ISSN 2350-1049

International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS)
Vol. 2, Issue 3, pp: (45-54), Month: July 2015 - September 2015, Available at: www.paperpublications.org

# Application of Weighted Least Squares Regression in Forecasting

Sulaimon Mutiu O.

Department of Statistics & Mathematics Moshood Abiola Polytechnic, Abeokuta, Ogun State, Nigeria

***Abstract:*** **This work models the loss of properties from fire outbreak in Ogun State using Simple Weighted Least Square Regression. The study covers (secondary) data on fire outbreak and monetary value of properties loss across the twenty (20) Local Government Areas of Ogun state for the year 2010. Data collected were analyzed electronically using SPSS 21.0. Results from the analysis reveal that there is a very strong positive relationship between the number of fire outbreak and the loss of properties; this relationship is significant. Fire outbreak exerts significant influence on loss of properties and it accounts for approximately 91.2% of the loss of properties in the state.**

***Keywords:*** **Fire Outbreak, Heteroscedasticity, Properties.**

## I. INTRODUCTION

The method of least squares is a standard approach in regression analysis to the approximate solution of overdetermined systems, i.e., sets of equations in which there are more equations than unknowns. "Least squares" means that the overall solution minimizes the sum of the squares of the errors made in the results of every single equation.

The most important application is in data fitting. The best fit in the least-squares sense minimizes the sum of squared residuals, a residual being the difference between an observed value and the fitted value provided by a model. When the problem has substantial uncertainties in the independent variable (the $x$ variable), then simple regression and least squares methods have problems; in such cases, the methodology required for fitting errors-in-variables models may be considered instead of that for least squares.

Least squares problems fall into two categories: linear or ordinary least squares and non-linear least squares, depending on whether or not the residuals are linear in all unknowns. The linear least-squares problem occurs in statistical regression analysis; it has a closed-form solution. The non-linear problem is usually solved by iterative refinement; at each iteration the system is approximated by a linear one, and thus the core calculation is similar in both cases.

One of the common assumptions underlying most process modelling methods, including linear and nonlinear least squares regression, is that each data point provides equally precise information about the deterministic part of the total process variation. In other words, it is assumed that the standard deviation of the error term is constant over all values of the predictor or explanatory variables. This assumption, however, clearly does not hold, even approximately, in every modelling application.

The usual linear regression model, $y = a + bx$ assumes that all the random error components are identically and independently distributed with constant variance. When this assumption is violated, then ordinary least squares estimator of regression coefficient looses its property of minimum variance in the class of linear and unbiased estimators. The violation of such assumption can arise in anyone of the following situations:

1. The variance of random error components is not constant.

2. The random error components are not independent.

3. The random error components do not have constant variance as well as they are not independent.

In such cases, the covariance matrix of random error components does not remain in the form of an identity matrix but can be considered as any positive definite matrix. Under such assumption, the ordinary least square estimate (OLSE) does not remain efficient as in the case of identity covariance matrix. The generalized or weighted least squares method is used in such situations to estimate the parameters of the model.

Unlike linear and nonlinear least squares regression, weighted least squares regression is not associated with a particular type of function used to describe the relationship between the process variables. Instead, weighted least squares reflects the behaviour of the random errors in the model; and it can be used with functions that are either linear or nonlinear in the parameters. It works by incorporating extra nonnegative constants, or weights, associated with each data point, into the fitting criterion. The size of the weight indicates the precision of the information contained in the associated observation. Optimizing the weighted fitting criterion to find the parameter estimates allows the weights to determine the contribution of each observation to the final parameter estimates. It is important to note that the weight for each observation is given relative to the weights of the other observations; so different sets of absolute weights can have identical effects.

Each term in the weighted least squares criterion includes an additional weight, that determines how much each observation in the data set influences the final parameter estimates and it can be used with functions that are either linear or nonlinear in the parameters.

In a weighted fit, less weight is given to the less precise measurements and more weight to more precise measurements when estimating the unknown parameters in the model. Using weights that are inversely proportional to the variance at each level of the explanatory variables yields the most precise parameter estimates possible. Weighting the sum of the squares of the differences may significantly improve the ability of the least square regression to fit the linear model to the data. Weighted least square is an efficient method that makes good use of small data set. It also shares the ability to provide different types of easily interpretable statistical intervals for estimation, prediction, calibration and optimization.

