.

Table 9: Analysis of variance table

		Sum of Squares	df	Mean Square	F	Sig
Between People		457445.219	1152	397.088		
Within People	Between Items	7.816E7	3	2.605E7	7.816E4	.000
	Residual	1152024.163	3456	333.340		
	Total	7.931E7	3459	22928.352		
Total		7.977E7	4611	17299.201		

3.1.3. Homogeneity test results

Raes *et al*. (2006) described homogeneity test as one based on the cumulative deviation from the mean. The homogeneity test hypothesis of gas pressure is defined as:

H_0: Data are statistically homogeneous.
H_1: Data are not homogeneous.

The null and alternate hypotheses were tested at 90%, 95% and 99% confidence interval (i.e. 0.1, 0.05 and 0.01) df as shown in Figure 1.
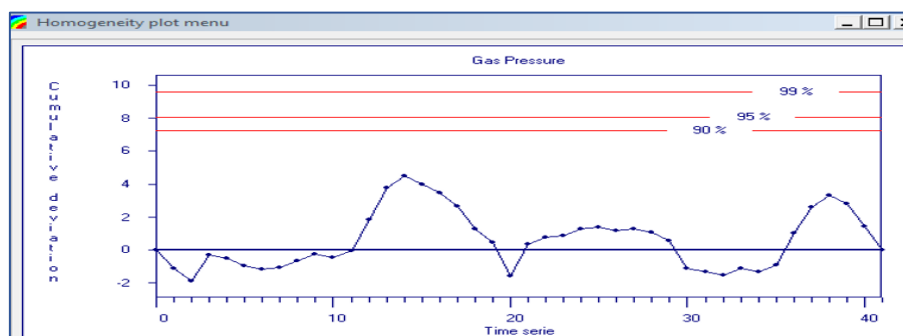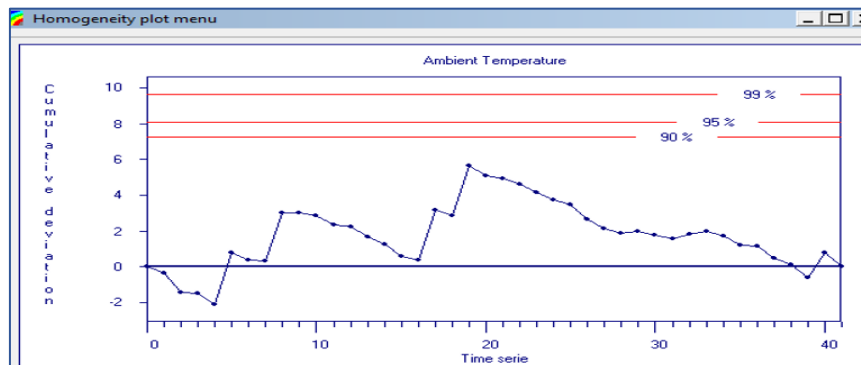

Figure 1: Homogeneity test of gas pressure data

The gas presure data is seen to fluctuated around the zero-center line of the residual mass curve in Figure 1, an indication that the data were statistically homogeneous. A further test of homogeneity was done using the homogeneity statistics to check the strength of the null hypothesis over the alternate hypothesis. Based on the result obtained, the null hypothesis (H_0) was accepted, and we concluded that the gas pressure data were statistically homogeneous at 90%, 95% and 99% confidence interval.

The homogeneity test hypothesis for ambient temperature is as follows:

H_0: Data are statistically homogeneous.
H_1: Data are not homogeneous.

The null and alternate hypotheses were tested at 90%, 95% and 99% confidence interval (i.e. 0.1, 0.05 and 0.01) df as shown in Figure 2.

Figure 2: Homogeneity test of ambient temperature data

The ambient temperature data (Figure 2) fluctuated around the zero-center line of the residual mass curve, an indication that the data were statistically homogeneous. The homogeneity statistics was used to check the strength of the null hypothesis over the alternate hypothesis. Based on the result obtained, the null hypothesis (H_0) was accepted, and it was concluded that the ambient temperature data were statistically homogeneous at 90%, 95% and 99% confidence interval.

The hypothesis of homogeneity test of compressed temperature difference data is:

H_0: Data are statistically homogeneous.
H_1: Data are not homogeneous.

