Journal of Economics and Sustainable Development
ISSN 2222-1700 (Paper) ISSN 2222-2855 (Online)
Vol.5, No.19, 2014

Climatic Effects on Major Oilseed Crops Production in Bangladesh: An Application of Multiple Regression Model

Mohammed Amir Hamjah
B.Sc. (Honours), M.S (Thesis) in Statistics
Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh.
Email: arif.stat.sust@hotmail.com.

Abstract
Climate has most vital effects on different agricultural crop production and sometimes climatic factors are main natural factors for crop production. In this study, it is tried to measure the climatic effects on different major oilseed crops production in Bangladesh. The main purpose of this study is to develop the Multiple Regression Model which is a well-established model to measure cause and effect relationship among the variables under study. To conduct this study, amount of land area corresponding year's production under study is the most important variable which should not be avoided because without including this variable appropriate model couldn't be fitted and from the study it is found that it has positively significant effects on each oilseed crop production. From the study, it is found that the value of R-square for Mustard, Linseed and Groundnut production are 0.9858, 0.9362 and 0.9039 respectively which are indicated that most of the variability can be explained by the climatic variables under study. From the global model validation test, it is clear that these models are valid linear regression model.
Keywords: Climatic Effects, Oilseed crops, Multiple Regression Model and Bangladesh.

1. Introduction

The group of crops containing oil and fatty acids in their seeds and the oil is extracted from their seeds either as vegetable or industrial oil are known as oilseed crops. The major oilseed crops grown in Bangladesh are mustard, sesame, groundnut and linseed. The minor oil crops are niger, soybean, sunflower, safflower and castor. The major contribution of oil comes from mustard (65%) followed by sesame (10.71%) and groundnut (invisible oil 10.5%) (Bangladesh Bureau of Statistics, 2004). In Bangladesh, oil crops have equally been neglected by farmers, researchers, extension workers and policy planners. Research Institutes never had ample resources at their disposal to follow a systematic Research and Development program that would bring the oilseed crops high on their priorities. Poor yielding capacity and volatile, though subdued, demand have pushed these crops onto marginal lands where they produce very poor yields since hardly any inputs are given or adequate management practices followed. Result is that production, area and productivity of most oilseed crops are declining.

Groundnuts are one of the major oilseed crops of Bangladesh, but yields are low when compared to the world average, with the result that Bangladesh produces only about 40% of its domestic oil consumption. Groundnuts are mostly used as ingredients for a number of industrially processed foods and contribute little to oil production. Groundnut is a major crop in the char lands of Bangladesh, but because of poor yields, farmers derive a limited income from the crop. The purpose of this project is to improve the livelihoods of poor and vulnerable farmers by introducing new affordable production technologies and by more effectively linking farmers with potential buyers. Mustard is an anpther important edible oilseed crops in Bangladesh, commonly known as sarisha that contain oil & protein in their seeds. It attains first position among oilseed crops in terms of both area & cultivation. It is well known for its versatile uses. Oil-cake is used as both cattle feed & organic fertilizer. Dry plant is also used as fuel.

Most agronomists believe that agricultural production will be mostly affected by the severity and pace of climate change, not so much by gradual trends in climate. If change is gradual, there may be enough time for biota adjustment. Rapid climate change, however, could harm agriculture in many countries, especially those that are already suffering from rather poor soil and climate conditions, because there is less time for optimum natural selection and adaption. Climate change and agriculture are interrelated processes, both of which take place on a global scale. Global warming is projected to have significant impacts on conditions affecting agriculture, including temperature, carbon dioxide, glacial run-off, precipitation and the interaction of these elements. These conditions determine the carrying capacity of the biosphere to produce enough food for the human population and domesticated animals. The overall effect of climate change on agriculture will depend on the balance of these effects. Assessment of the effects of global climate changes on agriculture might help to properly anticipate and adapt farming to maximize agricultural production.

2. Objective of the study

The main objective of the study is to develop the best Multiple Regression Model for measuring the significant of climatic effects on oilseed crops production in Bangladesh. The specific objectives of this study is to developing a Multiple Regression Model for measuring the climatic effects on different types of oilseed crops named as Mustard, Linseed and Groundnut production in Bangladesh.

3. Review of Literature

A lots of study have been done to measure the climatic effects on different agricultural crops production all over the world such as Hamjah (2014) has conducted an analysis to measure the climatic effects on Cotton and Tea production in Bangladesh by using Multiple Regression Model and here he has also measured the production efficiency due to climates using Stochastic Frontier Model. Hamjah (2014) has conducted another study to measure climatic effects on rice crop production in Bangladesh. Hamjah and Chowdhury (2014) has fitted different model to measure the climatic and hydrological effects on agricultural crop production in Bangladesh using Multiple Regression and ARIMAX model and hence they have made a comparative study between multiple Regression and ARIMAX model to select best model. Richard M. Adams, Brian H. Hurd, Stephanie Lenhart and Leary (1998) have conducted a study which reviews the extant literature on these physical and economic effects and interprets this in terms of common themes or findings. Shafiqur Rahman (2008) has made an analysis by which he has shown the significant effects of temperature on agricultural production by using regression and correlation analysis. Hag Hamad Abdelaziz, Adam Abdelrahman, Abdalla and Mohmmed Alameen Abdellatif (2010) have shown that shed light on the main constraints of crop production in the traditional rainfed sector in Umkdada district, North Darfur State (Sudan). The study used descriptive statistics and regression for data analysis. The results of regression analysis revealed that the crops produced in season 2006 were significantly affected by some factors.

4. Methodology

4.1. Classical Linear Multiple Regression Model

The multiple classical linear regression model is given by

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \dots\dots\dots. \beta_q X_{qi} + \varepsilon_i \quad i = 1,2,3, \dots, N \qquad (1)$$

Here, Y = Dependent variable, X_i's are independent variables, ε = stochastic error term, and β_0, β_1, β_2, … B_q are the model's parameter which are to be estimated.

There are five critical assumptions relating to Classical Linear Multiple Regression Model. These assumptions required to show that the estimation technique, Ordinary Least Squares (OLS), has a number of desirable properties, and also so that the hypothesis tests regarding the coefficient estimates could validly be conducted. These assumptions are (1) E (ε i) = 0 , The errors have zero mean, (2) Var (εi) = $\sigma^2 < \infty$, The values variance of the error is constant and have finite over all values of x_i, (3) Cov (ε i , ε j) = 0, The errors are statistically independent of one another, (4) Cov (ε ,i xi) = 0, There is no relationship between the error and the corresponding x, (5) $\varepsilon_i \sim N(0, \sigma^2)$, ε_i is normally distributed.

The following test also used to check whether the model assumption are satisfied or not

- **Shapiro–Wilk** (1965 by Samuel Shapiro and Martin Wilk.) normality test is used to check the normality assumptions of residuals of the fitted model.
- **Ljung-Box** (Box and Ljung, 1978) test is used to check autocorrelation among the residuals of the fitted model's residuals which is based on chisquare test.
- **Studentized Breusch-Pagan test** (Breusch and Pagan, 1979) is used to check whether the equality of variance of the fitted model's residuals are satisfied or not.

4.2. Global Test of Validity Checking for a Linear Model

An easy-to-implement global procedure for testing the four assumptions of the linear model is proposed. The test can be viewed as a Neyman smooth test (1937) and it only relies on the standardized residual vector. If the global procedure indicates a violation of at least one of the assumptions, the components of the global test statistic can be utilized to gain insights into which assumptions have been violated. The procedure can also be used in conjunction with associated deletion statistics to detect unusual observations.

This distributional assumption, together with the linear link specification in are enumerated as four distinct assumptions:

(A1) (*Linearity*) $E\{Y_i|X\} = x_i\beta$,where $\mathbf{x}_i$ is the *i*th row of X;
(A2) (*Homoscedasticity*) $Var\{Y_i|X\} = \sigma^2$, $i = 1,2,\dots,n$;
(A3) (*Uncorrelatedness*) $Cov\{Y_i,Y_j|X\} = 0, (i \neq j)$; and
(A4) (*Normality*) $(Y_1,Y_2,\dots,Y_n)|X$ have a multivariate normal distribution.

Assumptions (A3) and (A4) imply that, given X, Y_i, i = 1, 2, …, n are independent normal random variables. Without loss of generality, we assume that **X** is of full rank with $n > p$, so rank(**X**) = p. Under (A1)–(A4), the maximum likelihood (ML) estimators of β and σ^2 are given, respectively, by

$b=\hat{\beta}=(XtX)^{-1} XtY$ and $s^2=\sigma^2$=lnYt(I−P[X])Y;

Assessment of whether assumptions (A1)–(A4) are satisfied, based on the data (Y, X), has received considerable attention. Assessment procedures typically involve the standardized residuals **R**, herein defined according to

$$R_i = \frac{(Y_i - \hat{Y})}{s}$$

Where, $\hat{Y}_i$ is the fitted value of Y_i

Formal significance tests for (A1)–(A4) involve testing the null hypothesis (H_0) versus the alternative hypothesis (H_1), where

H_0 : Assumptions (A1)−(A4) all hold

H_1 : At least one of (A1)–(A4) does not hold.