It is to these effects that this study was undertaken to examine the application of the weighted least square regression in forecasting.

## II. STATEMENT OF THE PROBLEM

Weighted least square (WLS) regression is useful for estimating the values of model parameters when the response values have differing degrees of variability over the combinations of the predictor values. The problem of this study is modelling the loss of properties from fire outbreak by Simple Weighted Least Square Regression.

## III. SCOPE OF THE STUDY

This study covers data on fire outbreak and monetary value of properties loss across the twenty (20) Local Government Areas of Ogun state, Nigeria for the year 2010. The year 2010 data being the current data available through the Ogun State Statistical Year Book, 2011. The data collected is secondary in nature.

## IV. LITERATURE REVIEW

Ever since the seminal publications of Legendre (1805) and Gauss (1809), the method of least squares (LS) has been a main tool or approach of modern statistical analysis (Celmins, 1998; Kalman, 1960; Plackett, 1949; Plackett, 1950; Seal, 1967; Sprott, 1978; Stigler, 1981; Young, 1974).

In a regression problem with time series data (where the variables have subscript "$t$" denoting the time the variable was observed), it is common for the error terms to be correlated across time, but with a constant variance; this is the problem of "autocorrelated disturbances". For regressions with cross-section data (where the subscript "$i$" now denotes a particular individual or firm at a point in time), it is usually safe to assume the errors are uncorrelated, but often their variances are not constant across individuals. This is known as the problem of heteroscedasticity (for "unequal scatter"); the usual assumption of constant error variance is referred to as homoscedasticity. Although the mean of the dependent variable might be a linear function of the regressor, the variance of the error terms might also depend on the same regressor, so that the observations might "fan out" in a scatter diagram (Econometrics Laboratory 1999). This inherently assumes a given degree of heteroscedasticity in which the variance of $y$ increases proportionately with increased values of $x$.

In ordinary least squares, the estimated coefficients provide the regression equation that minimizes SSE = $\sum e_i^2$.

In weighted least squares (WLS), the estimated equation minimizes $\sum w_i e_i^2$ where $w_i$ is a weight given to the $i^{th}$ observation. The object is to minimize the sum of the squares of the random factors of the estimated residuals. If the weights are all the same constant, then we have ordinary least squares (OLS) regression. However, if the structure of the data suggests unequal weights are appropriate, then it would be inappropriate to ignore the regression weights. With one regressor, usually the regression weights are functions of that regressor (James R. Knaub, 2012).

**Uses of Weighted Least Squares (WLS):**

1. Weighted least squares can be used in the presence of non-constant variance. If the variance of the $i^{th}$ observation is $\sigma_i^2$, then weights $w_i = \frac{1}{\sigma_i^2}$ give theoretically correct results for standard errors of coefficients and the various significance tests. With $w_i = \frac{1}{\sigma_i^2}$ notice that $\sum w_i\, e_i^2 = \sum \frac{1}{\sigma_i^2} e_i^2 = \sum \left(\frac{e_i}{\sigma_i}\right)^2$. So, we're minimizing the sum of *squared standardized errors* when weights = 1/variance.

2. Weighted least squares is often used as the basis for doing "robust" regression in which outliers are given less weight than points that aren't outliers.

**Identifying the weights:**

The principal difficulty in practice is determining values for the weights. If we see non-constant variance in a plot of residuals versus fits, then we might consider using weighted least squares.

Most commonly, the pattern of non-constant variance is that either the standard deviation or the variance of the residuals is linearly related to the mean (the fits). This occurs theoretically in most skewed distributions, for instance.

The absolute residuals essentially are estimates of standard deviation. So we might plot absolute residuals versus fits. If this looks linear, we could fit a regression line (response = absolute residuals, predictor = fits) to the pattern. The predicted values from this regression could be viewed as smoothed estimates of the standard deviations of the points. So, our weights in a weighted least squares regression would be $w_i = \frac{1}{(\hat{S}_i)^2}$. Note that the "predicted" standard deviations would have to be squared in the weight function.