The null and alternate hypothesis were tested at 90%, 95% and 99% confidence interval (i.e. 0.1, 0.05 and 0.01) df as shown in Figure 3.
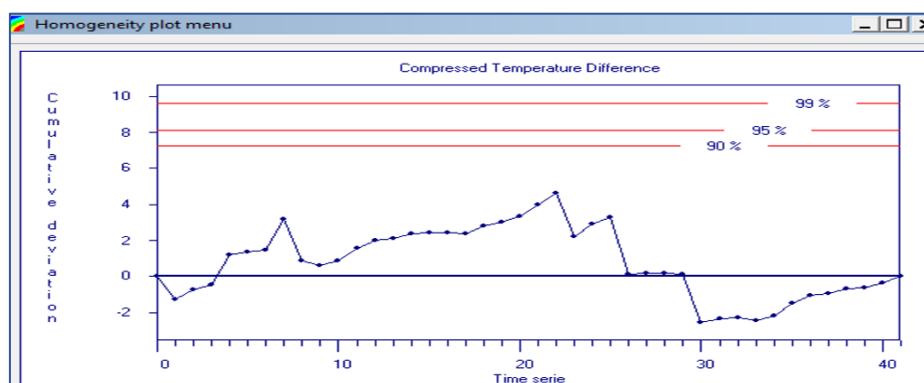

Figure 3: Homogeneity test of compresses temperature difference data

From Figure 3, the compressed temperature difference data fluctuated around the zero-center line of the residual mass curve, an indication that the data were statistically homogeneous. The homogeneity statistics was used to check the strength of the null hypothesis over the alternate hypothesis. Based on the result obtained, the null hypothesis (H_0) was accepted, and it was concluded that the compressed temperature difference data were statistically homogeneous at 90%, 95% and 99% confidence interval.

The hypothesis of homogeneity test of power generation data is:

H_0: Data are statistically homogeneous.
H_1: Data are not homogeneous.

The null and alternate hypothesis were tested at 90%, 95% and 99% confidence interval (i.e. 0.1, 0.05 and 0.01) df as shown in Figure 4.
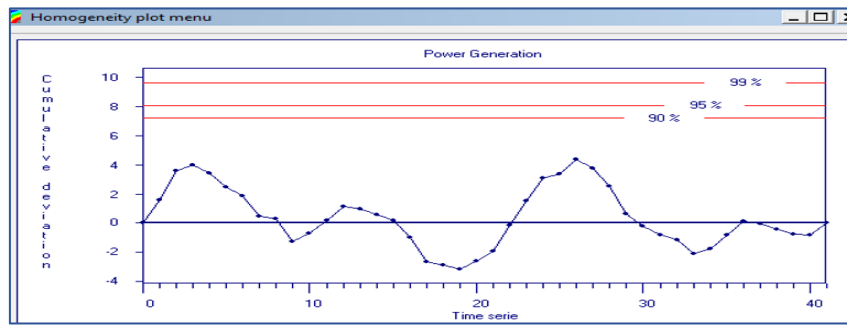
Figure 4: Homogeneity test of power generation data

Figure 4 showed that the power generation data fluctuated around the zero-center line of the residual mass curve, an indication that the data are statistically homogeneous. The homogeneity statistics was used to check the strength of the null hypothesis over the alternate hypothesis. Based on the result obtained, the null hypothesis (H_0) was accepted, and it was concluded that the power generation data were statistically homogeneous at 90%, 95% and 99% confidence interval.

3.2. Normality test results

The normality test was done for the one independent and three dependent variables using the JB test for normality statistical software. Figure 5 shows the results of the normality test of gas pressure.
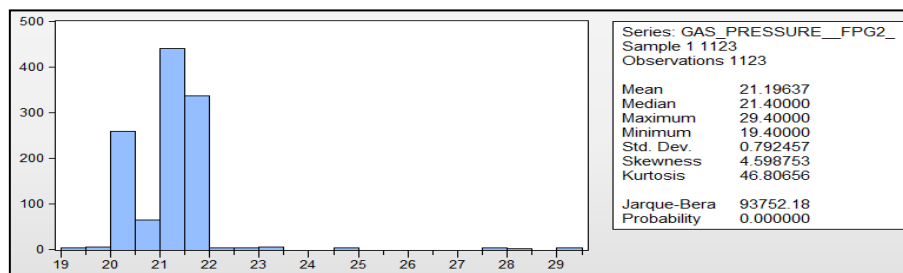

