

# **Applying Regression Model to Predict the Effects of Shaft and Bearing Deviations on the Clearance of the Assembly of Mating Parts**

Olorunnishola, A.A.G

*Department of Mechanical Engineering Technology, Federal Polytechnic, Ado-Ekiti*

---

**Abstract:** *This work presents formulation of linear regression model for the effect of shaft and bearing deviations on the clearance of the assembly of mating parts. The model validation confirmed the existence of statistical relationships between clearance and shaft and bearing deviations. Applying data collected from direct measurement of shafts and bearings produced in the workshop,  $R^2$  values of 99.7% was obtained for assembly of mating parts model ; thus, indicating that about 99.7% of the parts produced are interchangeable, thus reducing the probability of producing scraps to the barest minimum. The results obtained using critical parameters of assembly of mating parts and multiple linear regression models give good estimation and are validated with experimental values with an error much lesser than 4% error. The regression model also showed that given the shaft and bearing deviations, the expected clearance can be determined, thus, enabling the quality control personnel to significantly monitor and reduce variation in the assembly of mating parts.*

**Key words:** *Mating parts, Regression modeling, Quality control, Interchangeability*

---

## **I. Introduction**

According to Dale [3], quality control is the procedure of measuring the quality of output, mainly of manufactured goods against a specified standard. Quality control is used to prevent rather than to detect defectives. Stephen [12] says quality control is composite characteristics of engineering and manufacture that determine the degree to which it satisfies the need or use for which it is intended. In other words, statistical quality control is conformance to specifications and it is a simple statistical method which examining the extent to which the quality of the manufactured goods is satisfied without necessarily checking every item produced. Quality control involves setting quality standards, appraising conformance to these standards, acting when the standards are exceeded, and planning for the improvement in the standard [10]. In the manufacturing of goods, reliability (the ability to function consistently and effectively across time, appropriateness of configuration, fits and finish parts) and interchangeability (in case of replacement of parts), always come to mind. To achieve the aim of this paper, two questions were raised; (1) “how do we obtain quality control to the assembly of mating parts before it is released to the markets?” and (2) “what does statistical application by collection of data such as deviations, clearances, e.t.c., of the same component (mating parts) have to do with assuming it?” [5].

Several investigators have carried out studies and developed mathematical models concerning the problem of choosing optimum input parameters [7]. Regression techniques have a long history of use as forecasting tools in multiple disciplines. Regression models are also used for decision making [2]. Also a researcher has developed practical crash prediction regression models for assessing the long-range safety impact of alternative freeway networks for urban areas [13].

Therefore, the focus of this paper is to present a regression model that predicts the effect of shaft and bearing deviations on the clearance of the assembly of mating parts. This work will provide input data for reducing variation in assembling of mating parts through identifying, validating, and controlling statistically significant control factors (i.e. shaft and bearing deviations) that can influence clearance in the assembly of mating parts. This work will try to establish a significant relationship between the clearance and shaft and bearing deviations, and then formulate model to predict the effect of shaft and bearing deviations on the clearance of the assembly of mating parts.

## **II. Methodology**

It is assumed that relationships do exist between clearance and deviations in mating parts sizes. The construction of statistical models like linear regression can serve as a tool to verify or otherwise refute the presence of relationships between interacting variables [1]. Therefore, the technical approach adopted consists of three major steps. Firstly, linear regression model was formulated to represent anticipated relationships between clearance and shaft and bearing deviations. Secondly, this model was tested for validity and adequacy using statistical tools such as a hypothesis testing, Analysis of Variance (*ANOVA*), and Coefficient of determination,  $R^2$ .

Statistical Package of Social Sciences [11] was employed to determine the proportion of variation in the dependent variable  $R^2$ ; explained by the regression model. The values range from 0 to 1; the small value indicates that the model does not fit the data well. In Analysis of Variance (ANOVA), the sum of squares, degrees of freedom, and mean square are computed for two sources of variation, regression and residual. A model with a large regression sum of squares in comparison to the residual sum of squares indicates that the model accounts for most of variation in the dependent variable. The  $F$ -statistic is the regression mean square divided by the residual mean square and if the significance of the  $F$ -statistic is small (smaller than 0.05) then the independent variables performs a good work explaining the variation in the dependent variable.  $T$  statistic helps to determine the relative importance of each variable in the model.  $T$  values below  $-2$  or above  $+2$  give the favored impression. Regression Co-linearity Diagnostics helps to determine if there are any problems with co-linearity. Co-linearity is the undesirable situation where the correlations among the independent variables are strong. An important factor in co-linearity diagnostics is; condition indices. A condition index greater than 15 indicates a possible problem and an index greater than 30 suggests a serious problem with multi-co-linearity [4].