The squared residuals essentially are estimates of variances. We might plot squared residuals versus fits. If this looks linear, we could fit a regression line (response = squared residuals, predictor = fits) to determine smoothed estimates of the variance. Denote these estimates as $\hat{V}_i^{\,2}$ . Then, in a weighted regression, use weights $w_i = \frac{1}{\hat{V}_i^{\,2}}$.

**The nature of Weighted Least Squares (WLS) Regression:**

"All horses are animals, but not all animals are horses." (Socrates) - Analogously, all ordinary least squares (OLS) regressions are weighted least squares (WLS) regressions, but not all WLS regressions are OLS. That is, OLS regression is a special case of WLS regression. Many may use OLS as a default, and in some applications that might be good enough, but just because we do not know what weights are appropriate, it does not mean that one avoids assigning weights by using OLS, because we are *de facto* claiming that the weights are equal. That, in fact, is a very decisive assignment of weights. For establishment surveys, that is not a good assumption. See Brewer (2002). When we use regression through the origin, the strong weight assumption implicit in OLS regression is likely to be a highly faulty assumption.

An example of this confusion, from NIST (2009), a generally very nice handbook, is as follows: "The biggest disadvantage of weighted least squares, which many people are not aware of, is probably the fact that the theory behind this method is based on the assumption that the weights are known exactly. … It is important to remain aware of this potential problem, and to only use weighted least squares when the weights can be estimated precisely relative to one

ISSN 2350-1049
International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS)
Vol. 2, Issue 3, pp: (45-54), Month: July 2015 - September 2015, Available at: www.paperpublications.org

another." This is very misleading because it says that if one cannot estimate regression weights "precisely relative to one another" then one should always *assume* that they are *equal*. This may sometimes be true enough, but not for regression through the origin. Some information regarding the regression weights should be gleaned in any case. To use OLS actually *does* assume one knows the regression weights "precisely relative to one another."

Weights can be estimated from the data. (See Knaub(1993, 1997), Carroll and Ruppert(1988), Sweet and Sigman(1995), Steel and Fay(1995), and Brewer(1963), for example.) If the estimated weights cause an increase in estimated variance, it is not OK to pick another weight solely to lower the variance estimate. Such an estimate would not be justified unless there was a functional reason for it.

The object is to give less weight to the more uncertain data points, and those are generally the largest, but data near the origin can have disproportionately large measurement error in many practical situations. A good reason for using cutoff sampling is to avoid collecting data that are not reliable. Often with design-based sampling, the smallest observations are imputed by some model since they are either non-respondents or their responses do not 'pass the laugh test' (badly fail reasonable edits). However, from Knaub (2008), one may find a modified weight to be better for reasons of robustness.

**Benefits and Disadvantages of WLS:**

Weighted least squares is an efficient method that makes good use of small data sets. It also shares the ability to provide different types of easily interpretable statistical intervals for estimation, prediction, calibration and optimization. The main advantage that weighted least squares enjoys over other methods is the ability to handle regression situations in which the data points are of varying quality.

The biggest disadvantage of weighted least squares, is probably the fact that the theory behind this method is based on the assumption that the weights are known exactly. The exact weights are almost never known in real applications, so estimated weights must be used instead. The effect of using estimated weights is difficult to assess, but experience indicates that small variations in the weights due to estimation do not often affect a regression analysis or its interpretation. When the weights are estimated from small numbers of replicated observations, the results of an analysis can be very badly and unpredictably affected. This is especially likely to be the case when the weights for extreme values of the predictor or explanatory variables are estimated using only a few observations. It is important to remain aware of this potential problem, and to only use weighted least squares when the weights can be estimated precisely relative to one another.

Weighted least squares regression, is also sensitive to the effects of outliers. If potential outliers are not investigated and dealt with appropriately, they will likely have a negative impact on the parameter estimation and other aspects of a weighted least squares analysis.

