Volume VI Number 2, 2014

Alternative Approaches to Technical Efficiency Estimation in the Stochastic Frontier Model

H. D. Acquah[1], E. E. Onumah[2]

[1]Department of Agricultural Economics and Extension, University of Cape Coast, Ghana

[2]Department of Agricultural Economics and Agribusiness, University of Ghana, Ghana

Abstract

Estimating the stochastic frontier model and calculating technical efficiency of decision making units are of great importance in applied production economic works. This paper estimates technical efficiency from the stochastic frontier model using Jondrow, and Battese and Coelli approaches. Simulated data is employed to compare the alternative methods. Empirical results show a strong correlation between the alternative methods regardless of the differences in the actual values of the efficiency estimates. Mean technical efficiency is sensitive to the choice of estimation method. Analysis of variance and Tukey's test suggest difference in means between the efficiency scores from different methods. Battese and Coelli's approach produces more homogenous estimates of technical efficiency when compared with the Jondrow's mean or mode approach. Our results suggest that differences in conclusion are possible when the alternative methods of measuring technical efficiency are applied.

Key words

Stochastic frontier, technical efficiency, Tukey's test, mean and mode Approach.

Introduction

One way to evaluate decision making units (DMU) is by analysing whether or not they use their resources in an economically efficient manner. Economic efficiency is understood in terms of jurisdictions providing a maximum amount of output for a given level of inputs and is one potential means to evaluate DMUs. The most common efficiency concept is technical efficiency. Subsequently, numerous authors (Onumah, Acquah (2011), Onumah, Acquah (2010), Kumbhakar, (2002), Bravo-Ureta, Rieger (1991), Bagi and Huang (1983), Demir, Mahmud (2002), Kirkley, Squires and Strand (1995) have investigated technical efficiency of decision making units. Technical efficiency measures the conversion of inputs into outputs relative to best practice. In other words, given current technology, there is no wastage of inputs whatsoever in producing the given quantity of output.

However, various approaches co-exist to measure the technical efficiency of a decision making unit. For example, Jondrow et al. (1982) and Battese and Coelli (1988) provide alternative approaches to estimating technical efficiency in the stochastic frontier model. Previous studies measuring technical efficiency adopted either Battese and Coelli's approach or Jondrow's approach but not both. An exception is Hoyo et al. (2004) who applied Battese and Coelli's as well as Jondrow's mean approach. However, their study did not consider Jondrow's mode approach. A rigorous comparison of the Jondrow's mean and mode approaches with Battese and Coelli's approach to measuring technical efficiency is lacking in the literature. For researchers to assess the best alternative approach, it is imperative that a rigorous comparison of the methods is provided. Given that the alternative approaches differ methodologically, it is important to assess whether the different approach taken affects the outcome of efficiency studies or lead to differences in conclusion. The comparison of these approaches not only adds to the literature, but also deepens our understanding on inferences that can be derived when alternative methods of technical efficiency estimation are applied in production economics. Therefore this paper aims at measuring technical efficiency in the stochastic frontier model by applying the Jondrow et al. (1982) and Battese and Coelli (1988) approaches to simulated data.

Materials and methods

Technical Efficiency

Technical efficiency is defined as a measure of how well decision making units (DMU) convert inputs to output with a given technology and economic factors (Kumbhakar and Lovell, 2000). It is measured as the ratio of observed output (Y_i) to the corresponding frontier output (Y_i^*) with given levels of input and technology ($TE_i = Y_i / Y_i^*$). Therefore, technical inefficiency exists if a DMU produces below the production frontier. The measurement of technical efficiency and its underlying factors are of critical significance in production theory. Technical efficiency of a DMU and the degree of use of variable inputs determine the output and capacity utilization. Identifying the various factors affecting it allows stakeholders to take measures to limit or improve on it.