### 2.1 Variables Selection

The relationship of shaft deviation (SD) and bearing deviation (BD), both independent variables with clearance between the mating parts (CL), a dependent variable, is to be derived. The specific definitions and units of measurements of these variables are defined as follows; Shaft deviation (SD) represents the difference between the shaft size and the mean value of shaft measures in mm. Bearing deviation (BD) measures the difference between the bearing size and the mean value of bearing given in units of mm; clearance between the mating parts represents the difference between the bearing size and shaft size (CL) measures in mm.

The above factors are selected as control variables influencing clearance between the mating parts based on;

- i) The presence of physical or logical influence of these factors on the clearance of the mating parts. For example, as shaft deviation increases, clearance between the shaft and the bearing would increase or decrease as a result of variation in the assembly of mating parts [6].
- ii) It is predicted that as bearing sizes deviate from the specified value, it is logical to conclude that the clearance between the mating parts would be affected accordingly [9].

To justify the presence of such informative relationships between these factors, scatter diagram is used to clearly indicate the validity of initial selection of control variables (see Figure 1).

### 2.2 Models assumptions

The following assumptions are made:

- i) There was a linear relationship between the clearance and the related control variables (application of scatter diagrams).
- ii) That multi co-linearity was not present between the control variables (shaft and bearing deviations).
- iii) That the random errors (o) are independent and normally distributed with constant variance and zero mean.

## III. Formulation of Linear Regression Model (LRM)

### 3.1 Model formulation.

Based on the selected variables and model assumptions, the following linear regression model was formulated for the quality control of the assembly of mating parts.

**Model:** Assembly of mating parts.

$$Exp(CL / SD, BD) = b_o + b_1SD + b_2BD \quad \dots \text{Eqn. 1}$$

Where:

$Exp(CL / SD, BD)$  - is the expected value of clearance in mm given the independent variables SD and BD,  $b_o$  is intercept of model,  $b_1$  is regression coefficient associated with shaft deviation and  $b_2$  is regression coefficient associated with bearing deviation.

### 3.2 Statistical Tests on Regression Model Parameters

#### 3.2.1 Hypothesis 1: Testing model validity

Model I hypothesis for the assembly of mating parts is presented in equation. 2:

$$H_0 : b_j = 0, \quad j = 1, 2 \quad \dots \text{Eqn. 2}$$

If  $H_0$  is rejected, then  $H_1 : \text{at least one } b_j \neq 0$

The hypothesis is intended to test validity of the presence of a relation between clearance and any of the independent variables. If the null hypothesis is rejected, then there are some independent variables that do actually affect clearance of the assembled mating parts.

**3.2.2 Hypothesis II: Individual testing of coefficients of the linear regression model.**

**Hypothesis IIa** for Shaft deviation (SD) is as presented in equation 3.

$$H_0 : b_1 = 0 \text{ vs } H_1 : b_1 \neq 0 \quad \dots \text{Eqn. 3}$$

The null hypothesis assumed that there was no statistically significant relationship between the clearance of the assembled mating parts and shaft deviation (SD).

**Hypothesis IIb** for bearing deviation (BD) is as presented in equation.4:

$$H_0 : b_2 = 0 \text{ vs } H_1 : b_2 \neq 0 \quad \dots \text{Eqn. 4}$$

The null hypothesis assumed that there was no statistically significant relationship between the clearance of the assembled mating parts and bearing deviation (BD).

**IV. Data Collection**

The experimental data used in this work was adopted from Olorunnishola, et al [10] as presented in Table 1.

