Engineering and Applied Sciences
2016; 1(1): 5-12
http://www.sciencepublishinggroup.com/j/eas
doi: 10.11648/j.eas.20160101.12

Statistical Analysis of Factor Affecting Banana Production in Gamo Gofa District, Southern Ethiopia

Natnael Mamuye

Department of Statistics, College of Natural and Computational Sciences, Debre Berhan University, Debre Berhan, Ethiopia

Email address:
naty.buna@gmail.com

To cite this article:
Natnael Mamuye. Statistical Analysis of Factor Affecting Banana Production in GamoGofa District, Southern Ethiopia. *Engineering and Applied Sciences*. Vol. 1, No. 1, 2016, pp. 5-12. doi: 10.11648/j.eas.20160101.12

Received: December 14, 2015; **Accepted**: December 25, 2015; **Published**: June 20, 2016

Abstract: The banana plant, often erroneously referred to as a “tree”, is a large herb, with succulent, very juicy stem (properly “pseudo stem”) which is a cylinder of leaf petiole sheath’s, reaching a height of 20 to 25 fit (6-7.5meter) and arising from a fleshy rhizome. Starting from early 1980th, banana has been produced and consumed in Ethiopia but different researchers’ shows that the production of banana is declining in the country due to different reasons. This study focused on “Mirab Abaya zone” and has an objective of identifying the factors that affect the production of banana in the region. The subjects who are involved in the study are sample of banana producing farmers in Algae “Kebele” selected using simple random sampling technique. The study uses both descriptive statistical methods such as frequency distribution table and summary measures and inferential statistical methods mainly multiple regression analysis of the Cobb-Douglas production function using OLS technique was used to analyze the data obtained by using self administered questionnaire. As a result, the age of banana plants, family size, age of farmers and amount of labor force that used for banana farm were found to be statistically significant predictors of the production of banana in the region. Also in this study factors like gender, educational level of farmers, farm soil fertility, and distance from house to farm and amount of fertilizer that used on banana farm have no statistically significant impacts on the production of banana. Finally, the researchers recommended that it is the duty and responsibility of agricultural office to introduce new varieties of banana and to create awareness about the production of banana to the farmers to increase the productivity of the plant. Also farmers should replace the aged banana plant by the new one and the number of people who can take care of the plant has to be also increase to raise the productivity of the banana in the area.

Keywords: Factors, Banana, Regression, Douglas Production Function

1. Introduction

1.1. Background of the Study

The banana plant, often erroneously referred to as a “tree”, is a large herb, with succulent, very juicy stem (properly “pseudo stem”) which is a cylinder of leaf petiole sheath’s, reaching a height of 20 to 25 fit (6-7.5m) and arising from a fleshy rhizome. Suckers spring up around the main plant forming a clump (“stool”), the eldest suckers replacing the main plant when it fruits and dies and this process of succession continue independently. Tender, smooth, oblong, fleshy stalked leaves, numbering 4or5to15, are arranged spirally. They unfurl, as the plant grows, at the rate of one per week in worm whether, and extend upward and outward, becoming as much as nine fit (2.75) meter long and two fit (60cm) wide. It may be entirely green, green with maroon splotches, and green on the upper side [1].

According to Rivera [2], the ideal conditions for banana growing are the following:

- Soil has to be deep, friable, and rich in organic matter with complete nutrient and mineral elements, and has adequate moisture throughout the year; soil texture of 40% clay, 75% silt or 85% loam; soil pH of 6.5, soil topography of flat to rolling lands up to 45 degrees gradient;
- Tropical climate with temperature range of 22 to 32 degrees Celsius;
- Land elevation from sea level to 1,000 meters above; and.
- Minimal air movement.

Inputs in banana farming include suckers, corms and eye buds obtained from corms that can be used as planting materials; and fertilizer and labor that are utilized for land preparation and planting, fertilization, pruning, thinning, weeding and cultivation, flower and fruit management and harvesting.

Banana ranks fourth as the world's most important starch crop after cassava and sweet potatoes since. Its yields of carbohydrates per unit area are very high. It is the fourth most widely-grown food crop after rice, wheat and maize [3]. Approximately one-third of the bananas produced globally are grown in sub-Saharan Africa, where the crop provides more than 25% of food energy requirements for more than 100 million people. East Africa (including Burundi, Kenya, Rwanda, Tanzania, and Uganda) is the largest banana producing and consuming region in Africa. Uganda is the world's second largest producer after India, with a total of about 10.5×10^9 Kilogram per year [4]. It is estimated that over 75% of the country's farmers grow bananas on 1.5 million hectares, an equivalent of 38% of the total land under crops [5]. But in Uganda banana productivity has been declining, from more than 18kg, per bunch annually [5]. The average shift and productivity decline has been attributed to the increasing severity of production constraints. Particularly the declining of soil fertility, pests and diseases are the major causes for the reduction of banana production in some areas [6].