If a weighted least squares regression actually increases the influence of an outlier, the results of the analysis may be far inferior to an unweighted least squares analysis

**Derivation of coefficient for single regressor for WLS:**

*The Ordinary Least Square (OLS)*

In the simple linear regression model

$$y_i = \alpha + \beta x_i + e_i \qquad \text{\_\_\_\_\_(i)}$$

We compute the residuals, $e_i$ as

$$e_i = y_i - \alpha - \beta x_i \qquad \text{\_\_\_\_\_(ii)}$$

$$e_i = y_i - (\alpha + \beta x_i)$$

Or $$e_i = y_i - \hat{y}_i \qquad \text{\_\_\_\_\_(iii)}$$

Where $\hat{y}_i$ is the predicted value of $y_i$

Square of the residuals give

$$e_i^2 = (y_i - \alpha - \beta x_i)^2$$

ISSN 2350-1049

International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS)
Vol. 2, Issue 3, pp: (45-54), Month: July 2015 - September 2015, Available at: www.paperpublications.org

The *Ordinary Least Square function* gives

$$\sum e_i^2 = \sum(y_i - \alpha - \beta x_i)^2 \qquad \text{\_\_\_\_\_(iv)}$$

[*Sum of square of the residuals*]

For easy computation, let $Q$ be represented by the sum of square of the residuals, so that

$$Q = \sum e_i^2$$

$$Q = \sum(y_i - \alpha - \beta x_i)^2 \qquad \text{\_\_\_\_\_(v)}$$

We want to minimize $Q$ with respect to $\alpha$ and $\beta$. The least square estimate ($a$ and $b$) of $\alpha$ and $\beta$ is obtained by differentiating $Q$ and equate the derivative to zero (0).

Thus $$\frac{dQ}{d\alpha} = -2\sum(y_i - a - bx_i) = 0 \qquad \text{\_\_\_\_\_(vi)}$$

Or $$-2\sum(y_i - a - bx_i) = 0$$

$$\sum(y_i - a - bx_i) = 0$$

$$\sum y_i - \sum a - \sum bx_i = 0$$

$$\sum y_i - \sum a - b\sum x_i = 0$$

$$\sum y_i - na - b\sum x_i = 0$$

$$na + b\sum x_i = \sum y_i$$

$$na = \sum y_i - b\sum x_i$$

$$a = \frac{\sum y_i}{n} - b\frac{\sum x_i}{n}$$

$$\boldsymbol{a = \bar{y} - b\bar{x}} \qquad \text{\_\_\_\_\_(vii)}$$

Also $$\frac{dQ}{d\beta} = -2\sum(y_i - a - bx_i)x_i = 0 \qquad \text{\_\_\_\_\_(viii)}$$

Or $$-2\sum(x_i y_i - ax_i - bx_i^2) = 0$$

$$\sum(x_i y_i - ax_i - bx_i^2) = 0$$

$$\sum x_i y_i - \sum ax_i - \sum bx_i^2 = 0$$

$$\sum x_i y_i - \sum ax_i - b\sum x_i^2 = 0$$

$$\sum ax_i + b\sum x_i^2 = \sum x_i y_i$$

$$b\sum x_i^2 = \sum x_i y_i - \sum ax_i$$

$$b\sum x_i^2 = \sum x_i y_i - \sum ax_i$$

$$b\sum x_i^2 = \sum x_i y_i - a\sum x_i \qquad \text{\_\_\_\_\_(ix)}$$

On substituting (vi) into (viii) we have

$$b\sum x_i^2 = \sum x_i y_i - (\bar{y} - b\bar{x})\sum x_i$$

$$b\sum x_i^2 = \sum x_i y_i - \bar{y}\sum x_i + \bar{x}b\sum x_i$$

$$b\sum x_i^2 - \bar{x}b\sum x_i = \sum x_i y_i - \bar{y}\sum x_i$$

$$b(\sum x_i^2 - \bar{x}\sum x_i) = \sum x_i y_i - \bar{y}\sum x_i$$

$$\boldsymbol{b = \frac{\sum x_i y_i - \bar{y}\sum x_i}{\sum x_i^2 - \bar{x}\sum x_i}}$$

Or $$\boldsymbol{b} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \quad \text{\_\_\_\_\_(x)}$$

***The Weighted Least Square (WLS):***

In this method, the deviation between the observed and expected values of $y_i$ is multiplied by a weight $w_i$ where $w_i$ is chosen to be inversely proportional to the variance of $y_i$.