Figure 5: Normality test of gas pressure data

A skewness coefficient of 4.598753 and kurtosis value of 46.80656 shown in Figure 5 indicate that the gas pressure data is not normally distributed. For normality, the skewness coefficient should not be greater than 1 and the kurtosis should not be greater than 3 (Bai and Ng, 2005). JB value of 93752.18 and a probability (p-value) of 0.00% observed in Figure 5 indicates that the gas pressure data is not normally distributed. JB value >10 means that the null hypothesis is rejected at that level of significance (Das and Imon, 2016), meaning, the data did not come from a normal distribution. Since the JB test value is greater than 10 and the (p-value) is less than the 5% significant value, the null hypothesis was rejected and it was concluded that the data is not from a normal distribution.

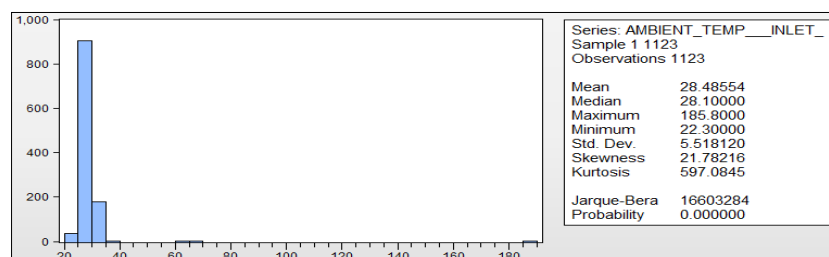The normality test result of ambient temperature is shown in Figure 6.

Figure 6: Normality test of ambient temperature data

A skewness coefficient of 21.78216 and a kurtosis value of 597.0845 observed in Figure 6 indicate that the ambient temperature data is not normally distributed. JB value of 16603284 and a probability

(p-value) of 0.00% observed in Figure 6 indicates that the ambient temperature data is not normally distributed. Since the JB test value is greater than 10 and the p-value is less than the 5% significant value, the null hypothesis was rejected and it was concluded that the data is not from a normal distribution.

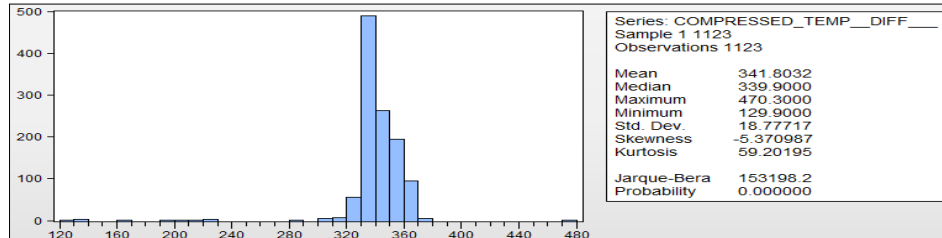Figure 7 shows result of the normality test of compressed temperature difference.

Figure 7: Normality test of compressed temperature difference data

A skewness coefficient of -5.370987 shows that the data is negatively skewed an indication that the data is not normally distributed. Kurtosis value of 59.20195 observed in Figure 7 is also an indication that the data is not from a normal population distribution. JB value of 153198.2 and a probability (p-value) of 0.00% also indicated that the compressed temperature difference data was not normally distributed. Since the JB test value is greater than 10 and the p-value is less than the 5% significant value, the null hypothesis was rejected and it was concluded that the data is not from a normal distribution.

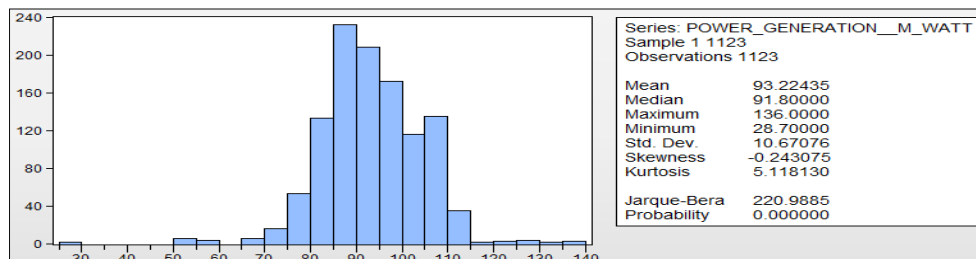Figure 8 shows result of the normality test of power generation data.