The time banana was introduced to Ethiopia is before 1980s. Until the early 1980s, maize, cotton and sweet potato were important crops produced by farmers in Arba Minch Zuria and Mirab Abaya districts of the Gamo Gofa zone in Southern nation nationalities and peoples region. During that period, the Arba Minch state farm had 62 hectare of land covered by dwarf Cavendish banana. Experts in the office of agriculture at Gamo Gofa Province made efforts to introduce banana to the "Lante producers cooperative", but it failed as the cooperative administrators at that time did not perceive banana as an important cash crop.

In 1984, a few experts restarted a dialogue to transform the mainly cereal based subsistence smallholder agriculture to a more market-oriented system by introducing irrigated banana. After repeated discussions with cooperative leaders and extension staff, banana was introduced on 4.2 hectare of the cooperatives land. Planting materials came from the state farm. However, the cooperative leaders and members were not fully convinced of banana's potential to improve their livelihoods; this was further fueled by the belief that banana impacts biological fertility if consumed. Continuous awareness creation about the benefits of banana changed the perception of cooperative members. Once the first introduction was made to Lante cooperative, scaling out to individual plots in Lante and Chano Mille Peasant Associations was carried out. Prisoners from Arba Minch helped transport and transplant suckers from the state farm to the farmer plots. The extension staff and administrators continuously monitored the pilot farmers.

The agro-ecology of Arba Minch was good for Cavendish banana; most of the suckers at the pilot farms bore bunches easily and gave good yields some 10 months after planting. The first harvest was transported and marketed in Addis Ababa. The farm gate price for a kilo of banana was 0.20 Ethiopian birr. Seeing this, the farmers understood the economic benefit of engaging in irrigated banana production. About five years later, most farm lands that had easy access to irrigation in Arba Minch were covered by dwarf Cavendish.

Currently banana in Ethiopia covers about 59.64% (53,956.16 hectares) of the total fruit area, about 68.00% (478,251.04 tones) of the total fruits produced, and about 38.30% (2,574,035) of the total fruit producing farmers [7]. On the other hand, about 68.72% (37,076.85 hectares) hectares of land covered by banana, about 77.53% (370,784.17 tones) of the banana produced and 22.38% (1,504,207) of the banana producers in Ethiopia are found in the Southern Nations Nationalities and Peoples' Regional State (SNNPRS) [7]. Gamo-Gofa, Bench Maji and Sheka zones are among the major banana producing zones of the SNNPRS, of which Gamo-Gofa zone alone covers over 70% of the total banana marketed across the major market outlets in Ethiopia [6].

The major commercial cultivars grown by small-scale growers across the survey areas are Dwarf Cavendish, Giant Cavendish, and Poyo, with Williams and Grand Naine recently coming into picture in Gamo-Gofa zone and across the large-scale commercial farms of Ethiopia. Others like Robusta and Butuzua are also among the recently released Cavendish banana cultivars released through Melkasa Agricultural Research Center. The rest are less popular land races grown to a very limited extent in certain localities across the country [8].

Although in Ethiopia, the living standard of the local value chain actors (banana producers, brokers, traders, retailers) and service providers (cooperatives, transporters) has substantially improved in recent years, they still face many challenges. For instance, yield per unit area of land is declining due to improper agronomic techniques such as overstocking, lack of soil amendments, improper irrigation techniques and mono-cropping. Previous studies in the area indicate that declining soil fertility has caused yield loss of 30-60%. Pests (fruit flies) and diseases (Banana xanthomonaswilt and Fusariumoxysporum*)* are potential challenges for banana production although they are not yet severe.

As a few studies are available on factors affecting banana production in the country, this study is intended to fill the gap by making statistical assessment on banana production in Gamo Gofa zone, Mirab Abaya woreda, in Algae Kebele. Thus, in this study factors that can possible affect banana production in the Kebele have been identified. Different statistical methods were used to analyze the primary data which have been obtained through questionnaire. Descriptive statistics was used to summarize the important features of the sampled population. Furthermore, the cob Douglas production function transformed into multiple linear regressions model was used to identify the factors that can affect banana production in the study area.