The first and second components for the test is given by

$$S_1 = \left\{\frac{1}{\sqrt[2]{6n}} \sum_{i=1}^{n} R^3{}_i\right\}^2; \quad S_2 = \left\{\frac{1}{\sqrt[2]{24n}} \sum_{i=1}^{n} (R^4{}_i - 3)\right\}^2$$

The third component for the test is given by

$$S_3 = \frac{\{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - \hat{Y})^2 R_i\}^2}{(\hat{\Omega} - b^t \hat{\Sigma}_x b - \Gamma \hat{\Sigma}_x{}^{-1} \Gamma)}$$

Where, $\hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y})^4$ and $\hat{\Sigma}_x = \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$; $\Gamma = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}\right)^2 (X_i - \bar{X})$

The Fourth component for the test is given by (the fourth component statistic requires a user-supplied n × 1 vector V, which by default is set to be the time sequence $V = (1, 2, \ldots, n)^t$.)

$S_4 = \frac{1}{\sqrt{2\,\hat{\sigma}^2\, n}} \sum_1^n (V_i - \bar{V})(R_i{}^2 - 1)$; where $\hat{\sigma}^2 = \frac{1}{n} \sum_1^n (V_i - \bar{V})^2$

The global test statistics is given by $G^2 = S_1 + S_2 + S_3 + S_4$

Now reject H_0, if $G^2 > G^2{}_{\alpha,4}$

5. Data Sources and Data Manipulation

The climatic information are available from the Bangladesh Government's authorized websites of Bangladesh Agricultural Research Council (BARC) named as ***www.barc.gov.bd.*** The oilseed crops datasets are also available from Bangladesh Agricultural Ministry's website named as ***www.moa.gov.bd.*** These dataset are available from the year 1972 to 2006. Climatic information was in the original form such a way that it is arranged in the monthly average information corresponding to the years from 1972 to 2006 according to the 30 climatic stations. The name of these stations are Dinajpur, Rangpur, Rajshahi, Bogra, Mymensingh, Sylhet, Srimangal, Ishurdi, Dhaka, Comilla, Chandpur, Josser, Faridpur, Madaripur, Khulna, Satkhira, Barisal, Bhola, Feni, MaijdeeCourt, Hatiya, Sitakunda, Sandwip, Chittagong, Kutubdia, Cox's Bazar, Teknaf, Rangamati, Patuakhali, Khepupara, Tangail, and Mongla. It is taken the month October, November, December, January and February as a "dry season" and March, April, May, June, July, August, September as a "summer season" considering the weather and climatic conditions of Bangladesh. Then, finally we take average seasonal climatic information of 30 climatic station corresponding to the year from 1972 to 2006. We take the average of 30 climatic area because of focusing overall country's situation and overall model fitting for whole Bangladesh.

6. Used Software for Analysis

This analysis has completely done using statistical programming based open source Software **"R"** for windows (version 2.15.1). The additional library packages used for analysis are lmtest, gvlma, car, etc.

7. Used Climatic Variables in This Study

sun.sum = Sunshine of the Summer Season, sun.dry = Sunshine of the Dry Season , clo.sum = Cloud Coverage of the Summer Season, clo.dry = Cloud Coverage of the Dry Season, max.tem.dry = Maximum Temperature of the Dry Season, max.tem.sum = Maximum Temperature of the Summer Season, min.tem.dry = Minimum Temperature of the Dry Season, min.tem.sum = Minimum Temperature of the Summer Season, rain.dry= Ammount of Rainfall of the Dry Season, rain.sum= Amount Rainfall of the Summer Season, rh.dry = Relative Humidity of the Dry Season, rh.sum= Relative Humidity of the Summer Season, wind.dry = Wind Speed of the Dry Season and wind.sum = Wind Speed of The Summer Season.

8. Results & Discussion

8.1. Multiple Regression Modeling of Mustard Production

The Multiple Regression Model to measure the climatic effects on Mustard production is given by

$$\begin{aligned} \text{mustard} = {} & \beta_0 + \beta_1 \text{mustard.area} + \beta_2\, \text{sun.sum} + \beta_3 \text{sun.dry} + \beta_4\, \text{clo.sum} + \beta_5 \text{clo.dry} + \beta_6\, \text{max.tem.dry} \\ & + \beta_7\, \text{max.tem.sum} + \beta_8\, \text{min.tem.dry} + \beta_9\, \text{min.tem.sum} + \beta_{10}\, \text{rain.dry} + \beta_{11} \text{rain.sum} \\ & + \beta_{12}\, \text{rh.dry} + \beta_{13}\, \text{rh.sum} + \beta_{14}\, \text{wind.dry} + \beta_{15}\, \text{wind.sum} + \varepsilon \end{aligned}$$

where, $\beta_1, \beta_2, \beta_3, \ldots, \beta_{15}$ are the partial regression coefficients, β_0 is the intercept term and ε = stochastic error term

The parameter estimates of Mustard production model are given in Table 1.

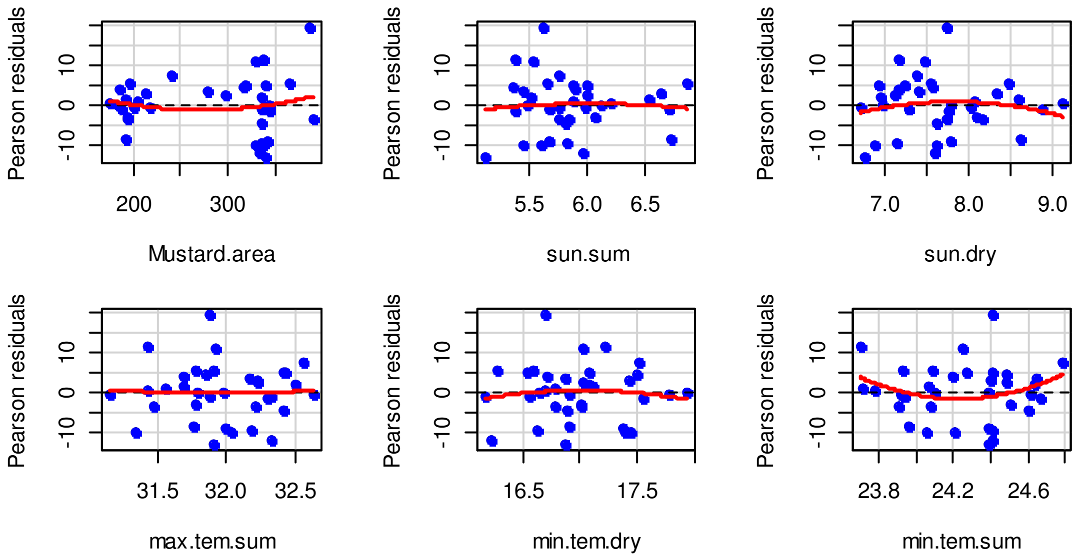
Table 1: Summary Statistics of Mustard production Model

Coefficients	Estimate	Std. Error	tvalue	Pr (>|t|)
(Intercept)	-1476.096	648.998	-2.274	0.0347
Mustard.area	0.689097	0.04587	15.022	<0.0001
sun.sum	7.034251	15.857	0.444	0.6623
sun.dry	-16.68547	8.8026	-1.896	0.0733
clo.sum	3.178672	20.6679	0.154	0.8794
clo.dry	-15.53642	20.8493	-0.745	0.4653
max.tem.dry	-4.323514	13.059	-0.331	0.7442
max.tem.sum	45.609605	23.4552	1.945	0.0668
min.tem.dry	13.07381	14.2104	0.92	0.3691
min.tem.sum	-21.4537	18.4434	-1.163	0.2591
rain.dry	-0.082968	0.15649	-0.53	0.6021
rain.sum	0.032709	0.07537	0.434	0.6692
rh.dry	-4.909475	2.92378	-1.679	0.1095
rh.sum	11.185241	5.85781	1.909	0.0714
wind.dry	11.272632	40.9504	0.275	0.7861
wind.sum	-5.693796	19.4434	-0.293	0.7728

From the Table 1, we observe that mustard.area, sun.sum, clo.sum, max.tem.sum, min.tem.dry, rain.sum, rh.sum and wind.dry have positive effects on Mustard production; and sun.dry, clo.dry, max.tem.dry, min.tem.sum, rain.dry, rh.dry and wind.sum have negative effects on Mustard production. Again, mustard.area, sun.dry, max.tem.sum, rh.dry and rh.sum have statistically significant effects on Mustard crop production at 10% level of significance.

From the fitted Multiple Regression model, the value of Multiple R-squared is 0.9858 which implies that 98.58% of the total variations can be explained by the regressor (climatic) variables and Adjusted R-squared is 0.9746 which indicates that 97.46% variation can be explained by the regressor variables after adjustments. Again, from the overall test, $Pr(|F_{(15, 19)}| \geq 88.12) < 0.000001$ implies that all climatic variables are not eqully significant effects on Mustard production at 5% level of significance.