For simple linear regression model $y_i = \alpha + \beta x_i + e_i$:

The *Weighted Least Squares function* is

$$\sum w_i e_i^2 = \sum w_i (y_i - \alpha - \beta x_i)^2 \quad \text{\_\_\_\_\_(xi)}$$

For easy computation, let $L$ be represented by the sum of square of the residuals, so that $L = \sum w_i e_i^2$

$$\Longrightarrow \quad L = \sum w_i (y_i - \alpha - \beta x_i)^2 \quad \text{\_\_\_\_\_(xii)}$$

We want to minimize $L$ with respect to $\alpha$ and $\beta$. The least square estimate ($a$ and $b$) of $\alpha$ and $\beta$ is obtained by differentiating $L$ and equate the derivative to zero (0).

Thus $$\frac{dL}{d\alpha} = -2\sum w_i (y_i - a - bx_i) = 0 \quad \text{\_\_\_\_\_(xiii)}$$

Or $$-2\sum w_i (y_i - a - bx_i) = 0$$

$$\sum w_i (y_i - a - bx_i) = 0$$

$$\sum w_i y_i - \sum a w_i - \sum b w_i x_i = 0$$

$$\sum w_i y_i - a\sum w_i - b\sum w_i x_i = 0$$

$$a\sum w_i = \sum w_i y_i - b\sum w_i x_i$$

$$\boldsymbol{a} = \frac{\sum w_i y_i - b\sum w_i x_i}{\sum w_i} \quad \text{\_\_\_\_\_(xiv)}$$

Also $$\frac{dL}{d\beta} = -2\sum w_i (y_i - a - bx_i)x_i = 0 \quad \text{\_\_\_\_\_(xv)}$$

Or $$-2\sum w_i (x_i y_i - a x_i - b x_i^2) = 0$$

$$\sum w_i (x_i y_i - a x_i - b x_i^2) = 0$$

$$\sum w_i x_i y_i - \sum a w_i x_i - \sum b w_i x_i^2 = 0$$

$$\sum w_i x_i y_i - a\sum w_i x_i - b\sum w_i x_i^2 = 0$$

$$b\sum w_i x_i^2 = \sum w_i x_i y_i - a\sum w_i x_i$$

$$b = \frac{\sum w_i x_i y_i - a\sum w_i x_i}{\sum w_i x_i^2} \quad \text{\_\_\_\_\_(xvi)}$$

On substituting (xiv) into (xvi) we have

$$b = \frac{\sum w_i x_i y_i - \left[\frac{\sum w_i y_i - b\sum w_i x_i}{\sum w_i}\right]\sum w_i x_i}{\sum w_i x_i^2}$$

$$b = \frac{\sum w_i x_i y_i - \left[\frac{\sum w_i x_i \sum w_i y_i - b(\sum w_i x_i)^2}{\sum w_i}\right]}{\sum w_i x_i^2}$$

$$b = \frac{\frac{\sum w_i \sum w_i x_i y_i - \left[\sum w_i x_i \sum w_i y_i - b(\sum w_i x_i)^2\right]}{\sum w_i}}{\sum w_i x_i^2}$$

$$b = \frac{\sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i + b(\sum w_i x_i)^2}{\sum w_i \sum w_i x_i^2}$$

International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS)
Vol. 2, Issue 3, pp: (45-54), Month: July 2015 - September 2015, Available at: www.paperpublications.org

$$b\sum w_i \sum w_i x_i^2 \;=\; \sum w_i \sum w_i x_i y_i \;-\; \sum w_i x_i \sum w_i y_i + b(\sum w_i x_i)^2$$

$$b\sum w_i \sum w_i x_i^2 \;-\; b(\sum w_i x_i)^2 \;=\; \sum w_i \sum w_i x_i y_i \;-\; \sum w_i x_i \sum w_i y_i$$

$$b[\sum w_i \sum w_i x_i^2 \;-\; (\sum w_i x_i)^2] \;=\; \sum w_i \sum w_i x_i y_i \;-\; \sum w_i x_i \sum w_i y_i$$

$$\boldsymbol{b} \;=\; \frac{\sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i}{\sum w_i \sum w_i x_i^2 - (\sum w_i x_i)^2} \qquad \_\_\_\_\_(xvii)$$