The concept of technical efficiency can be explained using a two input (x_1, x_2) - two output (y_1, y_2) production process. Bogetoft (2012) asserts that efficiency could be looked at from the angle where optimal inputs are combined to achieve a given level of output (an input-orientation) and where optimal output could be obtained given a set of inputs (an output-orientation). Grosskopf et al. (1994), points out that both measures provide the same technical efficiency scores when the assumption of constant returns to scale is applied. Technical efficiency in this study is considered from the angle where a DMU minimises the quantity of inputs used to achieve a constant output (an input-orientation). This idea draws from Farrell (1957) and is referred to as Farrell efficiency.

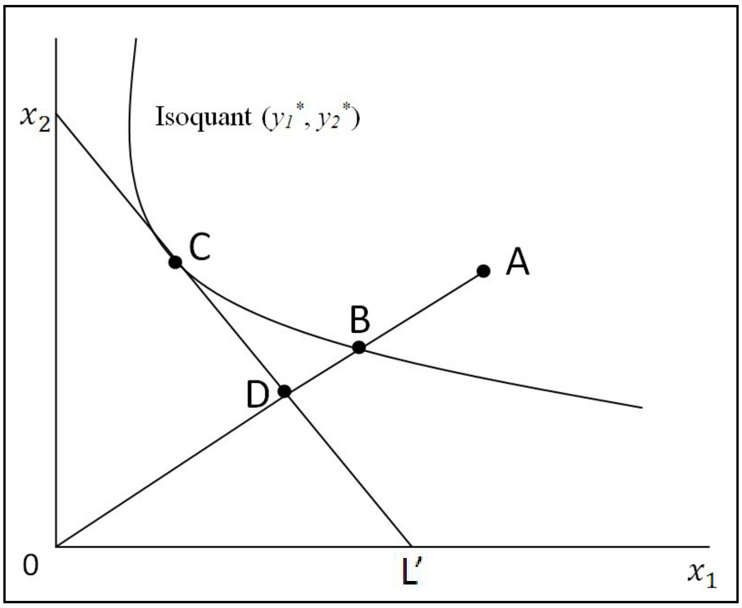
From Figure 1, the firm uses the combination of inputs defined by point A to produce a given level of output (y_1^*, y_2^*). The same level of output could have been produced by combining minimum level of inputs to produce (y_1^*, y_2^*) (Isoquant (y_1^*, y_2^*)). This is defined by point B and it lies on the isoquant. This is because isoquant shows the same level of output that an efficient firm can achieve by combining various quantities of inputs. Therefore, the input-oriented level of technical efficiency (y, x) is defined by OB/OA. The distance BA shows the technical inefficiency of the firm and it represents the amount by which all inputs could be proportionally reduced without a decrease in output. Noticeably, DB is the cost inefficiency and C is the economic efficiency point.

Alternative Approaches in Estimating Technical Efficiency

The stochastic frontier production function can be specified as:

$$y_i = f(x_i;\beta).\exp(v_i - u_i) = 1, 2, 3...n \qquad (1)$$

where y_i a scalar is output, x_i is a vector of inputs and β is a vector of parameters to be estimated. The composed error is made up of and u_i. v_i is



Source: Adapted from Coelli et al. (2005)

Figure 1: Input Oriented Efficiency Measure.

assumed to be independently and identically distributed (iid) and symmetric, and it represents statistical noise in the model. And $u_i \geq 0$ is a one side error term representing technical inefficiency. Note that u_i measures technical inefficiency in the sense that it measures the shortfall of output (y_i) from its maximal possible value given by the stochastic frontier. Technical efficiency can thus be measured as

$$TE_i = e^{-u} \tag{2}$$

Technical effciency of stochastic frontier models may be estimated by a maximum likelihood (ML) procedure. When using the ML procedure, the estimation is based on the mean or the mode of the conditional distribution of the inffciency error for each DMU (JLMS; Jondrow et al., 1982).