Residuals versus regressor Plots of Mustard production model are shown in the Figure 1.



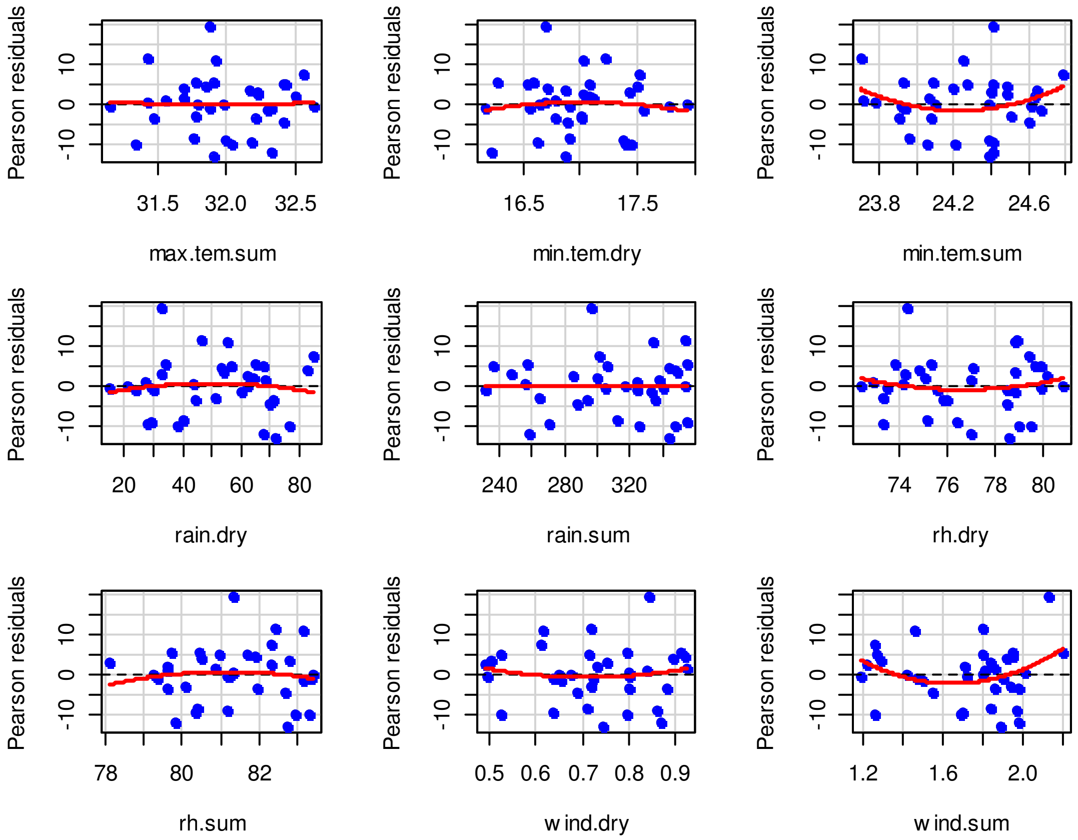
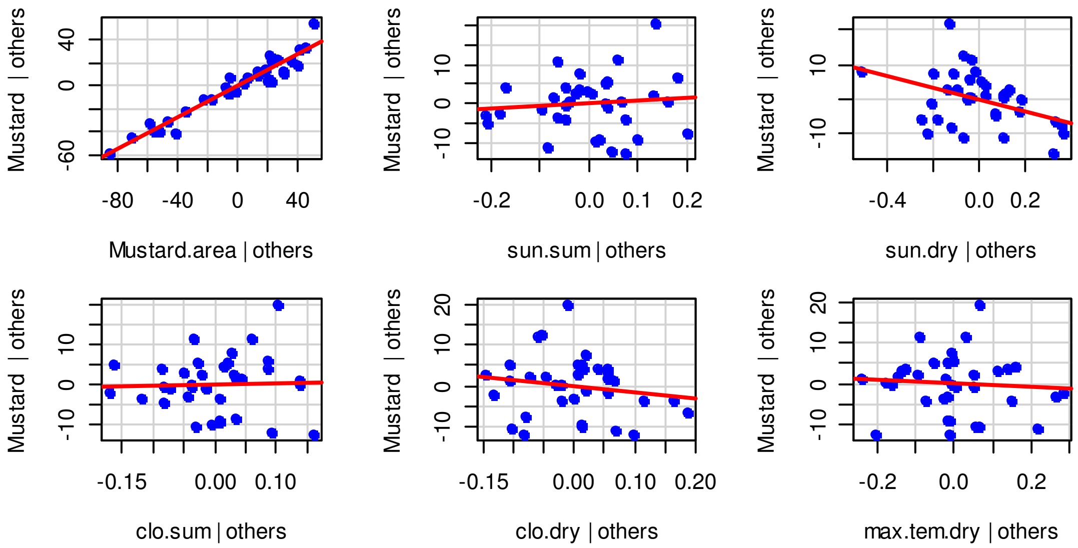


Figure 1: Residuals versus Regressor plot of Mustard Production Model

From the Figure 1 which shows that there is no curvature band or a nonlinear pattern in the initial stage in all "residuals versus regressors" plots and they create a horizontal band to make a decision that there is no curvature relationship between the response and each of the regressor variable. This indicates that there is no need to take any transformation on the predictor variables to make a linear relation between regressors and response variable. Added Variable Plots of the Mustard production's model are shown in the Figure 2.



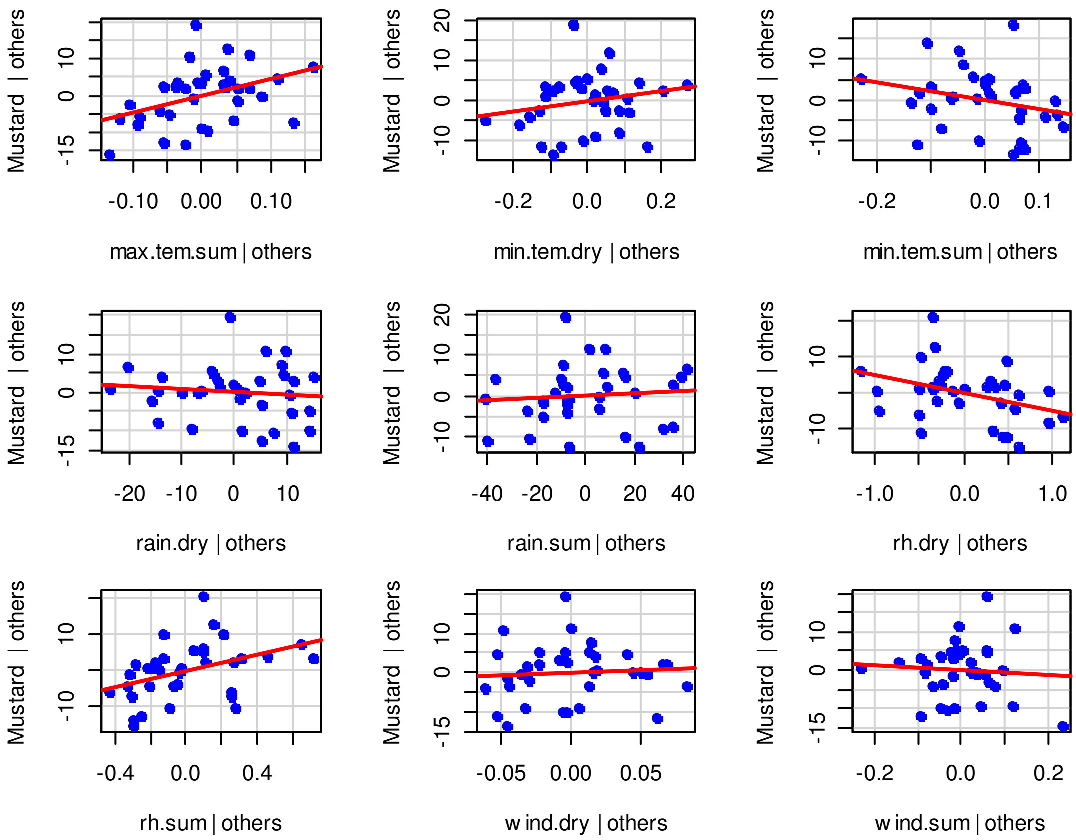


Figure 2: *Added Variable Plots for Mustard Productions Model*

From the Figure 2 which displays the partial relationship between the response (Mustard) residuals and each of the predictor's residuals of Mustard production model. All plots show that they follow a staright line with non-zero slope and there is no curvature relationship ammong the predictor's residuals and response's residuals. That is why, it can be said that each of the predictor variabls are added to the model with maintaining a linear relation, that is, this model is going to make a linear relationship between the response variable and the predictor variables to measure the climatic effects on Mustard production in Bangladesh.

Residuals Diaagonstics Plots are shown in the Figure 3.