**Coefficient for single regressor with zero intercept:**

*(Regrssion Through the Origin)*

When the linear regression model, $y_i = a + bx_i$ has zero intercept, then

$$\sum e_i^2 = \sum(y_i - \beta x_i)^2 \quad \text{and} \quad \boldsymbol{b} = \frac{\sum x_i y_i}{\sum x_i^2} \quad \text{(OLS)} \qquad \_\_\_\_(xviii)$$

Similarly $$\sum w_i e_i^2 = \sum w_i (y_i - \beta x_i)^2 \quad \text{and} \quad \boldsymbol{b} = \frac{\sum w_i x_i y_i}{\sum w_i x_i^2} \quad \text{(WLS)} \qquad \_\_\_\_(xix)$$

**Summary of results:**

For simple linear regression model:

$\boldsymbol{y_i = a + bx_i}$

| *Ordinary Least Square (OLS)* | *Weighted Least Square* |
|---|---|
| $b = \dfrac{n\sum x_i y_i \;-\; \sum x_i \sum y_i}{n\sum x_i^2 \;-\; (\sum x_i)^2}$ | $b \;=\; \dfrac{\sum w_i \sum w_i x_i y_i \;-\; \sum w_i x_i \sum w_i y_i}{\sum w_i \sum w_i x_i^2 \;-\; (\sum w_i x_i)^2}$ |
| $a = \bar{y} - b\bar{x}$ | $a = \dfrac{\sum w_i y_i \;-\; b\sum w_i x_i}{\sum w_i}$ |

If the weights, $w_i$, are all the same constant, then we have ordinary least squares (OLS) regression.

## V. MATERIALS & METHODS

**Research design:**

This research was designed to model the monetary loss of properties due to fire outbreak across the twenty (20) Local Government Areas of Ogun state using the Weighted Least Square Regression.

For the successful execution of this research work, secondary data on loss of properties and fire outbreak for the year 2010 was employed. This was extracted from the Ogun State Statistical Year Book, 2011. Data collected were analyzed electronically using SPSS 21.0.

**Techniques of data analysis:**

The data analysis techniques employed are the Weighted Least Square (WLS) Regression, Descriptive Statistics, Correlation (R) and Coefficient of Determination ($R^2$).

**Method of data analysis:**

In analyzing the data for WLS regression, the *Number of Fire Outbreak* per Local Government Area represents the predictor variable $(x)$ while the Loss of Properties represents the predicted variable $(y)$. The squared residuals from the OLS were used as estimates of variances in the weight $(w_i)$ function. Coefficients ($a$ and $b$) of the simple linear regression model, $y_i = a + bx_i$ are estimated using

$$b = \frac{\sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i}{\sum w_i \sum w_i x_i^2 - (\sum w_i x_i)^2} \qquad \text{and} \qquad a = \frac{\sum w_i y_i - b\sum w_i x_i}{\sum w_i}$$

International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS)
Vol. 2, Issue 3, pp: (45-54), Month: July 2015 - September 2015, Available at: www.paperpublications.org

## VI. RESULT

**TABLE 1:Descriptive Statistics[a]**

| | Mean | Std. Deviation | N |
|---|---|---|---|
| Loss of Properties | 546.7072 | 2.70165 | 20 |
| Fire Outbreak | 37.53 | .154 | 20 |

a. Weighted Least Squares Regression - Weighted by Weight

**TABLE 2:Correlations[a]**

| | | Loss of Properties | Fire Outbreak |
|---|---|---|---|
| Pearson Correlation | Loss of Properties | 1.000 | .955 |
| | Fire Outbreak | .955 | 1.000 |
| Sig. (1-tailed) | Loss of Properties | . | .000 |
| | Fire Outbreak | .000 | . |
| N | Loss of Properties | 20 | 20 |
| | Fire Outbreak | 20 | 20 |