The mean is

$$EV(u|\in) = \mu_* + \sigma_* \frac{\text{ø}(\mu_*/\sigma_*)}{\phi(\mu_*/\sigma_*)} \tag{3}$$

Where

$$\mu_* = -\in \frac{\sigma_u^2}{\sigma^2} = -\in \frac{\lambda^2}{1+\lambda^2} = -\in \gamma$$

$$\sigma_* = \sqrt{\frac{\sigma_u^2 \sigma_v^2}{\sigma^2}} = \frac{\lambda}{(1+\lambda^2)}\sigma = \sqrt{\gamma}(1-\gamma)\sigma^2$$

and ø(.) is the density function, and $\phi(.)$ the distribution function of a standard normal distribution. σ_v^2 is variance of the random error v, σ_u^2 is the variance of inefficiency term u, $\in$ is the total error, v-u, $\lambda = \sqrt{\frac{\sigma_u^2}{\sigma_v^2}}$ and $\sigma^2 = \sigma_v^2 + \sigma_u^2$.

When we substitute the estimated values for $\in$, σ^2, and λ then we have an estimate of u, call it $\hat{u}$, conditioned on the estimate of $\in$.

It can also be noted that

$$\frac{\mu_*}{\sigma_*} = -\in \frac{\sigma_u^2}{\sigma_2^2} \frac{\sigma}{\sigma_u \sigma_v} = -\in \frac{\sigma_u}{\sigma_v} \frac{1}{\sigma} = -\in \frac{\lambda}{\sigma} \text{ where } \lambda = \frac{\sigma_u}{\sigma_v}$$

Such that

$$EV(u|\in) = \sigma_* \left(\frac{\text{ø}(\in\lambda/\sigma)}{1-\phi(\in\lambda/\sigma)} - \in \frac{\lambda}{\sigma} \right) \tag{4}$$

The above equation can be simplified to

$$EV(u|\in) = \sigma_* \left(\frac{\text{ø}(\in_*)}{1-\phi(\in_*)} - \in_* \right) \text{where } \in_* = \in \frac{\lambda}{\sigma}. \tag{5}$$

The estimates calculated in Eq. (5) are equal to the estimates calculated in Eq. (3).

Another estimator is the mode of the conditional distribution, which can also be interpreted as a maximum likelihood estimator:

$$M(u|\in) = \begin{cases} \mu_* & for \in \leq 0, \\ 0 & for \in > 0. \end{cases} \tag{6}$$

$$M(u|\in) = \begin{cases} \mu_* & for \mu_* > 0, \\ 0 & for \mu* \leq 0. \end{cases} \tag{7}$$

So that we have

$$M(u_k|\in) = \max(0, \mu_{*i}) \tag{8}$$

As $EV(TE) = EV(e^{-u})$ is generally not equal to $e^{-EV(u)}$ yet another estimator has been proposed in Battese and Coelli, (1988).

$$TE = EV(e^{-u}/\in) = \frac{\phi(\mu_*/\sigma_* - \sigma_*)}{\phi(\mu_*/\sigma_*)} e^{\left(\frac{1}{2}\sigma_*^2 - \mu_*\right)} \tag{9}$$

This estimator is optimal in the sense of minimizing the mean square error. This is the one that is most often used in applied production economics work. The actual values of the efficiency estimates may somewhat differ between the three methods, but very little work has been done to shed light on the estimates based on the three different methods.

Comparing Technical Efficiency in SFA using Alternative Methods

In order to estimate firm specific technical efficiency using alternative methods, we generated simulated data from a stochastic frontier model of the form

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + v - u \tag{10}$$

The simulated data with a sample size of 60 is generated with design parameters as follows: $\hat{\beta}_0 = 2, \hat{\beta}_1 = 0.25, \hat{\beta}_2 = 0.75$. The inefficiency term is generated from the half normal distribution and remaining variables in the model are generated from the normal distribution with sample size of 60. Fundamentally, we denote y as the output of 60 agricultural firms and variables x_1 and x_2 as the input variables of these firms. Subsequently, the efficiency of the 60 firms from the simulated data are computed using alternative technical efficiency methods namely, Battese and Coelli (teBC), Jondrow's Mode approach (teMode) and Jondrow's Mean approach (teJ). For the purpose of brevity, we will denote Battese and Coelli approach, Jondrow's Mode approach and Jondrow's Mean approach by teBC, teMode and teJ respectively in the rest of this paper.