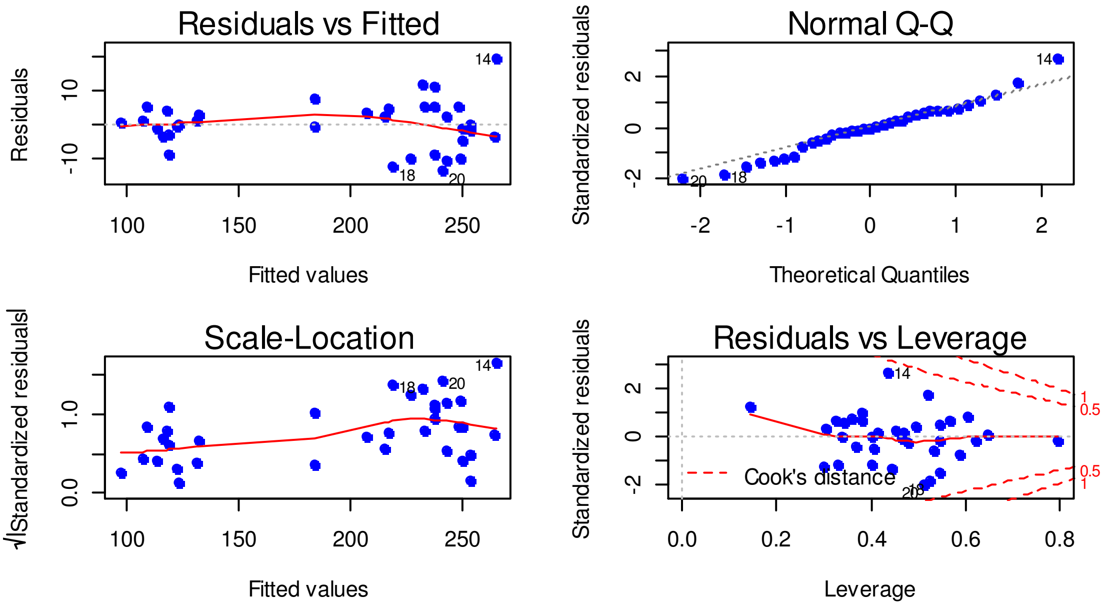


Figure 3: *Residuals Diagonstic Plots for Mustard Production Model*

From the Figure 3, we observe that

- All of the points are lied around the horizontal line and they try to create a horizontal band which implies constant variance among the residuals of the Mustard production model (top-left).

- almost all of the points try to create a horizontal band which indicates that residals have constant variance of the Mustard production model (bottom-left).
- almost all of the points are very closed to Q-Q line or on the Q-Q line which suggests that residuals are normally distributed of the Mustard production model (top-right).
- There is no leverage points and according to the cook's distance, all points are lied inside approximately 20% Cook's interval of the leverage points indicating there is no bad influence on the model properties of the Mustard production model (bottom-right).

Global validation test is used to check whether Mustard production model is valid or not. The test is performed at 5% level of significance on 4 degrees of freedom. The results from the test are shown in the Table 2.

Table 2: *Global Validation Checking for Mustard Production Model*

	Value	p-value	Decision
Global Stat	0.899	0.9247	Assumptions acceptable.
Skewness	0.329	0.5662	Assumptions acceptable.
Kurtosis	0.1711	0.6791	Assumptions acceptable.
Link Function	0.1061	0.7446	Assumptions acceptable.
Heteroscedasticity	0.2928	0.5885	Assumptions acceptable.

From the Table 2, it is observed that the p-value of Global stat is 0.9247 which strongly suggests that linearity of parameters, Homoscedasticity, Autocorrelation and Normality assumptions are very well managed in the fitted model, that is, this model is a valid linear linear regression model. Again, Skewness and Kurtosis of the fitted model's residuals are 0.329 and 0.1711respectively and their corresponding p-values for testing hypothesis are 0.5662 and 0.6791 which are suggested that the assumptions of the skewness and kurtosis are very well accepted to fit a linear model. Similarly, the link function assumptions is also satisfied to fit a linear model under validation checking of a linear regression model. At the same time, the heteroscedasticity assumptions is also accepted with the p-value of 0.5885 indicating homoscedasticity of variance. We can easily say that the fitted model is the best fitted valid Multiple Linear Regression model for measuring the climatic effects on Mustard production in Bangladesh.

Finally, from all of the test, assumptions of residuals like Homoscedasticity, Autocorrelation, Normality are very well satisfied and model validation test "Global Tesst" also satisfied all of the assuptions for a linear model indicating the fitted model is a valid linear regression model. Without any kind of loss of generality, it can be said that this is the best fitted Multiple Regression Model to measure the climatic effects on Mustarsd production based on sample data.

8.2. Multiple Regression Modeling of Linseed Production

The Multiple Regression Model to measure the climatic effects on Linseed production is given by

$$linseed = \beta_0 + \beta_1 linseed.area + \beta_2\,\mathrm{sun.sum} + \beta_3 \mathrm{sun.dry} + \beta_4\,\mathrm{clo.sum} + \beta_5 \mathrm{clo.dry} + \beta_6\,\mathrm{max.tem.dry} + \beta_7\,\mathrm{max.tem.sum} + \beta_8\,\mathrm{min.tem.dry} + \beta_9\,\mathrm{min.tem.sum} + \beta_{10}\,\mathrm{rain.dry} + \beta_{11}\mathrm{rain.sum} + \beta_{12}\,\mathrm{rh.dry} + \beta_{13}\,\mathrm{rh.sum} + \beta_{14}\,\mathrm{wind.dry} + \beta_{15}\,\mathrm{wind.sum} + \varepsilon$$

where, $\beta_1, \beta_2, \beta_3$, …. , β_{15} are the partial regression coefficients, β_0 is the intercept term and ε = stochastic error term.

The parameter estimates of the Linseed production model are given in Table 3.

Table 3: Summary Statistics of Linseed production Model

Coefficients	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-1304.077613	458.372067	-2.845	0.0104
linseed.area	0.397397	0.061417	6.47	<0.0001
sun.sum	-3.349327	11.403684	-0.294	0.7722
sun.dry	-10.136926	6.502692	-1.559	0.1355
clo.sum	-4.339881	15.474695	-0.28	0.7822
clo.dry	-4.364294	15.267105	-0.286	0.7781
max.tem.dry	-4.490873	9.497786	-0.473	0.6417
max.tem.sum	42.032386	16.455674	2.554	0.0194
min.tem.dry	13.145569	10.456576	1.257	0.2239
min.tem.sum	-26.63935	13.01178	-2.047	0.0547
rain.dry	-0.113716	0.114637	-0.992	0.3337
rain.sum	-0.002389	0.060811	-0.039	0.9691
rh.dry	-3.005748	2.133152	-1.409	0.175
rh.sum	9.755216	4.302231	2.267	0.0352
wind.dry	32.914243	30.752119	1.07	0.2979
wind.sum	38.458676	13.786574	2.79	0.0117

From the Table 3, it is observed that linseed.area, max.tem.sum, min.tem.dry, rh.sum, wind.dry and wind.sum have positive effects on Linseed production. Again, sun.dry, sun.sum, clo.dry, clo.sum, max.tem.dry,

min.tem.sum, rain.dry, rain.sum, rain.dry and rh.dry have negative effects on Linseed production. At the same time, linseed.area, max.tem.sum, min.tem.sum, rh.sum and wind.sum have statistically significant effects on Linseed crop production at 10% level of significance.
From the fitted Multiple Regression model, the value of Multiple R-squared is 0.9362 which implies that 93.62% of the total variations can be explained by the regressor variables and the value of Adjusted R-squared is 0.8857 which indicates that 88.57% variation can be explained by the regressor variables after adjustments. From the overall test, $\Pr(|F_{(15, 19)}| \geq 18.57) < 0.000001$ which implies that all the variables are not eqully significant effects on Linseed production at 5% level of significance.
Residuals diagnostic Plots of Linseed production's model are shown in the Figure 4.