Figure 8: Normality test of power generation data

A skewness coefficient of -0.243075 shows that the data is negatively skewed an indication that the data is not normally distributed. Kurtosis value of 5.118130 observed in Figure 8 is also an indication that the data is not from a normal population distribution. JB value of 220.9885 and a probability (p-value) of 0.00% as observed in Figure 8 also indicates that the power generation data is not normally distributed Since the JB test value is greater than 10 and the p-value is less than the 5% significant value, the null hypothesis was rejected and it was concluded that the data is not from a normal distribution.

3.3. Results of the Diagnostic Analysis of Data

The diagnostic statistical analyses done in this study include: Heteroskedasticity test using Breusch-Pagan Godfrey; Serial Correlation test using Breusch Godfrey; and Variance Inflation Factor (VIF).

3.3.1. Heteroskedasticity test
Result of heteroskedasticity test using Breusch-Pagan Godfrey method showed that (i) the calculated (p-value) based on the F-statistics is 0.0000; (ii) the calculated p-value based on Lagrange multiplier (LM) is 0.0000. Since the computed p-value based on F-statistics and Lagrange multiplier is less than 0.05 (P < 0.05), we rejected the null hypothesis of homoskedasticity and conclude that there is no heteroskedasticity in the data (Astivia and Zumbo, 2019).

3.3.2. Serial correlation test result
The result of serial correlation LM test using Breusch Godfrey method indicated that (i) the calculated p-value based on the F-statistics is 0.0000; and (ii) the calculated p-value based on LM is 0.0000.

Since the computed p-value based on F-statistics and LM is less than 0.05 (P < 0.05), we rejected the null hypothesis of serial correlation and concluded that there is the presence of serial correlation in the data.

3.3.3. The Variance Inflation Factor (VIF) result

The result of the calculated VIF for the selected variables was observed to be less than 10. Since the computed variance inflation factors (centred VIF) for the selected independent variables were less than 10, it was concluded that the variables were well correlated with the dependent variable, hence absence of multicollinearity (Montgomery, 2005). The Output of regression analysis is presented in Table 10. Finally, the reliance of the dependent variable on the selected independent variables was evaluated using the coded least square regression equation as shown in Eqn. (5).

$$(M/watts)\,C\,FPG2\,INLET\,CTD \qquad (5)$$

Table 10: Output of regression analysis

View	Proc	Object		Print	Name	Freeze		Estimate	Forecast	Stats	Resids

Dependent Variable: POWER_GENERATION__M_WATT
Method: Least Squares
Date: 01/01/20   Time: 14:43
Sample: 1 1123
Included observations: 1123

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	29.61506	9.196908	3.220111	0.0013
GAS_PRESSURE__FPG2_	0.197497	0.377239	0.523532	0.6007
AMBIENT_TEMP___INLET_	-0.247932	0.053707	-4.616362	0.0000
COMPRESSED_TEMP__DIFF___	0.194514	0.015927	12.21311	0.0000

R-squared	0.137577	Mean dependent var	93.22435
Adjusted R-squared	0.135265	S.D. dependent var	10.67076
S.E. of regression	9.922863	Akaike info criterion	7.431115
Sum squared resid	110180.3	Schwarz criterion	7.449010
Log likelihood	-4168.571	Hannan-Quinn criter.	7.437878
F-statistic	59.50232	Durbin-Watson stat	0.928874
Prob(F-statistic)	0.000000		

From the result of Table 10, with a regression p-value of 0.0013, it was concluded that the regression analysis was significant at 0.05 df. The Independent variables, namely; (INLET and CTD) were observed to have a very strong influence on the dependent variable when compared to FPG2. The poor regression terms such as coefficient of determination and adjusted coefficient of determination was apparently due to the presence of serial correlation in the used data. The overall regression equation was then generated as:

$$(M/watts) = 21.61506 + 0.197497(FGP2) - 0.247932(INLET) + 0.194514(CTD) \qquad (6)$$

3.4. Model development

Table 11 shows the estimated regression parameters of eqn. (2).

Table 11: Estimated Regression Parameters

S/N	Regression parameters	Estimated values
1	Coefficient of determination (R- Squared)	0.137577
2	Adjusted Coefficient of determination (Adjusted R-Squared)	0.135265
3	Sum of Error of Regression (S.E. of Regression)	9.922863
4	Residual Sum of Square	110180.3

From Table 11, the estimated values of regression parameters were very poor. An indication that the exact relationship between the dependent variable (power generation) and the selected independent variables (gas pressure, ambient temperature and compressed temperature difference) cannot be determined using linear regression model. Hence, non-linear model equations method (power function; quantile regression; robust regression; inverse function based on gamma distribution; and log function based on negative binomial distribution) were developed to determine the exact relationship between the dependent variable and the selected independent variables. Statistical software was used to develop the model equations.