Results and discussion

The results in Table 1 indicate that the efficiency scores of the firms derived using the 3 methods, ranged between 20 to 100%. Clearly, the actual values of the efficiency estimates differ between the three competing methods. Fundamentally, these differences in estimates may be attributed to the methodological differences in the estimators used. At lower levels of efficiency (<50%), teBC obtained 20 firms, teMode obtained 13 firms and teJ had 29 firms. At moderate levels of efficiency (50 to 79%), teBC reported 40 firms, teMode reported 19 and teJ reported 31 firms. At higher levels of efficiency (>80%), teMode recorded 28 firms whilst teBC and teJ had none (0 firms). Importantly, these results suggest that the different technical efficiency estimates provided by the different methods might have different policy implications since they imply different levels of firm capacity. Generally, the different methods lead to differences in conclusion.

The average efficiencies of the three methods are presented in Table 2 below. The average efficiencies tend to differ among the three methods studied. The teMode approach provided a higher mean efficiency of 73.12 this is followed by teBC and teJ approaches with 53.12 and 48.18 respectively. The coefficient of variation (CV) which is defined as the standard deviation expressed as a percentage of the mean is also examined. When a computed CV is less than 33% we say the data set is homogeneous. The teMode method tends to have the largest CV of 34.33%. This followed by teJ and teBC methods with CVs of 23.47 and 19.75 respectively. These results suggest that efficiency estimates from teMode is more variable when compared with efficiency estimates of teJ and teBC methods. Noticeably, teBC efficiency estimates has the smallest variability among the three methods. These results are consistent with Hoyo et al (2004) assertion that the Battese and Coelli approach (teBC) has a higher mean efficiency and a lower coefficient of variation when compared with the Jondrow's Mean approach (teJ).

Model	Mean	S.d	CV (%)
teBC	53.65	10.60	19.75
teMode	73.12	25.10	34.33
teJ	48.18	11.31	23.47

Source: own processing

Table 2: Average efficiencies with standard deviation (s.d) and coefficients of variation (CV) according to the different estimation procedures.

In order to investigate whether there is a significant difference in means between the efficiency scores from different methods, the analysis of variance (ANOVA) and Tukey's HSD (Honest Significance Difference) test were applied. The anova test (p-value=1.07e-13) suggest a significant difference among the three efficiency techniques as illustrated in Table 3. Using Tukey's HSD follow up test indicates that differences exist between teBC and teMode, and teJ and teMode as shown in Table 4.

	Df	Sum Sq	Mean Sq	F value	Pr (>F)
Method	2	20610	10305	35.52	1.07e-13***
Residuals	177	51349	290		

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Source: own processing

Table 3: Analysis of Variance of Technical Efficiency Estimates of Firms obtained with teBC, teMode and teJ.

Percent	teBC Freq		teMode Freq		teJ Freq	
	F		F	C. F	F	C. F
20-29	1	1	2	2	4	4
30-39	6	7	8	10	9	13
40-49	13	20	3	13	16	29
50-59	20	40	7	20	23	52
60-69	18	58	2	22	7	59
70-79	2	60	10	32	1	60
80-89	0	60	5	37	0	60
90-99	0	60	8	45	0	60
100	0	60	15	60	0	60

Source: own processing

Table 1: Frequencies (F) and cumulative frequencies (CF) of efficiency estimates of firms obtained with different estimation methods.

Method	diff	lwr	upr	p adj
teMode - teBC	19.4667	12.1166	26.8167	0.0000
teJ - teBC	-5.4667	-12.8167	1.8834	0.1869
teJ - teMode	-24.9333	-32.2834	-17.5832	0.0000

Source: own processing

Table 4: Tukey's Honest Significance Difference test of Technical Efficiency Estimates of Firms obtained with teBC, teMode and teJ.