1.2. Statement of the Problem

The banana farmers (producers) and other district stakeholders show that the production of banana is declining from year to year in different banana producing areas in Ethiopia. Declining of banana yield is a great concern to the farmers, because it is source of food security in many homes in the country, and is also the major income generating crop for the producers. To address the yield gap of banana in different areas of the country, there is a need to researcher to find out areas specific causes that have led to this yield reduction in the country specifically in Mirab Abaya zone Algae Kebele. Thus this study focus on Gamo Gofa zone and try to address the following research questions.

- How good is the production of Banana in the study area?
- What are the major factors that influence banana production in Gamo Gofa district ALGEA Kebele?

1.3. Objective of the Study

The general objective of this study is to identify factors that affect the production of banana in algae Keble.

The specific objectives of this study are:

- To assess the overall production of banana in the study area using descriptive statistics.
- To apply the transformed multiple linear regression model to identify the major factor that can affect banana production in Gamo Gofa district ALGEA Kebele.

2. Methods

2.1. Description of the Study Area

The study area implies the place where the study is conducted; this study was conducted in Southern Nation, Nationalities and Peoples region, GAMO GOFA ZONE, MIRAB ABAYA WOREDA, in ALGEA Kebele. The area was situated 450km south of capital city of Ethiopia and it is also 280km far from HAWASSA which is the capital city of SNNPR. Algae Kebele is bounded by ABAYA HAYIK in the east, BIRBIR town in the west, mole Kebele in the south and DOSHE Kebele in the north. It is in Great Rift Valley with average temperature 30degrees Celsius and annual average rain fall of 575milli meter. More over general elevation of the zone range from 60-330 meter above sea level. Weather condition of this area is considered as hot during winter season as compared with the other area in the southern regions of Ethiopia.

2.2. Sampling Techniques and Sample Size Determination

2.2.1. Target Population

Target populations of this study were those households who grow banana in GAMO GOFA ZONE, in MIRAB ABAYA WOREDA, in Algae Kebele.

2.2.2. Sampling Techniques

Sampling technique is a systematic or a system of taking small ratio of observation from large population with the aim of getting information of this large population from the small observation by using some statistical techniques [9]. In simple random sampling every unit in the population has an equal chance to be included in the sample population. In this study simple random sampling was used to select the entire sample because of the homogeneity of the population with respect to the characteristics of interest within the Kebele. That is the farmers in the Kebele receive the same counsel from agricultural experts, have equal access to rainfall, temperature and fertilizer etc.

2.2.3. Sample Size Determination

Even if there are a number of sample size determination formula approaches (like personal judgments and budgetary approach), the one in which the investigator interesting is based on precision with some confidence level. An appropriate sample is one of the means of gaining high precision and greater accuracy with minimum cost. That is in addition to the purpose of the study and population size, three criteria usually will need to be specified to determine the appropriate sample size: the level of precision, the level of confidence or risk, and the degree of variability in the attributes being measured [10].

The Level of Precision

The level of precision, sometimes called sampling error, is the range in which the true value of the population is estimated to be.

Degree of Variability

The degree of variability in the attributes being measured refers to the distribution of attributes in the population. The more heterogeneous a population, the larger the sample size required to obtain a given level of precision. The less variable (more homogeneous) a population is, the smaller the sample size will be.

Confidence level

There is always a chance that the sample you obtain does not represent the true population value. This risk is reduced for 99% confidence levels and increased for 90% (or lower) confidence levels. Thus to determine a sample size for this study, a results from previous study were used to estimate average and variance of yield of banana [8]. The study uses simple random sampling technique and accordingly the sample size determination formula adopted is given by [11].

$$no = \frac{(Z\alpha/2)^2 S^2}{e^2 \bar{X}^2} \quad (1)$$

Where no is the sample size, $Z\alpha/2$ is the value of standard normal distribution that gives an area of $\alpha/2$ to the right of it, e is the desired level of precision (in the same unit of measure as the variance), S^2 is the estimated variance of yield of banana obtained through pilot survey and $\bar{X}$ is the estimated mean of yield of banana. Thus using $\alpha = 5\%$ (95% confidence level), $e = 3\%, S^2 = 32.37$ and $\bar{X} = 19.5$ the sample size for this particular study is computed as,

$$no = \frac{(1.96)^2(5.69)^2}{(0.03(19.5))^2} = 364$$

Since the sample size is large relative to the population size, it can be adjusted by using the following equation,

$$n = \frac{no}{1+\frac{(no-1)}{N}} \tag{2}$$

$$n = \frac{no}{1+\frac{(no-1)}{N}} = \frac{364}{1+\frac{(364-1)}{2000}} = 308$$

Thus finally 308 farmers have been selected using random number table from the area to represent all household in the region.