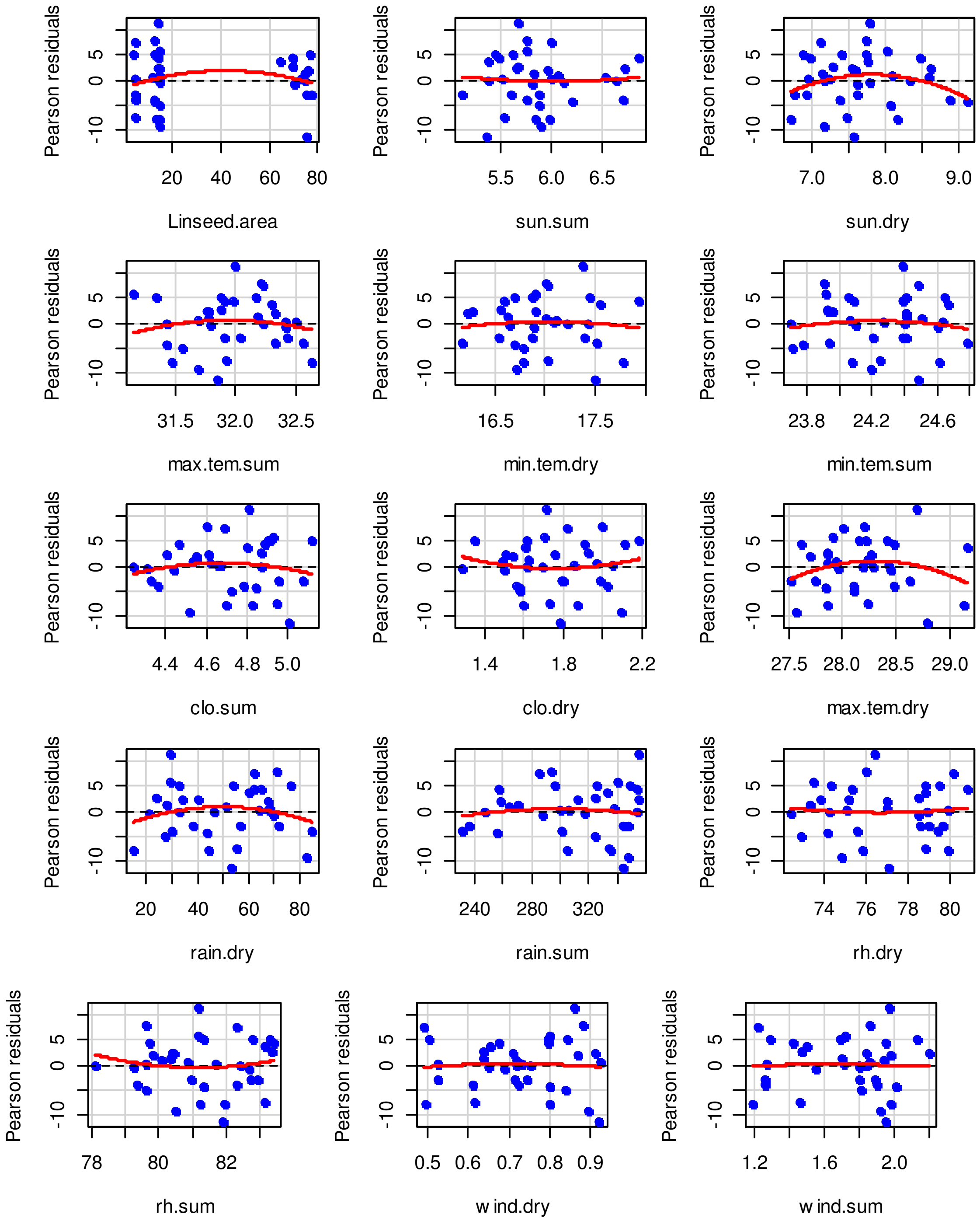


Figure 4: Residual versus Regressor plots of Linseed Production Model

From the Figure 4 which shows that all of the plots are trying to follow a horizontal band and they create a horizontal band to make a decision that there is no curvature relationship between the response and each of the other regressor variables. This also indicates that there is no need to take any further transformation among the

regeressor variable to make a linear relation between regressors and response variable.
Added Variable plots of the Linseed production model are shown in the Figure 5.

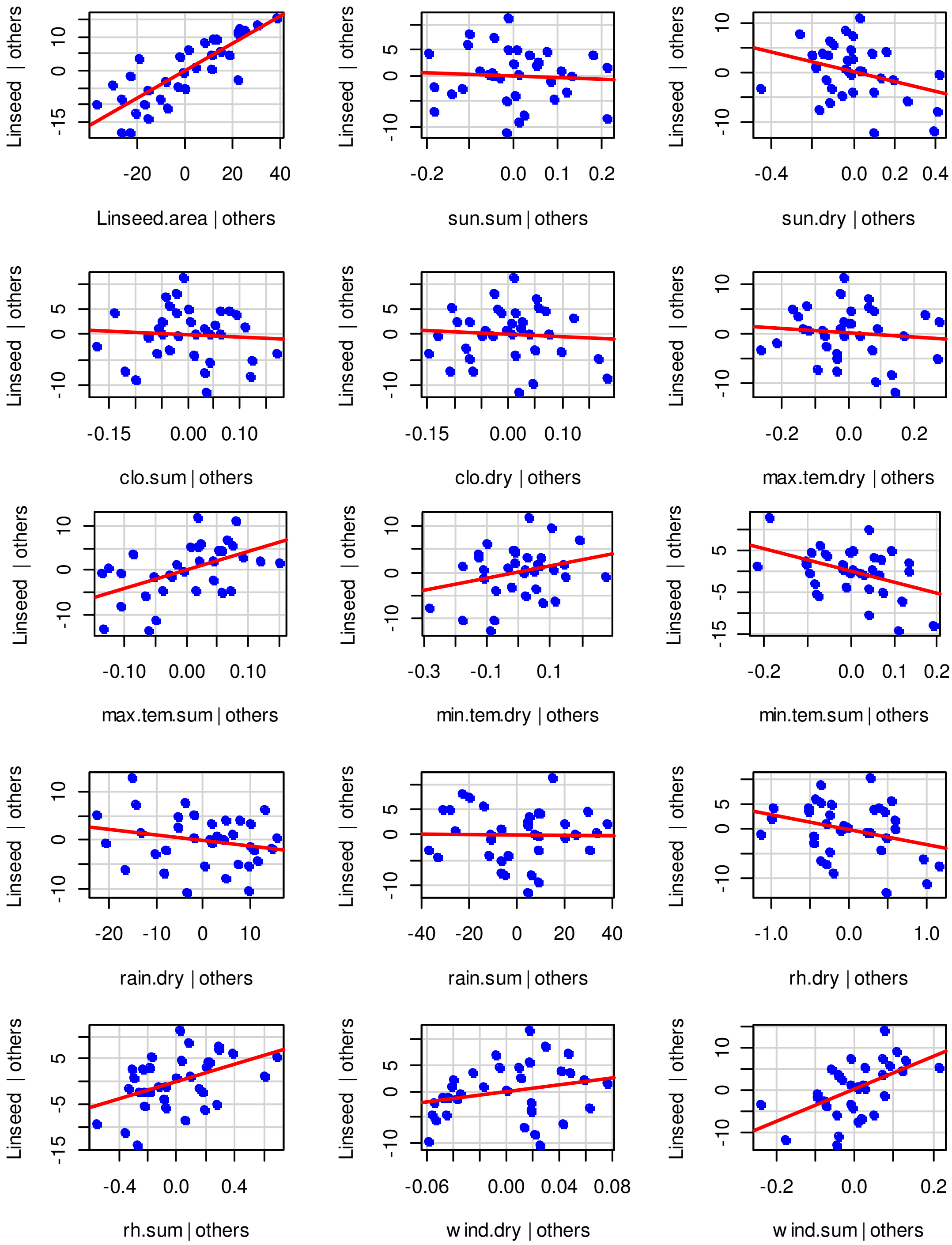


Figure 5: Added Variable Plots of Linseed Production Model

From the Figure 5 which displays the partial relationship between the response's (Linseed production) residuals and each of the predictor's residuals of Linseed production model. Each plot shows that they follow a staright line with non-zero slope and there is no curvature relationship ammong the predictor's residuals and response's residuals. That is why, it can be said that each of the predictor variabls are added to the model with maintaining a linear relation, that is, this model is going to make a linear relationship between the response variable and the predictor variables to measure the climatic effects on Linseed production in Bangladesh.
Residual Diagonstic plots of Linseed production model are shown in the Figure 6.

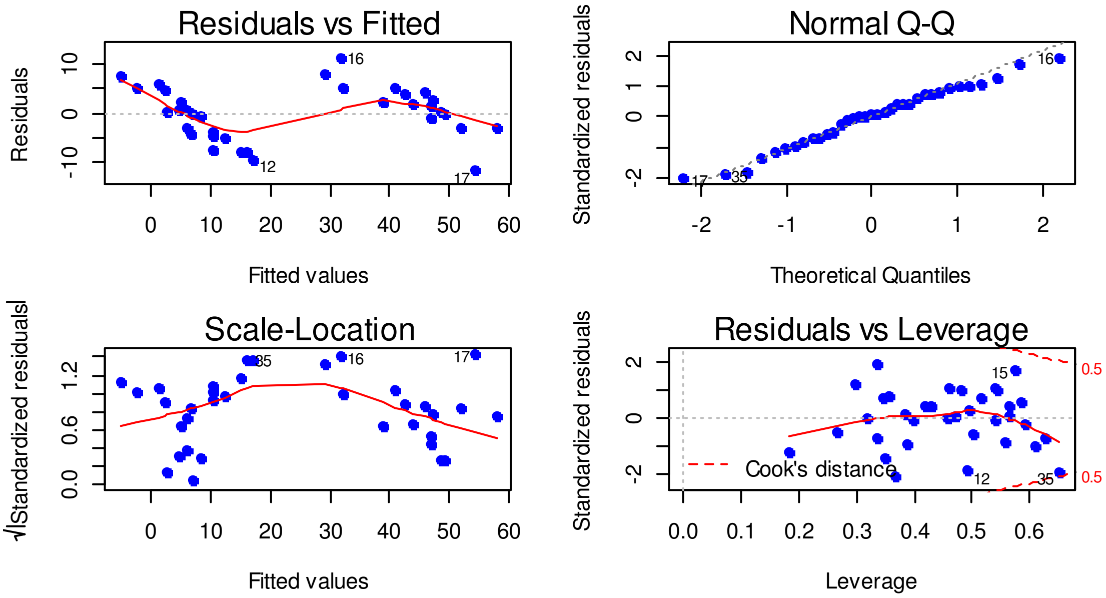


Figure 6: Residuals DiaagonsticPlots of Linseed Production Model

From the Figure 6, we observe that

- Most of the points are lied around the horizontal line and trying to create a horizontal band which implies constant variance among the residuals of the Linseed production model (top-left).
- almost all of the points try to create a horizontal band which indicates that residals have constant variance of the Linseed production model (bottom-left).
- almost all of the points are very closed to the Q-Q line or on the Q-Q line which suggests that residuals of the Linseed production model are normally distributed (top-right).
- There is a single leverage points and according to the cook's distance, it lies on the approximately 50% Cook's interval of the leverage points indicating it has a bit of influence on the model properties of the Linseed production model (bottom-right).