Table 1: Experimental data obtained by direct measurement of parts produced

Shaft size (mm) X	Shaft deviation (SD) (mm) $x - \bar{x}$	Bearing size (mm) X	Bearing deviation (BD) (mm) $x - \bar{x}$	Clearance of the assemblies (CL) (mm)
30.02	-.01	30.03	-.01	.01
30.02	-.01	30.04	.00	.02
30.03	.00	30.02	-.02	-.01
30.04	.01	30.02	-.02	-.02
30.03	.00	30.04	.00	.01
30.02	-.01	30.02	-.02	.00
30.02	-.01	30.08	.04	.06
30.03	.00	30.04	.00	.01
30.03	.00	30.05	.01	.02
30.02	-.01	30.04	.00	.02
30.03	.00	30.06	.02	.03
30.04	.01	30.03	-.01	-.01
30.02	-.01	30.02	-.02	.00
30.03	.00	30.04	.00	.01
30.01	-.02	30.06	.02	.05
30.02	-.01	30.05	.01	.03
30.01	-.02	30.03	-.01	.02
30.04	.01	30.02	-.02	-.02
30.03	.00	30.03	-.01	.00
30.00	-.03	30.05	.01	.05
30.03	.00	30.03	-.01	.00
30.00	-.03	30.05	.01	.05
30.01	-.02	30.02	-.02	.01
30.02	-.01	30.02	-.02	.00
30.02	-.01	30.02	-.02	.00
30.03	-.01	30.03	-.01	.00
30.03	.00	30.04	.00	.01
30.02	-.01	30.06	.02	.04
30.01	-.02	30.07	.03	.06
30.02	-.01	30.03	-.01	.01
30.05	.02	30.02	-.02	-.03
30.01	-.02	30.09	.05	.08
30.02	-.01	30.05	.01	.03
30.01	-.02	30.03	-.01	.02
30.05	.02	30.10	.06	.05
30.02	-.01	30.04	.00	.02
30.04	.01	30.02	-.02	-.02
30.00	-.03	30.04	.00	.04
30.04	.01	30.05	.01	.01
30.02	-.01	30.05	.01	.03
30.05	.02	30.06	.02	.01
30.04	.01	30.02	-.02	-.02
30.03	.00	30.02	-.02	-.01
30.01	-.02	30.02	-.02	.01
30.03	.00	30.06	.02	.03
30.02	-.01	30.04	.00	.02

30.04	.01	30.03	-.01	-.01
30.01	-.02	30.05	.01	.04
30.04	.01	30.07	.03	.03
30.03	.00	30.08	.04	.05

**V. Models Validation and Analysis**

SPSS (version 16.0) was used to validate the obtained data and the results are shown in Table 2.

Table 2: Summary of Regression Model for the assembly of mating parts.

Parameter	ANOVA			COLLINEARITY DIAGNOSTICS				RESIDUALS			
	Value	Parameter	Sum of squares	Parameter	Condition index	Coefficients	VIF	T-Statistic	Parameter	Mean ( $\mu$ )	Standard Deviation ( $\sigma$ )
R <sup>2</sup>	0.997	Regression	0.029	Constant (b <sub>0</sub> )	1.00	0.010	-	44.390	Standard Predicted value	0	1
F-Statistic	6972.661	Residual	0.000	SD (b <sub>1</sub> )	1.203	-0.994	1.001	62.012	Standard Residual	0	0.979
Significance of F-statistic	0.000	-	-	BD (b <sub>2</sub> )	1.526	1.006	1.001	98.650	-	-	-