a. Weighted Least Squares Regression - Weighted by Weight

**TABLE 3: Model Summary[b,c]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .955[a] | .912 | .908 | .82165 |

a. Predictors: (Constant), Fire Outbreak

b. Dependent Variable: Loss of Properties

c. Weighted Least Squares Regression - Weighted by Weight

**TABLE 4: Coefficients[a,b]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | -80.485 | 50.345 | | -1.599 | .127 |
| | Fire Outbreak | 16.713 | 1.221 | .955 | 13.690 | .000 |

a. Dependent Variable: Loss of Properties
b. Weighted Least Squares Regression - Weighted by Weight

The WLS regression model is thus: $y_i = -80.485 + 16.713x_i$

## VII. DISCUSSION OF RESULTS

**Descriptive Statistics:**

Table 1 indicates the mean number Fire Outbreak for the year across the twenty LGA of the State is approximately 38 with a standard deviation of 0.154. The mean Loss of properties for the year across the twenty LGA of the State is approximately ₦546,707,200.00 with a standard deviation of 2.70165.

**Correlations (R):**

Table 2 indicates that there is a very strong positive but imperfect relationship between the number of fire outbreak and the loss of properties, with a correlation value of 0.955. In addition, this correlation is indicated to be significant, with a ***sig.****value* of 0.000

**Coefficient of Determination ($R^2$):**

Table 3 indicates that approximately 91.2% of the variation in the loss of properties is being explained by fire outbreak, with a $R^2$ value of 0.912.

**Regression Coefficients:**

Table 4 indicates the regression coefficients ($a$ and $b$), where $a$ = –80.485 implies that when there is no fire outbreak we expect the loss of properties to be at a fall of ₦80,485,000.00. $b$ = 16.713 which implies that if fire outbreak increases by 1, the loss of properties is expected to increase by ₦16,713,000.00. This gives the WLS regression model as: $y_i = -80.485 + 16.713x_i$. The ***sig.****value* of 0.000 for the predictor variable (Fire outbreak) indicates that fire outbreak is a statistically useful predictor of loss of properties. In other words, fire outbreak is significant in predicting the loss of properties.

## VIII. CONCLUSIONS

From the results of this study, it can be concluded that there is a very strong positive relationship between the number of fire outbreak and the loss of properties; this relationship is however significant. An increase in the number of fire outbreak will lead to increase in monetary loss of properties and vice versa. In addition, fire outbreak exerts significant influence on loss of properties. Finally, fire outbreak accounts for approximately 91.2% of the loss of properties in the state.

## IX. RECOMMENDATIONS

1. Since an increase in the number of fire outbreak will lead to increase in monetary loss of properties, it is therefore important that the government setup efficient and effective fire service centres across each local government area of the state for effective response to any fire outbreak.

2. As part of effort to reduce loss of properties to fire outbreak, government and concerned bodies should frequently be orientating the populace on how best to prevent and manage fire outbreak.

## REFERENCES

[1] Brewer, K.R.W. (1963). "Ratio Estimation in Finite Populations: Some Results Deducible from the Assumption of an Underlying Stochastic Process," Australian Journal of Statistics, 5, pp. 93-105.

[2] Brewer, KRW (2002). Combined survey sampling inference: Weighing Basu's elephants, Arnold. London and Oxford University Press.

[3] Carroll, R.J., and Ruppert, D. (1988): Transformation and Weighting in Regression, Chapman &Hall.

[4] Celmins, A. (1998). The method of Gauss in 1799. Statistical Science, 13(2), 123-135.

[5] Econometrics Laboratory (1999): Regression Analysis Tutorial. University of California at Berkeley 22-26 March 1999.

[6] James R. Knaub (2012). Properties of Weighted Least Squares Regression for Cutoff Sampling in Establishment Surveys. Energy Information Administration.

[7] Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME-Journal of Basic Engineering, 82 (Series D), 35-45.

[8] Knaub, J.R., Jr. (1993). "Alternative to the Iterated Reweighted Least Squares Method: Apparent Heteroscedasticity and Linear Regression Model Sampling," Proceedings of the International Conference on Establishment Surveys (Buffalo, NY, USA), American Statistical Association, pp. 520-525.