2.3. Method of Data Collection

The study used cross-sectional primary data that were collected using a structured questionnaire administered by the researcher with the assistance of enumerators in face to face interviews. The interviews were supplemented with on-farm observations to harmonize the responses given. The data collected covered yield of banana per hectare and different farmers' socio-economic and demographic characteristics that included: age and gender of the household head, size of land (hectares), labor used in banana plots, farming experience (years), planting material cost, types of banana cultivar grown, intercrops, topography, years of education and farming experience, age and gender of the banana grower, household size and distance of farm to residence of the banana grower. The data were collected from in January, 2015. The data from questionnaires were entered into SPSS and cleaned to eliminate errors and then analyzed.

2.4. Method of Data Analysis

After data collection is over, the collected data was carefully edited, coded, tabulated and organized depending on the type of questions and the nature of the data before analysis. The data gathered from the sample was analyzed by using both descriptive statistical methods and inferential statistical methods.

Descriptive statistics: it is statistical methods that helps us to summarize a given sets of data using tables, diagrams, graphs and summary measures such as mean, median and standard deviation [9]. For this particular study mainly frequency distribution table was used to summarize the most important features of the sample data.

Inferential statistics: it is statistical method that helps us to make inference about the population parameters depending on the results obtained from sample data. Making statistical estimation and conducting statistical hypothesis testing are methods through which inferences are made. This study have used the non linear (intrinsically linear) model called the Cobb Douglas production function which have been later transformed into multiple linear regression model to identify the most significant predictors of banana production in MIRAB ABAYA zone ALGEA Kebele. Zenebe [8] have also used multiple linear regressions using Cobb-Douglass production function in order to identify the determinants of the banana productivity.

Regression statistical method is a conceptually simple method for investigating functional relationships among variables. The relationship is expressed in the form of an equation or a model connecting the response or dependent variable and one or more explanatory or predictor variables [12].

Multiple linear regression models: is a model that involves more than one explanatory variable. The goal of multiple linear regression models is to model a linear relationship between the response variable and several independent variables. The multiple linear regressions assume that the response variable is a linear function of the model parameters. The theory which involves production state that there is no direct linear relationship between the response variable (yield) and the independent variables (law of diminishing marginal returns). The law of diminishing marginal returns states that as units of a variable input are added to units of one or more fixed inputs, after a point, each incremental unit of the variable input produces less and less additional output [13].

In this study multiple linear regressions model using the cob-Douglas production function was used which can be given by the following model [13].

$$Y = \alpha X1^{\beta 1} X2^{\beta 2} X3^{\beta 3} \ldots . Xk^{\beta k} \varepsilon \tag{3}$$

By taking the logarithmic transformation, the above function can be transformed into linear model,

$$\ln Y = \ln \alpha + \beta 1 \ln X1 + \beta 2 \ln X2 + \beta 3 \ln X3 + \cdots + \beta k ln Xk + \ln \varepsilon \tag{4}$$

Where ln=logarithm to base 10.

Yi=yield in kilogram,

Age of banana in year, amount of labor used, amount of fertilizer used, soil type (soil fertility), educational status of farmers, and the distance from residence to farm in km, family size, age of farmers are all the independent variables.

ε is assumed to be a random error terms and it accounts for the failure of the model to fit the data exactly. Vector β are production coefficients or elasticity of production of the explanatory variable $X1, X2 \ldots Xk$; are unknownconstants to be determined (estimated) from the sample data using the least square estimation technique. Least square estimation technique estimates unknown parameters by minimizing the sum square of distance of each value from the fitted value and α is ultiplicative constant or intercept of the production plane.

Assumption of the multiple linear regression models

The validity of a statistical method, such as regression analysis, depends on certain assumptions. Assumptions are usually made about the data and the model. The accuracy of the analysis and the conclusions derived from an analysis depends crucially on the validity of these assumptions. For regression model the following assumptions are to be addressed.

Assumption about the form of the model

Linearity: The model should specify a linear relationship between dependent variable and the predictor variables. The linearity assumption is not as narrow as it might first appear. In the regression context, linearity refers to the manner in which the parameters and the error term enter into the equation, not necessarily to the relationship between the variables.