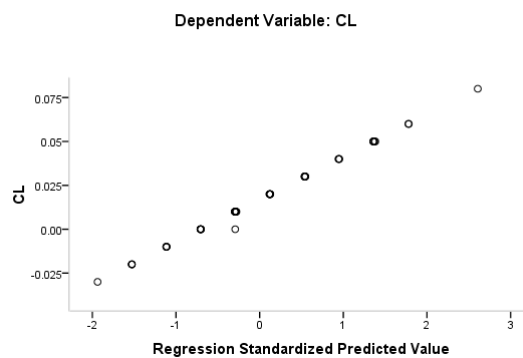


Figure 1: Scatter plot for mating parts assemblies

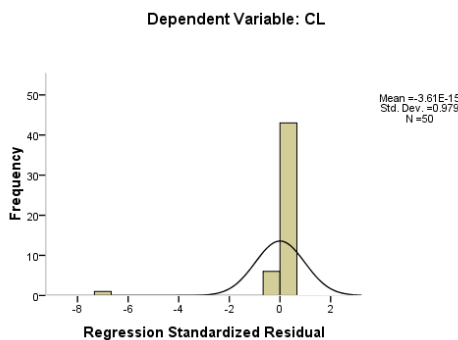


Figure 2: Histogram for mating parts assemblies

Scatter diagram shown in Fig.1 was plotted which clearly indicates the validity of initial selection of variables. The model summary shown in Table 2, gave a computed value for the  $R^2$  as 0.997, thus indicating that the regression was “strong” as about 99.7% of the variation in the clearance of the assemblies could be accounted for by the control variables. The ANOVA analysis in the regression result, shown in Table 2, gave a computed value for the F-statistic as 6972.661 while the corresponding table value of 3.40 at 0.05 level of significance (q) and (2,49) degrees of freedom showed that the multiple linear regression model was significant and valid. Also large regression sum of squares (0.029) in comparison to the residual sum of squares (0.00) indicated that the model accounts for most of variation in the dependent variable. The coefficients  $b_0$ ,  $b_1$ ,  $b_2$  shown in Table 2 are 0.010, -0.994, and 1.006, respectively; and the results of the T- test indicated that regression coefficients  $b_1$  and  $b_2$  were statistically significant and not equal to zero (as given by hypothesis ii) at 0.025 level of significance

and 49 degrees of freedom (table  $T$ -value= $T_{0.025, 49} = 2.064$ ). Therefore, the regression equation of clearance of the assemblies in mm can be given by equation 5. It should be noted that the assumptions made were valid for this model with respect to multi co-linearity and residuals' distribution, as seen from Table 2, the condition indexes values of 1.203 and 1.526 for SD and BD respectively. From Table 2 and Figure 2, the residuals' average was zero with standard deviation (0.979) of approximately 1.0 implying that residuals were actually independent. The variance inflation factor  $VIF$  of 1.001 indicated that multi co-linearity was not a problem in this application (i.e.  $VIF < 4$ ) Neave, [8], which clearly demonstrated that shaft deviation and bearing deviation were not significantly interacting factors. Finally, as it might appear to be a misleading result, the coefficient of shaft deviation was negative, implying that as shaft size increases, clearance of the assemblies (i.e. variation in the assembly of mating parts) decreases. For steady clearance of the assemblies, increasing the size of shaft would necessarily decrease variation in the assembly of mating parts, and thus, a negative coefficient in the clearance of the assemblies' equation.

$$Exp(CL / SD, BD) = 0.010 - 0.994 SD + 1.006 BD \quad \dots \text{Eqn. 5}$$

### VI. Simulation

A computer program based on the model equation 5 was developed by using JavaScript programming language to simulate the expected value of clearance of the assemblies as a function of shaft and bearing deviations. The simulated results are presented in Tables 3 below.

Table 3: Comparison of the experimental values of clearance and simulated clearance values

Shaft Deviation (SD) mm	Bearing Deviation (BD) mm	Clearance of the assemblies (CL) (mm) experimental	Clearance of the assemblies (CL) (mm) simulated	Percentage variations
-0.01	-0.01	.01	0.00988	1.2
-0.01	0.00	.02	0.01994	0.3
0.00	-0.02	-.01	-0.01012	1.2
0.01	-0.02	-.02	-0.02006	0.3
0.00	0.00	.01	0.01	0.0
-0.01	-0.02	.00	-0.00018	0.0
-0.01	0.04	.06	0.06018	-0.3
0.00	0.00	.01	0.01	0.0
0.00	0.01	.02	0.02006	-0.3
-0.01	0.00	.02	0.01994	0.3
0.00	0.02	.03	0.03012	-0.4
0.01	-0.01	-.01	-0.01	0.0
-0.01	-0.02	.00	-0.00018	0.0
0.00	0.00	.01	0.01	0.0
-0.02	0.02	.05	0.05	0.0
-0.01	0.01	.03	0.03	0.0
-0.02	-0.01	.02	0.01982	0.9
0.01	-0.02	-.02	-0.02006	0.3
0.00	-0.01	.00	-0.00006	0.0
-0.03	0.01	.05	0.04988	0.24
0.00	-0.01	.00	-0.00006	0.0
-0.03	0.01	.05	0.04988	0.24
-0.02	-0.02	.01	0.00976	2.4
-0.01	-0.02	.00	-0.00018	0.0
-0.01	-0.02	.00	-0.00018	0.0
-0.01	-0.01	.00	0.00988	0.0
0.00	0.00	.01	0.01	0.0
-0.01	0.02	.04	0.04006	-0.15
-0.02	0.03	.06	0.06006	-0.1
-0.01	-0.01	.01	0.00988	1.2
0.02	-0.02	-.03	-0.03	0.0
-0.02	0.05	.08	0.08018	-0.225
-0.01	0.01	.03	0.03	0.0
-0.02	-0.01	.02	0.01982	0.9
0.02	0.06	.05	0.05048	-0.96
-0.01	0.00	.02	0.01994	0.3
0.01	-0.02	-.02	-0.02006	0.3
-0.03	0.00	.04	0.03982	0.45
0.01	0.01	.01	0.01012	-1.2
-0.01	0.01	.03	0.03	0.0
0.02	0.02	.01	0.01024	-2.4
0.01	-0.02	-.02	-0.02006	0.3
0.00	-0.02	-.01	-0.01012	1.2
-0.02	-0.02	.01	0.00976	2.4
0.00	0.02	.03	0.03012	-0.4