3.4.1. Power function model

The model equation developed using the power function model is shown in Table 12.

Table 12: Result of power function model

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.196	3	.732	57.436	.000^a
	Residual	14.262	1119	.013		
	Total	16.458	1122			

a. Predictors: (Constant), Compressed Temp. Diff. (CTD), Ambient Temp. (INLET), Gas Pressure (FPG2)

b. Dependent Variable: Power Generation (M/WATTS)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.101	.399		7.763	.000
	Gas Pressure (FPG2)	.107	.097	.031	1.095	.274
	Ambient Temp. (INLET)	-.339	.034	-.278	-9.980	.000
	Compressed Temp. Diff. (CTD)	.383	.049	.222	7.883	.000

a. Dependent Variable: Power Generation (M/WATTS)

Results of the power function method as shown in Table 2 using ANOVA showed that the power function model is significant at 0.05df. Using the unstandardized coefficients, the power function equation was developed as:

$$(M/watts) = 22.22 \times (FPG2)^{0.107}(INLET)^{-0.339}(CTD)^{0.383} \tag{7}$$

3.4.2. Quantile regression method

The ANOVA result shown in Table 12 showed that the quantile regression model is significant at 0.05df. Using the unstandardized coefficients, the quantile regression equation is:

$$(M/watts) = -106.4318 + 0.096263(FPG2) - 1.560389(INLET) + 0.703719(CTD) \tag{8}$$

3.4.3. Robust regression method

Table 12 shows the ANOVA results which indicated that the robust regression model is significant at 0.05df. Using the unstandardized coefficients, the robust regression equation was developed as:

$$(M/watts) = -117.8962 + 0.123039(FPG2) - 1.776864(INLET) + 0.752693(CTD) \tag{9}$$

3.4.4. Inverse function method

From the ANOVA results in Table 12, the inverse function model is significant at 0.05df. Using the unstandardized coefficients, the inverse function equation developed is:

$$Z = 0.019253 - 2.48E(-05)(FPG2) + 3.86E(-05)(INLET) - 2.66E(-05)(CTD) \tag{10}$$
$$(M/watts) = {}^1/_Z \tag{11}$$

3.4.5. Log function method

The ANOVA analysis in Table 12 indicated that the log function model is significant at 0.05df. Using the unstandardized coefficients, the log function equation was developed as:

$$Z = 3.963366 + 0.002036(FPG2) - 0.002265(INLET) + 0.001733(CTD) \tag{12}$$
$$(M/watts) = Exp(Z) \tag{13}$$

The summary of the developed mathematical models are shown in Table 13, and can be used to predict the power generation.

Table 13: Summary of developed mathematical equations (models)

S/N	Method	Developed mathematical equation
1	Linear Regression	$(M/watts) = 21.61506 + 0.197497(FGP2) - 0.247932(INLET) + 0.194514(CTD)$
2	Robust Regression	$(M/watts) = -117.8962 + 0.123039(FPG2) - 1.776864(INLET) + 0.752693(CTD)$
3	Quantile Regression	$(M/watts) = -106.4318 + 0.096263(FPG2) - 1.560389(INLET) + 0.703719(CTD)$
4	Power Function	$(M/watts) = 22.22 \times (FPG2)^{0.107}(INLET)^{-0.339}(CTD)^{0.383}$
5	Inverse Function	$Z = 0.019253 - 2.48E(-05)(FPG2) + 3.86E(-05)(INLET) - 2.66E(-05)(CTD)$ $(M/watts) = 1/Z$
6	Log Function	$Z = 3.963366 + 0.002036(FPG2) - 0.002265(INLET) + 0.001733(CTD)$ $(M/watts) = Exp(Z)$

4.0. Conclusions

The relationship between the output of gas power plant and the quantitative variables were not linearly related, but could be best described by a non-linear regression model. The study provided a veritable and laudable process to systematically identify factors that are capable of influencing the generation of power in gas power plants. One dependent (power generation) and three independent variables (Gas Pressure, Ambient Temperature, Compressed Temperature) were used for this analysis. The three critical numeric variables were observed to play a key role in assessing the relationship between input and output parameters in gas power plant. The study also showed that the ambient temperature and compressed temperature difference had very strong influence on the dependent variable when compared to the gas pressure variable.