Assumptions about the error term

1. ***Normality:*** The error term is a random variable and should have a normal distribution with mean vector equal to zero and variance-covariance matrix (δ^2). Symbolically, this is shown as$\mathcal{E} - N(0, \delta^2)$.
2. ***Homoskedasticity*:** The error variance (δ^2) is assumed to be constant. The variance should not increase as asthevalueoftheindependentvariablesincreases.
3. ***Non autocorrelation:*** The error term should also have no correlation to each other. The errors associated to different observations should be independent to each other. The deviation of$yi$from the line should not affect the deviation of yj from the line. That is correlation ($\mathcal{E}i,\ \mathcal{E}j$) =0.

Assumption about the independent variables

The X matrix should be full rank. That is the independent variables must be linearly independent to each other

Measuring the quality of fit (assessing model adequacy).

After fitting a linear model relating Y to *X's,* we are interested not only in knowing whether a linear relationship exits, but also in measuring the quality of the fit of the model to the data. The quality of the fit can be assessed by one of the following highly related mechanisms

i. *Testing the regression coefficients associated with all the independent variables*

Testing the null hypothesis that all the regression coefficients are zero against the alternative which state that at least one of the regression coefficients are different from zero. Rejection of the null hypothesis indicates the adequacy (quality) of the model to fit the data very well.

ii. Determining the coefficient of determination

Coefficient of determination is the percentage of the total variation in the response variable which is explained by the presence of the independent variable in the model. Coefficient of determination has values in between 0 and 1 and can also be expressed in percent having values ranging from 0 to 100. Usually high value of R^2 indicates the quality of the model to fit the data very well and low values of R^2 indicates the poor model to fit the data.

iii. Checking if there are any violations of regression assumptions

Adequacy of the model is also related with the regression assumptions set. We reemphasize that the regression assumptions should be checked before drawing statistical conclusions from the analysis (e.g., conducting tests of hypothesis or constructing confidence or prediction intervals) because the validity of these statistical procedures hinges on the validity of the assumptions.

3. Results

One of the most challenging or rewarding task in survey research comes after the data have been collected and coding, editing and preliminary processing have been completed. This section contains statistically analysis results and interpretation of the findings.

Summarization of most important characteristics

A data fom sample of 308respondents in Mirab Abaya districtwas collected to achieve the main objective of this study. The different characteristics of the sampled population were summarized by using measures such as mean,standard deviation and variance if it is continuous in nature. For categorical characteristics of the sampled population, a table which contains percent was used to make summary of it.

Table 1. *Descriptive statistics for some selected continuous variables, Samara 2015.*

Characteristics	Minimum	Maximum	Mean	Std. Deviation	Variance
age of farmers	20	58	36.18	8.379	70.202
family size	1	8	3.23	1.628	2.650
year of banana plant	2	12	5.32	1.792	3.210
the amount of yield of banana in ton/hectare/year	15	41.49	23.45	9.68	93.7
distance from house to farm in km	1.00	7.00	2.7142	1.30123	1.693
farm size in hectare	.25	1.50	.5151	.33990	.116
the amount of fertilizer that used in farm	20	100	39.35	16.636	276.752
the amount labour force that used for banana farm	1	6	3.17	1.353	1.831
the amount of birr that you get from your farm	300	4800	2119.83	958.529	9.188E5

The above table indicates that the minimum age of farmers involved in the sample data is 20 years while the average age of the farmers participated in the sample is 36.89 years. The minimum age of banana plant is 2year and also the average age of banana plant is 5.32 years. The minimum amount of yield is 15 ton per hectare. The average amount of yield is 23.45 ton per hectare and the standard deviation of amount of yield is 9.68 ton per hectare. Farmers on average cover their 0.339 hectare land by banana plant. Furthermore the average number of person who spent taking care of their banana plant is 1.353.

Table 3 summarizes the categorical characteristics by using percents. It can be observed that 52.6% of the respondents were male. Surprisingly about 14.7% of the respondents have attended their education at least up to college level. Most of the farmers plant a type of banana called Kenya banana (76.7%). In addition to that 55.2% of the sample farmers told us they have used traditional fertilizer.