[9] Knaub, J.R., Jr. (1997). "Weighting in Regression for Use in Survey Methodology," InterStat, April 1997, http://interstat.statjournals.net/. (Note shorter, but improved version in the 1997 Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 153-157.)

ISSN 2350-1049

International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS)
Vol. 2, Issue 3, pp: (45-54), Month: July 2015 - September 2015, Available at: www.paperpublications.org

[10] Knaub, J.R., Jr. (2008a). "Cutoff vs. Design-Based Sampling and Inference For Establishment Surveys," InterStat, June 2008, http://interstat.statjournals.net/.

[11] NIST(2009). National Institute of Standards and Technology Engineering Statistics Handbook, with SEMATECH contractor: NIST/SEMATECH Engineering Statistics Internet Handbook, downloaded on November 24, 2009, from URL http://www.itl.nist.gov/div898/handbook/pmd/section1/pmd143.htm.

[12] Plackett, R. L. (1949). A historical note on the method of least squares. Biometrika, 36(3/4), 458-460.

[13] Plackett, R. L. (1950). Some Theorems in Least Squares. Biometrika, 1/2, 149-157.

[14] Seal, H. L. (1967). Studies in the History of Probability and Statistics. XV: The Historical Development of the Gauss Linear Model. Biometrika, 54, 1-24.

[15] Steel. P. and Fay, R.E. (1995). "Variance Estimation for Finite Populations with Imputed Data," ASA Survey Research Methods Section Proceedings, 1995, pages 374-379. http://www.amstat.org/sections/SRMS/Proceedings/

[16] Stigler, S. M. (1981). Gauss and the invention of Least Squares. The Annals of Statistics, 9(3), 465-474.

[17] Sweet, E.M., and Sigman, R.S. (1995). "Evaluation of Model-Assisted Procedures for Stratifying Skewed Populations Using Auxiliary Data," ASA Survey Research Methods Section Proceedings, 1995, pages 491-496. http://www.amstat.org/sections/SRMS/Proceedings/

[18] Young, P. (1974). Recursive Approaches to Time Series Analysis. Bulletin / Institute of Mathematics and its Applications, 10(May/June), 209-224.

[19] Online http://www.itl.nist.gov/div898/handbook/index.htm

## APPENDIX

**TABLE 5: Monetary loss of properties due to fire outbreak across the twenty (20) Local Government Areas of Ogun state in year 2010**

| LGA | Number of Fire Outbreak | Loss of properties (m'₦) | Weights |
|---|---|---|---|
| ABEOKUTA NORTH | 55 | 1520.60 | 1.88552E-06 |
| ABEOKUTA SOUTH | 80 | 660.30 | 4.91234E-06 |
| ADO ODO OTA | 20 | 1040.45 | 2.07085E-06 |
| EWEKORO | 20 | 47.52 | 1.12589E-05 |
| IFO | 15 | 95.04 | 2.8696E-05 |
| IJEBU EAST | 12 | 47.52 | 2.60576E-05 |
| IJEBU NORTH | 13 | 480.15 | 1.99361E-05 |
| IJEBU NORTH EAST | 15 | 47.52 | 1.82322E-05 |
| IJEBU ODE | 50 | 760.30 | 0.000989852 |
| IKENNE | 25 | 95.04 | 1.01209E-05 |
| MEKO AFON | 15 | 47.52 | 1.82322E-05 |
| IPOKIA | 15 | 47.52 | 1.82322E-05 |
| OBAFEMI OWODE | 16 | 760.30 | 4.60859E-06 |
| ODEDA | 15 | 47.52 | 1.82322E-05 |
| ODOGBOLU | 20 | 95.04 | 1.59356E-05 |
| OGUN WATERSIDE | 12 | 95.04 | 4.54206E-05 |
| REMO NORTH | 12 | 95.04 | 4.54206E-05 |
| SAGAMU | 15 | 1240.45 | 1.08794E-06 |
| YEWA NORTH | 15 | 190.07 | 0.000119061 |
| YEWA SOUTH | 15 | 200.07 | 0.000150012 |