-0.01	0.00	.02	0.01994	0.3
0.01	-0.01	-0.01	-0.01	0.0
-0.02	0.01	.04	0.03994	0.15
0.01	0.03	.03	0.03024	-0.8
0.00	0.04	.05	0.05024	-0.48

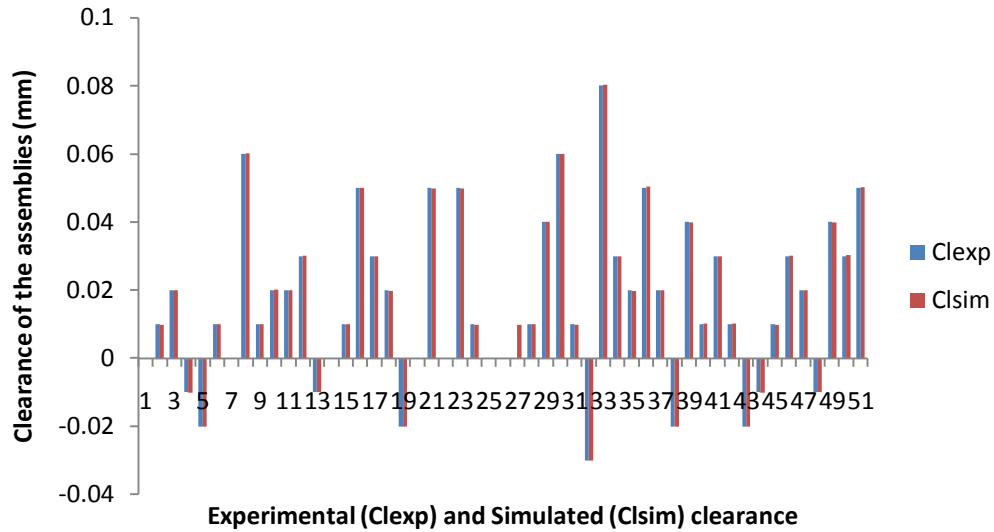


Figure 3: The relationship between the experimental and simulated clearance values