Analysis of inferential statistics/

Multiple regression; yield of banana versus age of farmers, gender…

Table 2. *Descriptive statistics for some selected categorical variables, Samara 2015.*

Attribute	Categories	Percent
Gender	Male	52.6
	Female	47.4
Educational status	Uneducated	19
	Elementary	32.8
	High school	33.6
	College and above	14.7
Banana varieties	Cooking banana	6.9
	Kenya banana	76.7
	White banana	10.3
	Habesha banana	6.0
Use of traditional fertilizer	yes	55.2
	no	44.8

Multiple linear regressions using the cob Douglas production function was applied to assess the relationship between yields of banana which is continuous dependent variables with the selected independent variables. IBM SPSS statistics 20 was used to perform the multiple regression model using step wise variable selection methods. Before the main parameter estimation table is presented the summary statistics of the model is given the table below.

The value of Coefficient of determination is 0.877 which indicates that 87.7% of the total variation in yield of banana is attributed by the factors which are included in the model. This large value of coefficient of variation is an indication for the model to fit the data very well. On the contrary the standard error of the estimates is not small. This may indicate some problem in the model if it does; it will be find out during assumptions checking.

Table 3. *Model summary statistics, Samara 2015.*

Coefficient of determination	Adjusted Coefficient of determination	Standard Error of the Estimate
.877	.862	62.335

Table 4. *Regression results showing the effects of selected explanatory variables on banana yield using Cobb-Douglas functional form, Samara 2015.*

Attributes	Estimates	Standard error of the estimates	t-value	P-value	VIF
(Constant)	201.219	55.153	3.648	.000	
age of farmers	-1.955	.869	-2.250	.027	2.789
Gender(female)	17.109	13.672	1.251	.214	1.133
educational level	6.492	7.162	.907	.367	1.483
family size	14.635	4.720	3.101	.002	2.792
year of banana plant	-13.719	4.498	-3.050	.003	1.136
what is your farm soil fertility	1.804	8.648	.209	.835	1.263
distance from house to farm in km	-4.755	5.195	-.915	.362	1.347
farm size in hectare	5.983	32.017	6.434	.000	1.602
the amount of fertilizer that used in farm	.319	.413	.773	.441	1.154
the amount labor force that used for banana farm	4.742	6.911	6.474	.000	1.198
cooking banana(1)	-8.684	22.745	-.382	.703	2.649
nich banana(2)	25.241	18.658	1.353	.179	1.832

P-value for analysis of variance table=0.000.

Statistical test of individual predictors

The fourth column of the above table provides us the P-value associated with each coefficient. Compare the P-value with the given level of significance to test the significance of the individual predictors to the model. The hypotheses that are used to test individual predictors is stated as

Ho; $\beta i=0$ (the i^{th} factor is not statistically significant predictor of yield of banana)

Ha; $\beta i\neq 0$ (the i^{th} factor is statistically significant predictor of yield of banana)

Decision rule: Reject Ho if the P-value is less than the pre-selected level of significance (5%). since the P-value associated with age of farmers (0.027) is less than the selected level of significance (0.05), there is no enough evidence not to reject the null hypotheses. Therefore it can be conclude that age of farmer is statistically significant predictor of yield of banana in Mirab Abaya district. The coefficient-1.955 shows that for a one year increase in age of farmers, the production of banana is decreased by 1.955 percent assuming the effect of other independent variables holding constant.

The corresponding p-value for the amount of labor force is 0.000 which is less than the preselected level of significance. The null hypothesis is rejected and it can be conclude that the amount of labor force that used on banana farm has statistically significant impact on the yield of banana. Because a one percent increases in family labor, leads to an increase in banana output by 4.74 percent.

The corresponding p-value for year of banana plant is 0.003 which is less than α-value (0.003< 0.05). Thus we have enough evidence to reject the null hypothesis and conclude that year of banana plant has statistically significant impact on the yield of banana. The coefficient year of banana plant is-13.719 which shows that if the banana plant increase by one year, the production of the plant decrease by 13.71 percent assuming the effect of other independent variables holding constant similarly the corresponding p-value for family size is 0.002 which is less than the pre-selected level of significance (0.002<0.05), thus we have enough evidence to reject the null hypothesis and conclude that family size has statistically significant impact on the production of banana. The coefficient 14.635 shows that a one percent increase in family size, leads to an increase in banana output by 14.635 percent where all other variables hold constant.

The p-value for farm size is also less than the pre-selected level of significance indicating that farm size is statistically significant predictor of yield of banana. The coefficient is also positive which implies that as farm size of farmers increases by one percent, yield of banana also expected to increase by 5.9 percent keeping all other independent variables constant.

The p-values associated with gender, amount of fertilizer, farm soil fertility, educational level of farmers and the distance from house to farm are all not less than the pre-selected level significance. Thus gender, amount of fertilizer, farm soil fertility, educational level of farmers and the distance from house to farm are all found to be statistically insignificant predictors of yield of banana at 5% level of significance in Mirab Abaya district.