Table 5 provides the results of the correlation analysis between the actual values of the efficiency estimates from the three different methods. Though the actual values of the estimates differ among the methods but the estimates based on the three methods are highly correlated. The presence of a strong positive correlation among the efficiency estimates, suggest that the methods can be applied concurrently to provide a holistic view of firm specific efficiency analysis. Similarly, Bogetoft and Otto (2011) notes that the actual values of estimates differ among the methods but the estimates based on the three methods are highly correlated.

	teBC	teMode	teJ
teBC	1.0000	0.9706	0.9978
teMode	0.9706	1.0000	0.9670
teJ	0.9978	0.9670	1.0000

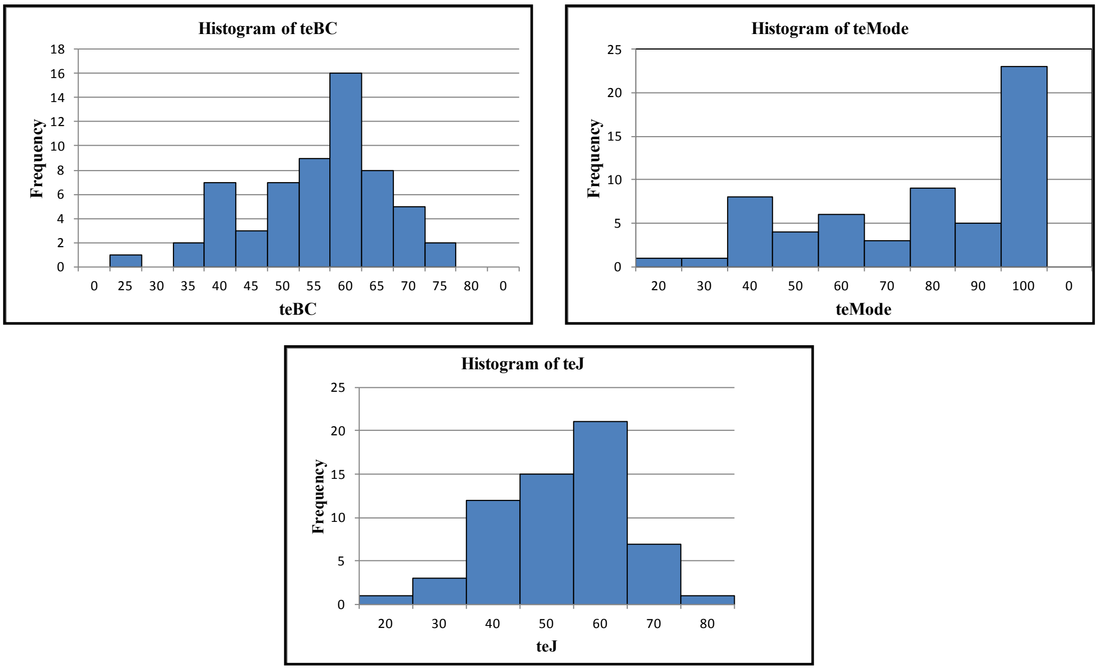
Source: own processing

Table 5: Correlation analysis of the efficiency estimates between the different methods.

The histogram of the efficiency estimates of the 3 methods differ in the shape of their distribution. Noticeably, the histogram of teBC shows a more uniform distribution of efficiency estimates when compared to those of teJ and teMode histograms in Figure 2.

The results of our analysis based on a sample size of 60 were consistent with those based on a sample size of 200 generated from the same stochastic frontier model specified in equation 10. At higher levels of efficiency (>80%), teMode recorded 69 firms whilst teBC and teJ obtained 0 firms respectively given a sample size of 200 as illustrated in Table 6. Similarly, as indicated in Table 1, at higher levels of efficiency (>80%), teMode recorded 15 firms whilst teBC and teJ obtained 0 firms respectively using a sample size of 60.

Using a sample size of 200, teMode approach provided a higher mean efficiency of 61.54 followed by teBC and teJ approaches with 47.21 and 41.41 respectively. Similarly, with a sample size of 60 as indicated in Table 2, teMode resulted in higher mean efficiency of 73.12 followed by teBC and teJ approaches with 53.12 and 48.18 respectively. These results suggest that teMode resulted in higher mean technical efficiency estimates than the teBC and teJ regardless of sample sizes. Furthermore, using the sample size of 200, teMode method tends to have the largest CV of 49.05%. This is followed by teJ and teBC methods with CVs of 35.74



Source: own processing

Figure 2: Histograms of efficiency estimates of the different methods.

Percent	teBC Freq		teMode Freq		teJ Freq	
	F	C. F	F	C. F	F	C. F
0-9	1	1	2	2	2	2
10-19	4	5	12	14	19	21
20-29	22	27	26	40	30	51
30-39	35	62	23	63	33	84
40-49	39	101	16	79	50	134
50-59	54	155	22	101	46	180
60-69	40	195	12	113	17	197
70-79	5	200	18	131	3	200
80-89	0	200	11	142	0	200
90-99	0	200	12	154	0	200
100	0	200	46	200	0	200

Source: own processing

Table 6: Frequencies (F) and cumulative frequencies (CF) of efficiency estimates of firms obtained with different estimation method.

and 30.54 respectively. These results suggest that efficiency estimates from teMode is more variable when compared with efficiency estimates of teJ and teBC methods across the different sample sizes of 60 and 200 respectively. Similarly, the analysis of variance (ANOVA) and Tukey's HSD (Honest Significance Difference) test were applied to investigate the difference in means between efficiency scores from different methods. The results of both the ANOVA test and Tukey's HSD based on a sample size of 200 presented in Tables 7 and 8 suggest a significant difference among the three efficiency techniques.

	Df	Sum Sq	Mean Sq	F value	Pr (>F)
Method	2	42930	21465	48.12	<2e-16***
Residuals	597	266299	446		

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Source: own processing

Table 7: Analysis of Variance of Technical Efficiency Estimates of Firms obtained with teBC, teMode and teJ.

Method	diff	lwr	upr	p adj
teMode - teBC	14.33	9.3677	19.2923	0.0000
teJ - teBC	-5.795	-10.7573	-0.8327	0.0172
teJ - teMode	-20.125	-25.0873	-15.1627	0.0000

Source: own processing

Table 8: Tukey's Honest Significance Difference test of Technical Efficiency Estimates of Firms obtained with teBC, teMode and teJ.

This significant difference among the three efficiency estimates hold across the different sample sizes of 60 and 200. Though the actual values of the estimates differ among the methods, estimates based on the three methods are highly correlated as indicated in Table 9.

	teBC	teMode	teJ
teBC	1.0000	0.9718	0.9980
teMode	0.9718	1.0000	0.9771
teJ	0.9980	0.9771	1.0000

Source: own processing

Table 9: Correlation analysis of the efficiency estimates between the different methods.

Conclusion

Previous research has developed alternative methods of estimating technical efficiency. In this study simulated data is employed to compare the alternative methods of calculating technical efficiency in the stochastic frontier model. The results show that though the actual values of the efficiency estimates differ between the methods, there exists a strong positive correlation among the efficiency estimates based on the three methods. Mean technical efficiency is sensitive to the choice of estimation method. On the basis of analysis of variance and Tukey's test this study finds significant difference in means between the efficiency scores from different methods. Furthermore, the efficiency estimates of the Battese and Coelli's approach has the smallest variability when compared with the Jondrow's mean or mode approach. An implication for efficiency analysis is that the Battese and Coelli's approach is more adequate to provide efficiency estimates with less variability. These results suggest that differences in conclusion are possible when the alternative methods of measuring technical efficiency are applied. These results hold in both small and large samples. Furthermore, the differences in technical efficiency estimates provided by the different methods might have different policy implications since they imply different levels of firm capacity. In the light of the findings it is necessary for further research to extensively investigate the mathematical and intuitive reasons underlying the differences in estimates derived from the different technical efficiency measures.

Corresponding author:
Dr. Henry De-Graft Acquah
Department of Agricultural Economics and Extension
University of Cape Coast, Ghana
Phone:+233 245543956, E-mail: henrydegraftacquah@yahoo.com

References

[1] Bagi, F. S. , Huang, C. J. Estimating technical Efficiency for Individual Farms in Tennesse. Canadian Journal of Agricultural Economics. 1983, 31 (2), p. 249–256.

[2] Battese, G. E., Coelli, T. J. Predictions of firm-level technical efficiencies with generalized frontier production function and panel data. Journal of Econometrics, 1988, 38(3), p. 387-399. ISSN 0304-4076.

[3] Bogetoft, P. Performance Benchmarking Measuring and Managing Performance. Springer, New York. 2012, ISBN 978-1-4614-6042-8.

[4] Bogetoft, P., Otto, L.. Benchmarking with DEA, SFA and R. Springer, New York. 2011. ISBN 978-1-4614-2772-8.

[5] Bravo-Ureta, B. E., Rieger, L (1991) Dairy Farm Efficiency Measurement Using Stochastic Frontiers and Neoclassical Duality. American Journal of Agricultural Economics, 1991, 73 (2). p. 421-428.

[6] Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., Battese, G. E. An Introduction to Efficiency and Productivity Analysis, Springer, New York. 2005. ISBN-13: 978-0387-24266-8.

[7] Demir, N. Mahmud, S. F. (2002): Agro-Climatic Conditions and Regional Technical Inefficiencies in Agriculture. Canadian Journal of Agricultural Economics. 2002, 50(3), p. 269–280. ISSN 0008-3976.

[8] Farrell, M. J. The measurement of productive efficiency. Journal of Royal Statistical Society, 1957, A 120. p. 253-281.

[9] Grosskopf, F. R. S., Lovell, C. A. K. Production Frontiers. Cambridge, MA: Cambridge University Press. 1994. ISBN-13: 978-0521072069.

[10] Hoyo, J. J. G., Espino, D. C., Toribio, R. J. Determination of technical effciency of fisheries by stochastic frontier models: a case on the Gulf of Cadiz (Spain). ICES Journal of Marine Science, 2004, 61, p. 416-421. ISSN:1054-3139, E-ISSN:1095-9289.

[11] Jondrow, J., Lovell, C. A. K., Materov, I. S., Schmidt, P. (1982). On the estimation of technical efficiency in the stochastic frontier production function model. Journal of Econometrics, 1982, 19, p. 233-238. ISSN 0304-4076.

[12] Kirkley, J. E. Squires, D., Strand, I. E. Assessing Technical Efficiency in Commercial Fisheries: The Mid-Atlantic Sea Scallop Fishery. American Journal of Agricultural Economics, 1995, 77 (3), p. 686-697.

[13] Kumbhakar, C. S. Specification and Estimation of Production Risk, Risk Preferences and Technical Efficiency American Journal of Agricultural Economics, 2002, 84 (1), p. 8-22. ISSN: 0002-9092, E-ISSN: 1467-8276.

[14] Kumbhaker, S. C., Lovell, C. A. K. Stochastic Frontier Analysis. Cambridge University Press, Cambridge, UK. , 2000, ISBN-13: 978-0521666633.

[15] Onumah, E. E., Acquah, H. D. A Stochastic Production Investigation of Fish Farms in Ghana in Agris on line Papers in Economics and Informatics, 2011, Vol. III, No. 2, p. 55-65. ISSN 1804-1930.

[16] Onumah, E. E., Acquah, H. D. Frontier Analysis of Aquaculture Farms in Southern Sector of Ghana in World Applied Sciences Journal, 2010, 9 (7), p. 826-835. ISSN: 1818-4952.