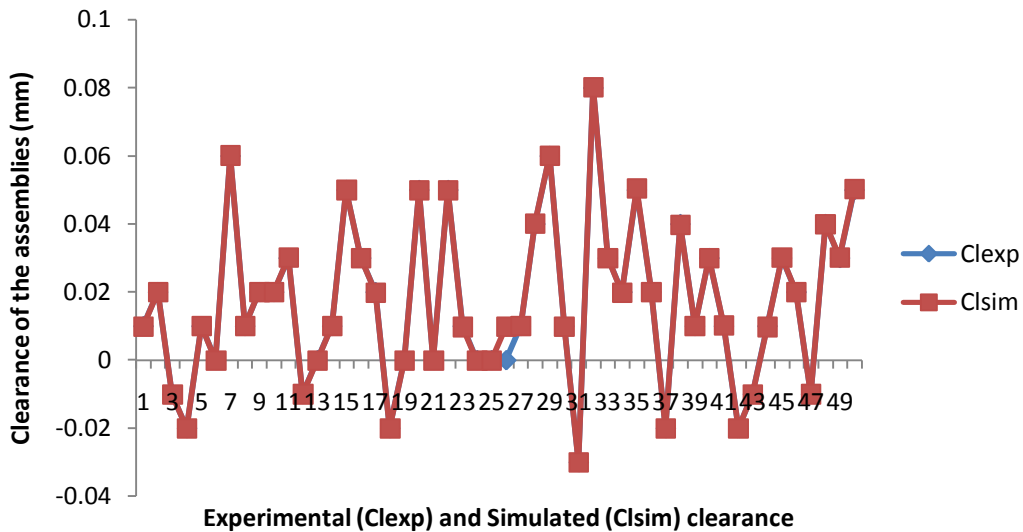


Figure 4: Plot of experimental clearance versus simulated clearance values

Table 3 is representing the experimental value of clearance and the estimated value by the regression model. The percentage variations shown in Table 3 show that the regression model has approximately same results as was obtained from the experimentations. It was also observed that, the critical parameters Shaft and Bearing deviations are the main factors which affect the clearance of the assemblies. Figures 3 and 4 representing the relationship between the experimental and regression modeled clearance of the assemblies. The figures show that the model gives error less than 2%.

### VII. Conclusion

The regression model developed in this paper can effectively predict the expected values of clearance based on shaft and bearing deviations. The results obtained using critical parameters of assembly of mating parts and multiple linear regression models give good estimation and are validated with experimental values with an error much lesser than 4% error. Results show that this technique easily captures the intricate relationship between various process parameters and can be readily integrated into existing manufacturing environment [7]. The results further established a significant relationship between clearance and shaft and bearing deviations of

the assembly of mating parts, thus, call for additional efforts for integrating this model (method for ensuring the conformance of mating parts to specifications) inside the production facility such as quality control and production management. The model will enable quality control personnel to significantly monitor and reduce variation in the assembly of mating parts. In addition the model will ensure that above 90% of the parts produced are interchangeable, thus reducing the probability of producing scraps to the barest minimum.

### References

- [1] Amjed, A. (2004). "A statistical Approach Linking Energy Management to Maintenance and Production Factors". *Journal of quality in maintenance engineering*, Vol. 9, No.1, P 25.
- [2] Andrew, G., Matt, S., and Valerie, C. (2003). "Regression Modeling and Meta-Analysis for Decision Making: A Cost-Benefit Analysis of Incentives in Telephone Surveys", 1, Pp 1-13.
- [3] Dale, H.B. (1986). **Quality Control**. Prentice-Hall International Edition, U.S.A. Pp 1-10.
- [4] Egbo, G. and Olorunnishola, A.A.G (2007). Analysis of the effects of equipment failure and production rates on electrical energy consumption in Ashaka Cement Plc. Gombe State. *Journal of Pure and Applied Sciences*, Vol. 10, No.1.
- [5] Harris, P.J. (1981). **Manufacturing Technology 3**. London, Butter Worth and Co. (Publisher) Ltd.
- [6] Harry, F. and Steven, C.A. (1994). **Statistical Concepts and Applications**. New York, Cambridge University Press. Pp 39-58.
- [7] Kumar, P. and Singh, R. (2012). Neural Networks and Regression Modeling of Eco-friendly Melting Furnace Parameters using Bio-fuels. [Electronic version]. *International Journal of Computer Applications* 43. (1), 10-12.
- [8] Neave, H. R. (1978), **Statistics Tables for Mathematicians, Engineers, Economists and the Behavioural and Management Sciences**. George Allen and Unwin Publishers Ltd. London, Pp 41-62.
- [9] Olorunnishola, A.A.G and Egbo, G. (2007). "Models to Predict the Effect of Failure and Production Rates on Electrical Energy Consumption in Ashaka Cement Plc. Gombe State". *Journal of Pure and Applied Sciences*, Vol. 10, No.1, Pp 77-83.
- [10] Olorunnishola, A.A.G, Datau, S.G. and Anbua, E.E. (2010). "Application of Statistical Quality Control to the Assembly of Mating Parts". *Yankari Journal*, Vol. 6, Pp 41-45.
- [11] SPSS, (2006), "Statistical Package of Social Sciences Manual Version 16.0", [www.spss.com](http://www.spss.com).
- [12] Stephen, B. V., (1993), **Statistics for Engineering Problem Solving**. Iowa state University, Pirs Publishing Company, Boston, Pp. 5-10.
- [13] Vasin, K., Arun, C., Joseph, E.H., and Mary, S.Y. (2008). "Planning Level Regression Models for Crash Prediction on Interchange and Non-Interchange Segments of Urban Freeways". *Journal of Transportation Engineering*, 134 (3).