Testing the overall model adequacy

The analysis of variance table gives us the sources of variation in a regression model with their respective degree of freedom, sum of squares of variation, mean square of the variation, the corresponding F-statistic and the p value. These table is used to test the overall model adequacy at 5% level of significance. The hypothesis

Ho; all the model parameters are zero (the model is not adequate).

Ha; at least one of the model parameter is different from zero (the model is adequate).

The p-value for the analysis of variance is (0.00) which is less than the pre-selected level of significance (5%), the null hypothesis is rejected and it can be conclude that the model is adequate to fit the data.

Checking the assumptions of multiple linear regressions

Test for Multicolinerity.

One of the strong assumptions that lie under MLR model is no Multicolinerity. It means that the independent variables must not be correlated with each other. One method of detecting Multicolinerity is variance inflation factor (VIF). If the VIF of each independent variable is less than 5, there is no correlation among the explanatory variables in the model [16]. From the table 4 of the last column, one can see that the variance inflection factor (VIF) of each explanatory variable is less than 5; this implies that there is no Multicolinerity problem in the model

Residual plots

The residual plots such as histogram of the residuals and scatter plot of fitted values against the residuals have been used in this study to check if there is any violation of the model assumptions. Both plots indicates that there is no series model defects in the above multiple regression model.

4. Discussion

In previous sections, the results obtained from both descriptive statistics and the multiple linear regression using cob-Douglas production functions were presented. This section tries to discuss the finding of this paper by relating with other similar works

The result of this paper indicates that the average amount of yield was23.45 ton per hectare with standard deviation 9.68 ton per hectare in the sample data. This result has slight difference with the result presented by Zenebe [8] which found that the average yield of banana to be 19.5 ton per hectare with standard deviation 5.69 in the sample data selected from some banana producing area in Ethiopia.

This paper also found out that age of farmer is statistically significant predictor of yield of banana in Mirab Abaya district. The negative coefficients indicate that there is negative relationship between age of farmers and banana production in the study area. This result is well supported by Kainga [17] who studied the socio economic determinants of banana production in Bayelsa State in southern Nigeria. However, the same study also conclude that educational attainment and family size were negatively related to the output of banana in the Bayelsa state as opposed to the result of this study which conclude that educational level of farmers have no impact on banana production and family size has positive impact on banana production in Mirab Abaya district.

Age of banana plant is statistically found to be significant with negative coefficient indicating that as the banana plant gets older; its production performance will decrease in some extent. This conclusion was also shared by Tekle [14] which reported that lack of improved varieties has been one of the primary sources of lower banana production in Jinka, south western Ethiopia.

The coefficient of the dummy variable representing gender in banana production suggested that gender of household head is not statistically significant in production of bananas. This is similar to previous studies in Kenya which found gender of the household head not statistically significant in determining banana production in the study area [15].

Furthermore farm size as expected is found to be statistically significant predictor of banana production but land fertility and amount of fertilizer were both found to be statistically insignificant predictor of banana production. This is may be related with the homogeneity of the land in fertility in the area.

Several studies have identified the determinants of yield for banana production. Apart from inputs of production, other factors have influenced yield levels. Yamano [18] examined the integration of dairy and banana farms in Uganda and the effects of selected determinants on banana yield. Results of the study revealed that the amount of organic matter, plot size, tenancy, number of male household members, farm assets, land altitude, population density, and rainfall were statistically significant predictors of banana productions.

Bates [19] result's of the multiple regression analysis of the Cobb-Douglas production function using OLS technique also showed that higher utilization of fertilizer and labor; tenure status in favor of owner operators; operation of diversified farms; and establishment of ideal farm-specific characteristics such as clay loam or sandy clay loam for soil type, distance between hills of $\geq$ 20 meter^2 and shorter distance from farm to residence would significantly and positively affect banana yield in oriental Mindoro, Philippines.

5. Conclusions and Recommendation

The result of this study will lead us to conclude and recommend the following points

A result of data fom sample of 308 farmers in Mirab Abaya district revealed that the average amount of yield was23.45 ton per hectare with standard deviation 9.68 ton per hectare. This number is somewhat better than the results obtained from previous study conducted in other part of Ethiopia. Results of the Cobb-Douglas model showed that family size and farm size were found to be positively and statistically significant predictor of banana production in Mirab Abaya district, in Algae Kebele. Whereas Age of banana plant and age of farmers are negatively and statistically significant predictor of banana production in the area.