References

Astivia, O. L. O. and Zumbo, B. D. (2019). Heteroskedasticity in Multiple Regression Analysis: What it is, How to detect it and How to solve it with Application in R and SPSS. *Practical Assessment, Research, and Evaluation*, 24(1), pp. 1 – 17.

Bai, J. and Ng, S. (2005). Tests for Skewness, Kurtosis, and Normality for Time Series Data. *American Statistical Association Journal of Business & Economic Statistics*, 23(1), pp. 49 – 60.

Bowman, K. O. and Shenton, B. R. (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2. *Biometrika*, 64, pp. 243- 450.

Das, K. R. and Imon, A. H. M. R. (2016). A Brief Review of Tests for Normality. *American Journal of Theoretical and Applied Statistics*, 5(1), pp. 5 - 12.

Enete, C. I. and Alabi, M. O. (2011). Potential Impacts of Global Climate Change on Power and Energy Generation. *Journal of Knowledge Management, Economics and Information Technology*, 6, pp. 1-14.

Iwuamadi, O. C., and Dike, D. O. (2012). Empirical Analysis of Productivity of Nigerian Power Sector. *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)*, 3(4), pp. 24 – 38.

Izuwan, N. C. (2017). Improving Natural Gas Distribution and Management in Nigeria. *International Journal of Scientific and Engineering Research*, 8(7), pp. 330 - 344.

Levi, D. B., Julie, E. K., Olsen, J. R., Pulwarty, R. S., Raff, D.A., Turnipseed, D. P., Webb, R. S and Kathleen D. W. (2009). Climate Change and Water Resources Management: A Federal Perspective. *Circular*, 1331, pp. 1-72.

Luis, R., Bolonio, D., Mazadiego, L. F. and Valencia-Chapi, R. (2019). Long-Term Electricity Supply and Demand Forecast (2018–2040): A LEAP Model Application towards a Sustainable Power Generation System in Ecuador. *Sustainability*, 11(5316), pp. 2-19.

Montgomery, D. C. (2005). *Design and Analysis of Experiments*. 6th Edition, New York, John Wiley and Sons, Inc. USA.

Nwokeji, G. U. (2007). The Nigerian National Petroleum Cooperation and the Development of the Nigerian Oil and Gas Industry: History, Strategies and Current Directions. Baker Institute and Japan Petroleum Energy Centre.

Onochie, U. P., Obanor, A. and Aliu, S. A. (2015). Electricity Crisis in Nigeria: The Way Forward. *American Journal of Renewable and Sustainable Energy*, 1(4), pp. 180-186.

Oricha, J.Y. and Olarinoye, G.A. (2012). Analysis of Interrelated Factors Affecting Efficiency and Stability of Power Supply in Nigeria. *International Journal of Energy Engineering*, 2(1), pp. 1 – 8.

Orogun, B. O. (2015). Natural Gas to Power in Nigeria, the Practices and the Way Forward for Sustainable Development. *Sustainable Development and National Resources*, University of Manitoba.

Oyedepo, S. O. (2012). Efficient Energy Utilization as a tool for Sustainable Development in Nigeria. *International Journal of Energy and Environmental Engineering*, 3(11), pp. 1-12.

Oyedepo, S.O., Fagbenle, R. O. and Adefila, S. S. (2015). *Assessment of Performance Indices of Selected Gas Turbine Power Plants in Nigeria*. Energy Science & Engineering, Society of Chemical Industry and John Wiley & Sons Ltd., U. S. A.

Raes, D., Willens, P. and Gbaguidi (2006). Rainbow-A Software Package for Analyzing Data and Testing the Homogeneity of Historical Data Sets, 1, pp. 1-15.

Rapu, C. S., Adenuga, A. O., Kanya, W. J., Abeng, M. O., Golit, P. D., Hilili, M. J., et al. (2015). Analysis of Energy Market Conditions in Nigeria. *Occasional Paper,* No. 55, Central Bank of Nigeria, Garki, Abuja, Nigeria.

Sambo, A. S., Garba, B., Zarma, I. H. and Gaji, M. M. (2010). Electricity Generation and the Present Challenges in the Nigerian Power Sector. Energy Commission of Nigeria, Abuja-Nigeria.