Global model validation test is used to check whether Linseed production model is valid or not. The test is performed at 5% level of significance on 4 degrees of freedom. The results from the test are shown in the Table 4.

Table 4: Global Validation Checking of Linseed Production Model

	Value	p-value	Decision
Global Stat	0.543826	0.9691	Assumptions acceptable.
Skewness	0.208271	0.6481	Assumptions acceptable.
Kurtosis	0.246879	0.6193	Assumptions acceptable.
Link Function	0.081482	0.7753	Assumptions acceptable.
Heteroscedasticity	0.007194	0.9324	Assumptions acceptable.

From the Table 4, we observe that the p-value of Global stat is 0.9691 which strongly suggests that linearity of parameters, Homoscedasticity, Autocorrelation and Normality assumptions are very well managed in the fitted model, that is, this model is a valid regression linear model. Again, Skewness and Kurtosis of the fitted model's residuals are 0.208271 and 0.246879 respectively and their corresponding p-values for testing hypothesis are 0.6481 and 0.6193 which are suggested that the assumptions of the skewness and kurtosis are very well accepted to fit a linear model. Similarly, the Link Function assumptions is also satisfied to fit a linear model with p-value = 0.7753 under validation checking of a linear regression model. At the same time, the heteroscedasticity assumptions is also accepted with the p-value of 0.9324 suggesting homoscedasticity of variance. It can easily be said that the fitted model is the best fitted valid Multiple Linear Regression model for measuring the climatic effects on Linseed production in Bangladesh.

Finally, from all of the formal test, assumptions of residuals like Homoscedasticity, Autocorrelation, Normality are very well satisfied and model validation test "Global Tesst" also satisfied indicating the fitted model is a valid linear regression model. Without any kind of loss of generality, it can be said that this is the best fitted Multiple Regression Model to measure the climatic effects on Linseed production based on sample data.

8.3. Multiple Regression Modeling of Groundnut Production

The Multiple Regression Model to measure the climatic effects on Groundnut production is given

$$groundnut = \beta_0 + \beta_1 groundnut.area + \beta_2 \text{ sun.sum} + \beta_3 \text{sun.dry} + \beta_4 \text{ clo.sum} + \beta_5 \text{ clo.dry} +$$

$$\beta_6 \text{ max. tem. dry} + \beta_7 \text{ max. tem. sum } + \beta_8 \text{ min. tem. dry } + \beta_9 \text{ min. tem. sum} + \beta_{10} \text{ rain. dry } + \beta_{11}\text{rain. sum} + \beta_{12} \text{ rh. dry } + \beta_{13} \text{ rh. sum } + \beta_{14} \text{ wind. dry} + \beta_{15} \text{ wind. sum} + \varepsilon$$

where, $\beta_1, \beta_2, \beta_3$, …. , β_{15} are the partial regression coefficients, β_0 is the intercept terms and ε = stochastic error term.

The parameter estimates of Groundnut production model are given in Table 5.

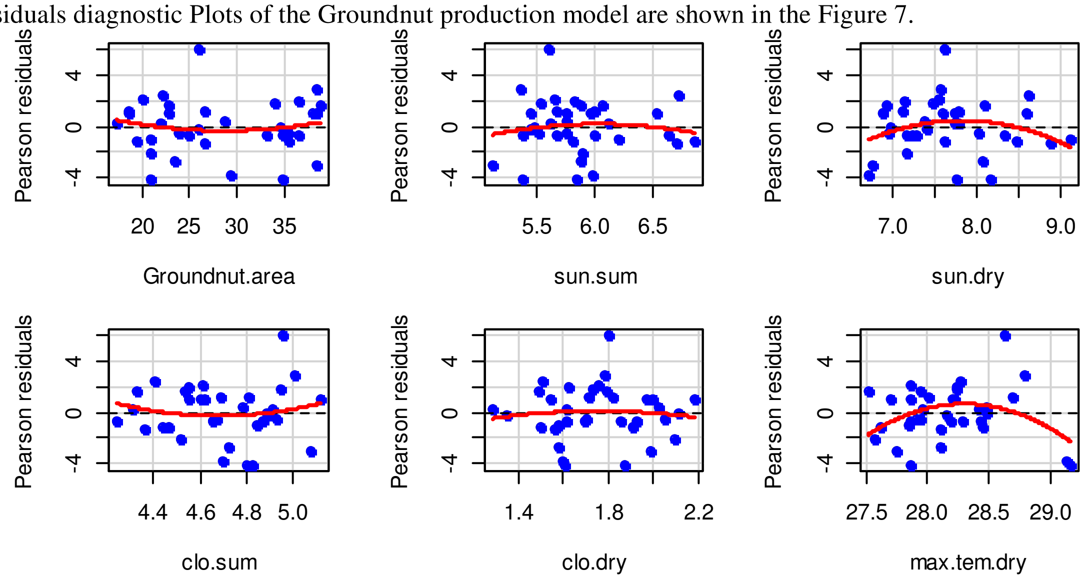
Table 5: Summary Statistics of Groundnut Production Model

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	-487.86045	187.71009	-2.599	0.0176
groundnut.area	0.60652	0.10615	5.714	<0.0001
sun.sum	3.96823	4.67126	0.849	0.4062
sun.dry	-1.68138	2.61599	-0.643	0.5281
clo.sum	13.84384	6.20481	2.231	0.0379
clo.dry	-3.13006	6.23768	-0.502	0.6216
max.tem.dry	2.99507	3.87904	0.772	0.4495
max.tem.sum	14.03989	6.73774	2.084	0.0509
min.tem.dry	-1.02466	4.24313	-0.241	0.8118
min.tem.sum	-7.01653	5.29464	-1.325	0.2008
rain.dry	0.01147	0.0467	0.246	0.8086
rain.sum	0.01188	0.02456	0.484	0.6342
rh.dry	0.02232	0.88258	0.025	0.9801
rh.sum	0.85537	1.71256	0.499	0.6232
wind.dry	8.58433	12.30915	0.697	0.494
wind.sum	4.7598	5.64059	0.844	0.4093

From the Table 5, we observe that groundnut.area, sun.sum, clo.sum, max.tem.dry, max.tem.sum, rain.dry, rain.sum, rh.dry, rh.sum, wind.dry and wind.sum have positive effects on Groundnut production. Again, sun.dry, clo.dry, min.tem.dry and min.tem.sum have negative effects on Groundnut production. At the same time, groundnut.area, clo.sum and max.tem.sum have statistically significant effects on Groundnut crop production at 5% level of significance.

From the fitted Multiple Regression model, the value of Multiple R-squared is 0.9039 which implies that 90.39% of the total variations can be explained by the regressor (climatic) variables and that of Adjusted R-squared is 0.8281 which implies that 82.81 % of total variations can be explained by the regressor variables after adjustments. Again, from the overall test, $Pr(|F_{(15, 19)}| \geq 11.92) < 0.0001$ which implies that all variables are not equly significant effects on Groundnut production at 5% level of significance.

Residuals diagnostic Plots of the Groundnut production model are shown in the Figure 7.