However gender, land fertility, amount of fertilizer and the distance of the farm from the farmers home are all statistically insignificant predictors of banana production in Mirab Abaya, Algae Kebele. Finally, the researchers recommended that it is the duty and responsibility of agricultural office to introduce new varieties of banana and to create awareness about the production of banana to the farmers to increase the productivity of the plant. Also farmers should replace the aged banana plant by the new one and the number of people who can take care of the plant has to be also increase to raise the productivity of the banana in the area.

References

[1] Bekunda and Seguya S.: Soil fertility management in the banana-based agriculture of central Uganda; farmer's constraints and options: *African Crop Science journal*, vol.7 pp.54-60; 2001.

[2] Rivera, R. A.: Philippine Banana Production and Marketing Available online at http://hvcc.da.gov.ph/pdf/banana_phil_prodn_market.pdf; Accessed on 9, February 2015.

[3] Bates M. B and Flordeliza A. L: Factors affecting yield performance of banana farms inoriental Mindoro, Philippines, J. *ISSAAS* Vol. 16, No. 1:110-120; 2010.

[4] Wachira P. M, Kimenju J. W, Kiarie J. W, Kihurani A. W, Mwaniki S. W and Gathaara V. N.: Incidence of pests and diseases affecting banana in a commercial banana production setting in Kenya, *International Journal of Research In Agriculture and Food Sciences*, ISSN 2311-2476;2013.

[5] Leena T., Maina M., and Steffen A.: Xanthomonas Wilt. Atreat to Banana production to east and central Africa: *The American Phytopathological Society*. Vol. 93 No. 5; 2009.

[6] CFC (Common Fund for Commodities): Development of organic banana production and export in Sudan and Ethiopia to the Middle East and Europe. FC/CC/34/FISGB/10. Appraisal Report, Addis Ababa, Ethiopia; 2004.

[7] CSA (Central Statistical Agency of Ethiopia): Agricultural Sample Survey. Report on Area and Production of Major Crops. Volume I, VII and VIII. Statistical Bulletin 578. Addis Ababa, Ethiopia; 2014.

[8] Zenebe W, Ali M, Derbew B, Zekarias S, Adam B: Assessment of Banana Production and Marketing in Ethiopia, *International Journal of Sciences:* Basic and Applied Research (IJSBAR), Volume 24, No 3, pp 283-307; 2015.

[9] Bluman: Elementary statistics, Astep by step approach, 8th edition; 2012.

[10] Miaoulis, George, and R. D. Michener: An Introduction to Sampling. Dubuque, Iowa: Kendall/Hunt Publishing Company; 1976.

[11] Cochran, W. G.: Sampling Techniques, 2nd Ed., New York: John Wiley and Sons, Inc.; 1963.

[12] Walter A. Shewhart and Samuel S. Wilks: Regression analysis by example, 4th edition, john Wiley&sons, inc., publication; 2006.

[13] David L. Debertin: Agricultural Production Economics, 2nd edition, ISBN 0-02-328060-3; 2012.

[14] Tekle Y., Wondewosen S., Zemach S., Tibebu S., Abraham S., Woineshet S: Adaptability study of banana (Musa paradisiacal var. sapiertum) varieties at Jinka, southern Ethiopia: *American Journal of Agriculture and Forestry*; 2(6): 250-255, 2014.

[15] Shellmith Mugo: Factors Influencing Tissue Culture Banana Output and it Impact on Income in Nyamusi Division, Nyamira North District, Kenya: *International Journal of Sciences: Basic and Applied Research (IJSBAR*), ISSN (Online): 2307-4531; 2013.

[16] Maddala, G. S: Introduction to Econometrics, 3rd edition, John Wiley and Sons, Chich-ester; 2001.

[17] Kainga P. E., Okorji C. E. & Nweze N. J: Socio-economic determinants and productivity in banana and plantain production, *Globaljournal of Biology, Agriculture and health science*, ISSN: 2319 – 5584; 2014.

[18] Yamano, T: Dairy-Banana Integration and Organic Fertilizer Use in Uganda. Discussion Paper: 08-03. GRIPS Policy Information Center. Available online at http://www3.grips.ac.jp/~pinc/data/08-03.pdf. Accessed on 17, sebtember, 2015.

[19] Bates M. B. and Flordeliza A. L: Factors affecting yield performance of banana farms inoriental Mindoro, Philippines, *J. ISSAAS,* Vol. 16, No. 1:110-120; 2010.