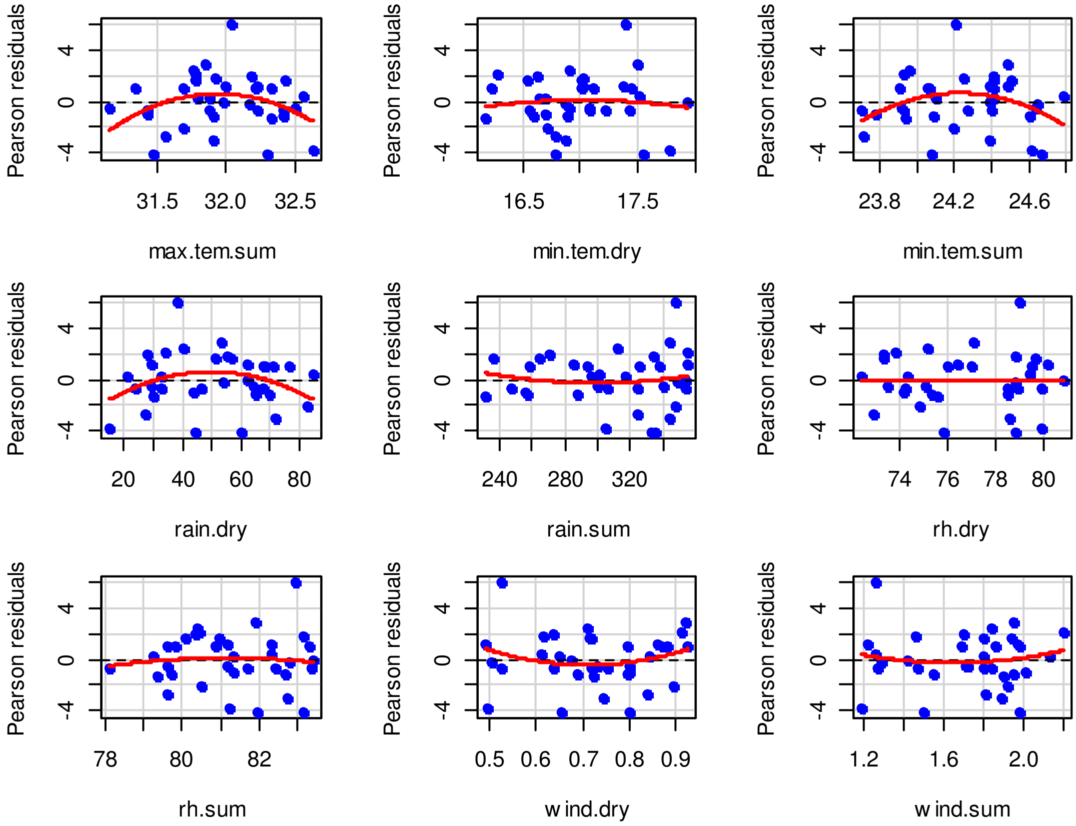
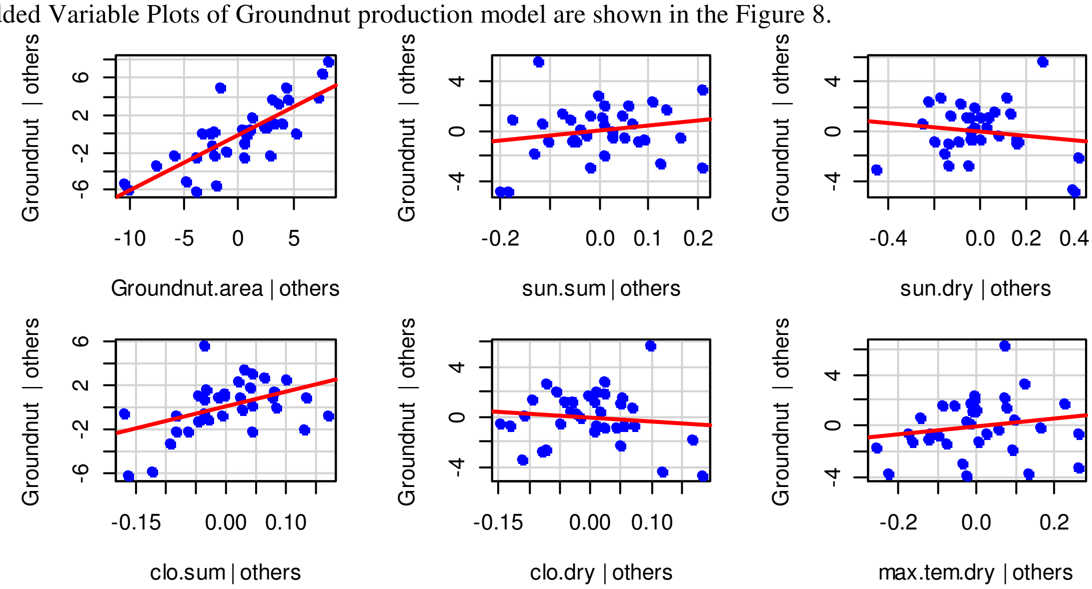


Figure 7: Residuals versus Regressor Plots of Groundnut Production Model

From the Figure 7 which shows that each of the plots are trying to follow a horizontal band and create a horizontal band to make a decision that there is no curvature relationship between the response and each of the other regressor variable. This indicates that there is no need to take any further transformation on the regressor variable to make a linear relation between regressors and response variable.

Added Variable Plots of Groundnut production model are shown in the Figure 8.



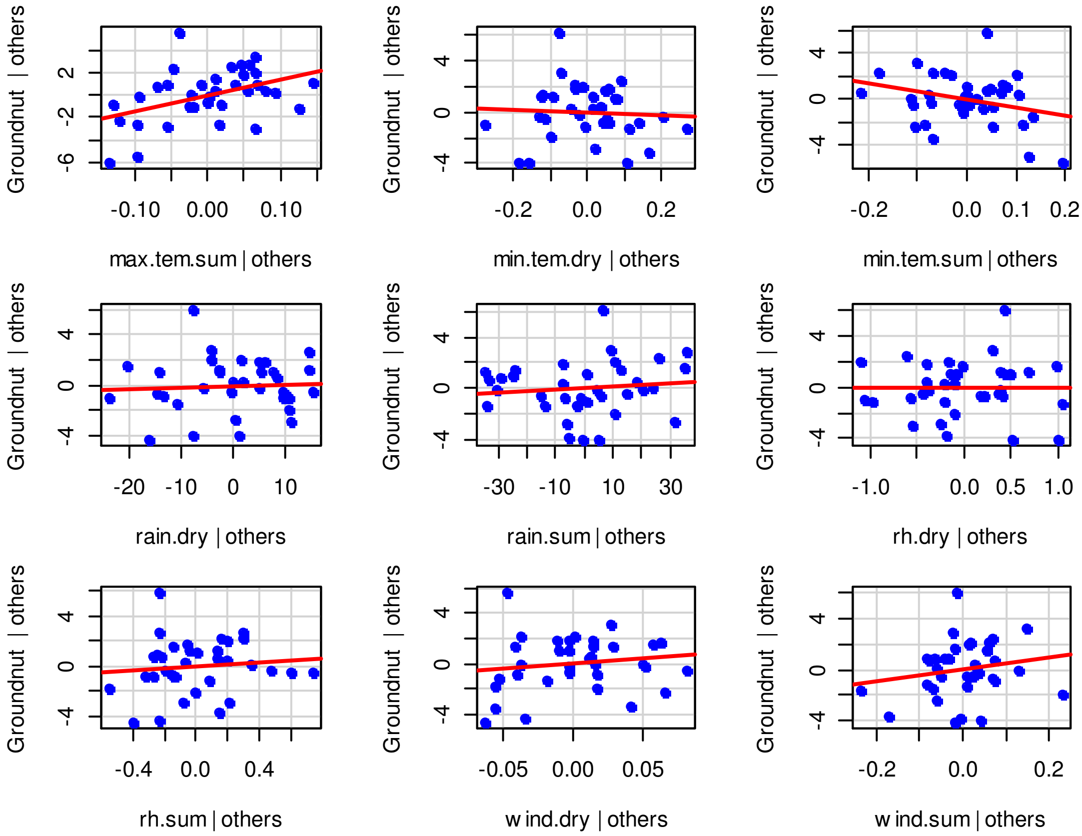


Figure 8: Added Variable Plots of Groundnut Production Model

From the Figure 8 which displays the partial relationship between the response (Groundnut) residuals and each of the predictor's residuals of Groundnut production model. All plots show that they follow a staright line with non-zero slope and there is no curvature relationship ammong the predictor's residuals and response's residuals. That is why, it can be said that each of the predictor variables are added to the model with maintaining a linear relation, that is, this model is going to make a linear relationship between the response variable and the predictor variables to measure the climatic effects on Groundnut production in Bangladesh.

Residuals Diaagonstic Plots are shown in the Figure 9.

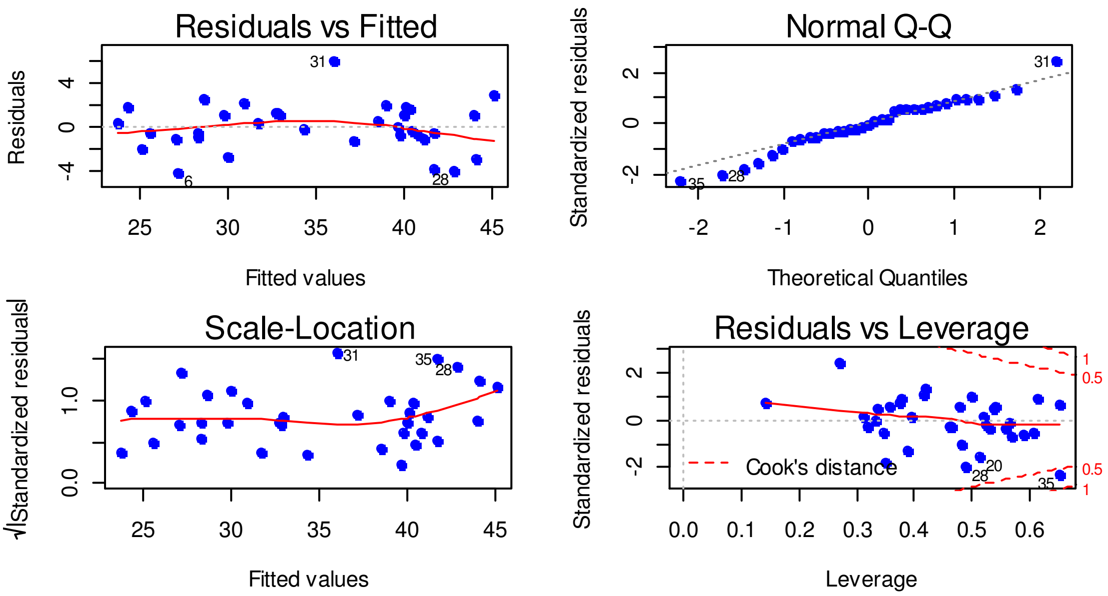


Figure 9: Residuals DiaagonsticPlots of Groundnut Production Model

From the Figure 9, we observe that

- All of the points are lied around the horizontal line and created a horizontal band which implies constant variance among the residuals of the Groundnut production model (top-left).

- almost all of the points try to create a horizontal band which indicates that residals have constant variance of the Groundnut production model (bottom-left).
- almost all of the points are very closed to Q-Q line or on the Q-Q line which suggests that residuals are normally distributed of the Groundnut production model (top-right).
- There is a single leverage points and according to the cook's distance, it lies on the approximately 50% Cook's interval of the leverage points indicating it has a bit of influence on the model properties of the Groundnut production model (bottom-right).

Global model validation test is used to check whether Groundnut production model is valid or not. The test is performed at 5% level of significance on 4 degrees of freedom. The results from the test are shown in the Table 6.

Table 6: *Global Validation Checking of Groundnut Production Model*

Test Statistic	Value	p-value	Decision
Global Stat	5.70221	0.22252	Assumptions acceptable.
Skewness	0.09584	0.75687	Assumptions acceptable.
Kurtosis	0.61439	0.43314	Assumptions acceptable.
Link Function	3.39343	0.06546	Assumptions acceptable.
Heteroscedasticity	1.59854	0.20611	Assumptions acceptable.

From the Table 6, we observe that the p-value of Global Stat of Groundnut production model is 0.22252 which suggests that linearity of parameters, Homoscedasticity, Autocorrelation and Normality assumptions are very well managed in the fitted model, that is, this model is a valid linear linear model. Again, Skewness and Kurtosis of the fitted model's residuals are 0.09584 and 0.61439 respectively and their corresponding p-values for testing hypothesis are 0.75687 and 0.43314 which are suggested that the assumption of the skewness and kurtosis are very well accepted to fit a linear model. Similarly, the Link Function assumptions of Global test is also satisfied to fit a linear model under validation checking of a linear regression model. At the same time, the heteroscedasticity assumptions is also accepted with the p-value = 0.20611 suggesting homoscedasticity of variance. It can easily be said that the fitted model is the best fitted valid Multiple Linear Regression model for measuring the climatic effects on Groundnut production in Bangladesh.

Finally, from all of the test, assumptions of residuals like Homoscedasticity, Autocorrelation, Normality are very well satisfied and model validation test "Global Tesst" also satisfied all of the assuptions for a linear model indicating the fitted model is a valid linear regression model. Without any kind of loss of generality, it can be said that this is the best fitted Multiple Regression Model to measure the climatic effects on Groundnut production based on sample data.

The formal test to check the model's assumptions for all model given bellow in the Table 7.

Table7: Results from the Formal Test of All Fitted Model

Name of The Test	P-values of the Crop Production's Model		
	Groundnut	Linseed	Mustard
Studentized Breusch-Pagan Variance Test	0.7142	0.3101	0.1104
Box-Ljung Autocorrelation Test	0.5414	0.9497	0.8252
Shapiro-Wilk Normality Test	0.285	0.9118	0.3477

- Level of Significance = 5%

From the Table 7, it is clear that residuals of the each fitted Multiple Regression model for Mustard production have constant variance, have no auto-correlation problems and they follow normal distribution at 5% level of significance.

Conclusion and Recommendations

The main objective of the study is to develop the best Multiple Regression Model to measure the climatic effects on different types of oilseed crops production in the Bangladesh. To serve this purpose, Climatic information are divided into two season to measure seasonal effects named as dry season which covers the months October, November, December, January and February; and summer season which covers the months March, April, May, June, July, August and September considering the climatic condition of Bangladesh. To conduct this study, amount of land to the corresponding years is the most vital variable which should be considered. From the study, it is obvious that this variable have statistically positively significant effects on the corresponding production. That is why, we consider this variable in the study. From the study, it is found that the value of Multiple R-squared of Mustard production model is 0.9858 indicating 98.58% of the total variations can be explained by the climatic variables. Again, the value of Multiple R-squared of Linseed production is 0.9362 which implies that 93.62% of the total variations can be explained by the regressor variables. At the same time, the value of Multiple R-squared of Groundnut production is 0.9039 implying 90.39% of the total variations can be explained by the regressor (climatic) variables under study. It is clear from all these models that they are capable of explaining the most of the variability by the climatic variable, that is, these models are good model. From the global validation test, the p-values for mustard, Linseed and Groundnut production are 0.9247, 0.9691 and 0.22252 respectively. From the global test, it is clear that these fitted models are valid linear regression model. From the study, it is found that sunshine of the dry season has negatively significant effects on mustard production. Again, minimum temperature of summer season has negatively significant effects on Linseed production. So, government or policy maker should keep their eyes on this sector if there be any chance to improve this effect.

After conducting these analyses, the following recommendations can be made such as

- The policy makers and researchers could use these model to make a decision for oilseed crops production under consideration of climatic effects on oilseed production.
- Similar regional models could be further studied to find variations of the models.
- The climatic zone similar to Bangladesh could also be compared in the future studies.

References

[1] Hamjah M.A. and Choydhury M.A.K. (2014), Determinants of Crop Production in Bangladesh: Measuring Climatic & Hydrological Effects of Agricultural Productions in Bangladesh, LAP Lambert Academic Publishing, Germany.

[2] Deressa and Hassan (2007), “Economic Impact of Climate Change on Crop Production in Ethiopia: Evidence from Cross-section Measure”, *Journal of African Economies*, 18(4), 529–554.

[3] Hamjah M.A. (2014), “Climatic Effects on Cotton and Tea Productions in Bangladesh and Measuring Efficiency using Multiple Regression and Stochastic Frontier Model Respectively”, *Mathematical Theory and Modeling*, 4(3), 86-98

[4] Hamjah, M.A. (2014), Climatic Effects on Rice Crop Productions in Bangladesh Using Multiple Regression Model and Measuring Production Efficiency Due to Climates Using Stochastic Frontier Model”, *Journal of Economics and Sustainable Development,* 5(8), 199-218.

[5] Hamjah, M.A. (2014), Measuring Climatic and Hydrological Effects on Cereal Crop Production in Bangladesh Using Multiple Regression and Measuring Efficiency Using Stochastic Frontier Model”, *Mathematical Theory and Modeling,* 4(4), 12-32.

[6] Shafiqur Rahman (2008), “Effect of Global Warming on Rainfall and Agriculture Production’, *International Review of Business Research*, l 4(4), 319-329.

[7] Aggarwal, P. K.: (2003), “Impact of climate change on Indian agriculture”, *J. Plant Biology* 30(2), 189–198.

[8] J. R. Andersen (1979), “Impacts of climatic variability on Australian agriculture: a review”, *Marketing and Agricultural Economics*, vol. 47, pp. 147-177.

[9] Doglas C. Montegomery, Elizabeth A. Peck, G. Geoffery Vining, (3rd edition), (2004), Introduvtion to Linear Regression Analysis, John Wiley and Sons, Inc, New York.

[10] Gujarati D N, (4th editions), (2003), Basic Econometris, McGraw-Hill Companies Inc., New York.

[11] Julian J. Faraway, July (2002), Practical Regression and Anova using R, www.stat.lsa.umich.edu/~faraway/book

[12] Basak, K., Jayanta. (2009). “Climate Change Impacts on Rice Production in Bangladesh: Results from a Model” Unnayan Onneshan-The Innovators, *A Center for Research and Action on Development*.

[13] Breusch, T.S.; Pagan, A.R. (1979). "Simple test for heteroscedasticity and random coefficient variation". *Econometrica* (The Econometric Society) 47 (5): 1287–1294

[14] E. A. and Pena E. H. Slate (2003), Global Validation of Linear Model Assumption.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar