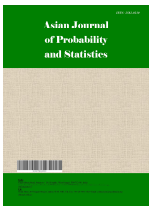

Asian Journal of Probability and Statistics

5(4): 1-10, 2019; Article no.AJPAS.52735
ISSN: 2582-0230

Factors Affecting the Productivity of the Big Onion in Hambantota District during the Off Season

M. K. M. Sulochana[1] and L. S. Nawarathna[2*]

[1]Department of Physical Science, Rajarata University of Sri Lanka, Mihintale, Sri Lanka.
[2]Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2019/v5i430145
Editor(s):
(1) Dr. Jiteng Jia, School of Mathematics and Statistics, Xidian University, China.
Reviewers:
(1) R. Shenbagavalli, India.
(2) Philbert Modest Luhunga, Tanzania.
Complete Peer review History: http://www.sdiarticle4.com/review-history/52735

Original Research Article

Received: 10 September 2019
Accepted: 13 November 2019
Published: 20 November 2019

Abstract

Aim: The main aim of this study is to identify the factors affecting the big onion productivity of Hambantota district during the off-season. Moreover, we identify the average productivity per acre from Hambantota district and compare it with the other areas that cultivated the big onion. Further, identify the main issues encountered in big onion cultivation in Hambantota and identify the critical contributing factors for the big onion cultivation in this area.
Place and Duration of Study: During the off seasons in 2015 to 2016 in Hambantota District.
Methodology: Sample data was collected from 201 farmers in Hambantota district. Multiple linear regression model was used to identify the factors affecting the big onion productivity in Hambantota district during the off-season. The normality assumption of the regression model was checked using Kolmogorov–Smirnov test, Shapiro Wilk normality test and Skewness and Kurtosis test. Pearson, Spearman's Rank and Partial correlation tests were used to check the correlations between variables. Mean absolute percentage error (MAPE) and Symmetrical Mean absolute percentage error (SMAPE) values were used to validate the fitted model.
Results: By the multiple linear regression model main factors affecting the productivity of big onion in Hambantota area were Seasonal Months, Monthly Income, Subsidies Fertilizer and Cultivated Quantity. And the R-squared value was most like to 80% and this means these independent variables were described 80% of the dependent variable. Model accuracies were reported as 98.48% and 98.49% from

**Corresponding author: E-mail: lakshikas@pdn.ac.lk;*

MAPE and SMAPE respectively. Therefore, this multiple linear regression model was suitable for this study. Further, the model determined the affected factors for the big onion cultivation in Hambantota district during the off-season.

Conclusion: Hambantota district average productivity was less than other areas. Big onion productivity of Matale is more than 2 times greater than big onion productivity of Hambantota. Off season big onion cultivation in Hambantota district is not very effective because of the average productivity is less than other areas in Sri Lanka.

Keywords: Big onions; mean absolute percentage error (MAPE); multiple linear regression analysis; productivity; symmetrical mean absolute percentage error (SMAPE).

1 Introduction

Big Onion crop was introduced to Sri Lanka by the British in 1855 and economical cultivation was introduced by the Department of Agriculture (DOA) during the 1950's and over the past years, the crop performance was evaluated in many parts of the country and it was observed that big onions can be grown economically during every Maha Kannaya (peak season) in almost all parts of the country. Big onion one of the important cash crops grown in the dry and intermediate zone of Sri Lanka and has become an important item in our diet. Cultivation of the crop, which has specific environmental requirements, is presently restricted to a few agro-ecological regions of the Dry Zone. The crop is well adapted to the countries, which have a long day (> 12 hours/day) and dry conditions. Short day onions can produce bulbs at low latitudes, where photoperiods remain close to 12 hours throughout the year [1,2]. Production of onions in the tropics is beset due to various problems. These production constraints include biotic stresses, such as diseases, insect pests etc, and abiotic stresses like high temperature, prevailing photoperiod, high rainfall etc [3]. The conditions in Sri Lanka are marginal to the crop compared to the other countries, where the onion is grown. It should be possible, however, to cultivate it in other areas if appropriate production technologies are developed in respect to adapted varieties, fertilizer and water management, and cultural practices [4]. In general, the big onion can be established using seedlings. Use of good quality planting materials is the key to establish a successful crop. Adverse climatic conditions prevailing in the dry zone, critically affect the crop throughout the growing period. As a result, the crop fails even at the nursery stage [5]. Quality of seedling is one of the critical factors determining the crop yield. The cultivation of the big onion requires about 600-person days of labour per hectare (ha). It is estimated 45 million labour units are employed annually in onion production [6]. The onion crop can be established on different types of beds such as raised beds or sunken beds depending upon the soil type and irrigation facilities. Raised beds could also lower the bulb rot. Though the national annual requirement of big onion production is about 235,000 MT but the local production is only 104,000 MT. Therefore, more than 120,000 MT are imported annually to fulfil the gap (131,000 MT in 2015). The total annual imported bill on the onion is Rs. 5,765 million while the total projected expenditure is Rs. 50 million. Thus, increasing onion production will provide gainful employment as well as saving scarce foreign exchange. Total big onion production is only 98,900 MT in 2014 [7,8]. So, Sri Lankan government have successfully cultivated big onions during the off season in the Hambantota District. That is around 600 hectares of land, and they will reap the harvest by the end of May. This has proven that big onions can be successfully cultivated in Hambantota as in other districts. According to the government decision if they are able to spread the cultivation of this crop in both the Moneragala and Hambantota districts in a land area of 10,000 hectares, government will be able to meet 100% of our domestic demand [9]. However, at present the cultivation of the big onion is confined only to Matale, Anuradhapura, Puthalama, Pollonnaruwa, Mahawelli, Jaffna and Hambantota Districts during the Maha Kannaya (peak season). Here Hambantota district cultivated Big Onion in off-season. Hambantota district provided 31, 000 MT big onion annually.

According to the big onion promotion program around 700 ha (hectare) in Hambantota, and 150 ha in Moneragala have been targeted for big onion cultivation during the off season and the total targeted big onion productivity extent was 1,097 ha in Sri Lanka. According to this Sri Lankan government targeted to

achieve the domestic big onion demand by extent big onion cultivation to Hambantota and Monaragala district during off-season [10].

Based on a research about socio economic factors affecting the productivity of green gram in Sri Lanka, the researcher used data transformation, ordinary multiple regression models to determine those socio-economic factors. According to this age, fertilizer cost, labor cost affected to the productivity of green gram [11,12].

From this study farmers can get an idea about their off-season cultivation, and they can identify the factors affecting to their yield. After knowing factors, they can find the solutions for those factors and get more productivity during their off-season big onion cultivation. Further, government can pay attention to those affecting factors and increase the subsidies and other agricultural facilities to the farmers.

This research is carried out with the following specific and general objectives. The main general objective of this study is to identify the main factors affecting to the productivity of the big onion cultivation of the Hambantota area during the off season. Hambantota district is situated in Southern province of Sri Lanka and it is the biggest district in the Southern province. Hambantota features a tropical wet and dry climate. There is no true dry season, but there is significantly less rain from January–March and again from June–August. The heaviest rain falls in October and November. The city sees on average roughly 1,050 millimeters (41 in) of precipitation annually. Average temperatures in Hambantota change little throughout the year, ranging from 26.3°C (79.3°F) in January to 28.1°C (82.6°F) in April and May. Based on the Department of Census and Statistic's; Census of Population and Housing 2012 Hambantota has a population of 596,617 and there are 293,567 males and 303,050 females.

The specific objective of this study is to give the recommendation to improve the big onion cultivation in the Hambantota area and specific objectives are to study the average yield of big onion cultivation in Hambantota area with the other big onion cultivated areas and identify the main issues encountered in big onion cultivation in Hambantota. Also identify the critical contributing factors.

This research studies the off-season onion cultivation in Hambantota district. The scope covers only the Hambantota area of big onion farmers. Therefore, this research has been limited to the onion farmers of the Hambantota area. This research studies the factors affecting to the yield of the big onion cultivation in Hambantota. Therefore, the responses were collected from the local onion farmers from the Hambantota area. Thus, 200 big onion farmers were considered as a sample for this study and all farmers are from Beralihela, Weeravila, Ambalantota, Bandagiriya, Lunama, Yodhakandiya and Meegahajandura West and Meegahajandura East agrarian services divisions (ASDs). All these agrarian service divisions belong to Hambantota district. These data were collected by the Hector Kobbekaduwa Agrarian Research and Training Institute (HARTI) in 2015-16.

This article is organized as follows, In Section 2, we describe the nature of the data set that we used to analyze, Section 3 outlines the Multiple regression Analysis used for analysis with specific reference to step wise model along with its applicability for the current interest of study. Section 4 expands on Statistical analysis conducted with the aid of SPSS statistical software and key results derived. To sum up in Section 5, the Conclusion states the main findings with direct comparison against the only previous study conducted in Sri Lanka and against the big onion productivity of other areas in Sri Lanka and also against literature published worldwide and further discuss the limitations of the finding for the readers reference.

2 Methodology

This study is based on 201 big onion farmers data from eight Agrarian Division Centers (ADC) in Hambantota district which collected from Department of Agriculture Sri Lanka (DOA). The dependent variable is productivity of the big onion per acre and thirteen independent variables that are considered in this study are age of farmers, gender of farmers, agrarian division of farmers, educational level, seasonal month, land type, cultivating quantity (acres), watering days, participate for big onion cultivation programs,

known about water affecting ways, known about suitable land quantity, subsidies seeds quantity from government and subsidies polythene quantity. Table 1 shows the abbreviate description about the variables that are consider in this study.

Table 1. Description of variables

Variable symbol	Variable name	Description of variable
Y	Big Onion Productivity	Big onion yield per acre
X1	Agrarian Service center	Agrarian service center of farmers
X2	Age	Age of farmers
X3	Educational Level	Educational level of farmers
X4	Seasonal Month	Seasonal month of the yield
X5	Land Type	Land type that farmers used for cultivation
X6	Watering Days	Number of dates for watered for cultivation
X7	Income	Farmers' income per month
X8	Society	Participated agricultural programs
X9	Cultivated Quantity	Cultivating quantity (acres) used for cultivation
X10	Subsidies Fertilizers	Subsidies fertilizer quantity
X11	Subsidies Seed	Subsidies seeds quantity from government
X12	Subsidies Polythene	Subsidies polythene quantity from government

Here, among these independent variables there are some categorical variables. There are Age, Educational Level, Agrarian Service Center, Land Type, and Society. Further, variable Age has categorized to age groups and they are 0-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70 and above 70. Also, Educational Level has categorized into groups and numbered those groups. And they were 1 (not attended school and unable to read and write), 2 (Not attended school but able to write and read), 3 (Grade 1 to 5), 4 (Grade 6-11), 5 (Pass O\L examination), 6 (Pass A\L Examination), 7 (Undergraduate), 8 (Graduate or higher), 9 (Diploma) and 10 (Others). Like this Agrarian Service Centers were categorized and numbered them and they were 1 (Beralihela), 2 (Weeravila), 3 (Ambalantota), 4 (Bandagiriya), 5 (Lunama), 6 (Yodhakandiya) and 7 (Meegahajandura West) and 8 (Meegahajandura East). As like this, other categorical variables were categorized and numbered. After that descriptive analysis were done and got a brief idea about these variables.

Every variable was examined separately to identify their behavior. Then the combinations of categories in variables were analyzed to identify the relationship of each other. SPSS (Statistical Package for Social Sciences) software was used for analyzing data. The data were arranged into tabular manner and each variable was represented by using charts to identify the spread of data. The graphs were drawn, and descriptive statistics were obtained for each variable. Multiple linear regression model was used to find out the best factor combinations for numerical response variables for the productivity of the big onion. A model was fitted to predict productivity of the big onion using multiple regression techniques.

First, preliminary analysis was done for all variables. Before use the multiple linear regression model correlation between variables was done using Pearson correlation, Spearman's Rank correlation and Partial correlation.

Initially, multicollinearity have to check before use the regression model. But there is no clear-cut criterion for evaluating multicollinearity of linear regression models. Correlation coefficients of independent variables can be computed. But high correlation coefficients do not necessarily imply multicollinearity. We can make a judgment by checking related statistics, such as tolerance value or variance inflation factor (VIF),

Eigenvalue, and condition number. If VIF ,1/ (1 - R2) is greater than 10 multicollinearities can be considered as statistically significant. When we use Variance Inflation Factor (VIF), it measures the impact of collinearity among the variables in a regression model. Values of VIF that exceed 10 are often regarded as indicating multicollinearity, but in weaker models values above 2.5 may be a cause for concern. The Variance Inflation Factor (VIF) quantifies the severity of multi-collinearity in an ordinary least squares regression analysis and provides an index to measures the variance of an estimated regression coefficient.

The VIF formula,

$$VIF\ (\beta\hat{\imath}) = \frac{1}{1 - R_i^2}$$

where R_i^2 is the coefficient of determination of regression. If $(\beta\hat{\imath})$ >1 then multicollinearity is high and 10 has been proposed as a cut off value.

Regression analysis is a quantitative research method which is used when the study involves modelling and analyzing several variables, where the relationship includes a dependent variable and one or more independent variables. One of the main occasions where such analysis is used is to understand the relationship between independent variables and a dependent variable. Regression model, basically, specifies the relation of dependent variable (Y) to a function combination of independent variables (X) and unknown parameters (β).

i.e., Y ≈ f (X, β). It is of the form, $Y=a+bX$ where a is the slope of the line, b is the y-intercept.

They can be estimated using the equations given below.

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$a = \frac{\sum y - b(\sum x)}{n}$$

A regression model will be fitted to predict factors affecting for the productivity of big onions in Hambantota district during off-season.

$$Y=\beta_0+ \beta_1X_1+ \beta_2X_2+\beta_3X_3+\cdots \beta_kX_k+ \varepsilon$$

where X_k is the last selected variable using stepwise variable selection method. Symbol ε is the error term and $\beta 0$ is the intercept and β's are coefficients for selected variables.

Normality test was used to check sample data set has been drawn from normal distribution population. There are several methods have to test normality, such as Kolmogorov-Smirnov test, chi-square's test, graphical tests, analysis of Skewness and Kurtosis. Data transformation used for the variables that has not behavior as normal.

Linear regression analysis was used to get summary results of the Big onion Productivity data set. In summary analysis report, check the variable are whether significance or not. If variable significant, P-value should be less than 0.05 otherwise the variable not significant. After that selected significant variable use for the prediction of big onion productivity of this data set. The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), is measured, in order to find the prediction accuracy of the proposed model. Accuracy can be explained in 100-MAPD value.

$$MAPE = \frac{100}{n} * \sum_{t=1}^{n} \frac{|A_t - F_t|}{|A_t|}$$

where At- Actual value, Ft- Predict value.

Also use Symmetric Mean Absolute Percent Error (SMAPE) as an alternative method.

$$SMAPE = \frac{200}{N} * \sum_{i=1}^{N} |\frac{x_i - f_i}{x_i - f_i}|$$

where xi- Observation value, fi- Forecast value.

3 Results and Discussion

Using multiple regression model, the fitted multiple linear regression model for log value of yield from one acre is,

ln Y = 8.248 + 0.085*ln X9 - 0.086*ln X10 – 0.012*X4 + 0.0001*X7

where X4 = Seasonal Month, X7 = Income, X9 = Cultivated Quantity, X10 = Subsidies Fertilizers.

Table 2 represented the P – values, coefficients and the VIF values that obtained by the regression model.

Table 2. Results of the regression model

Variable	Coefficient	P – value	VIF
Constant	8.248		
Agrarian Service Centre	0.008	0.292	1.215
Age	-0.001	0.703	1.228
Educational Level	-0.007	0.617	1.208
Seasonal Month	-0.012	0.003	1.143
Land Type	0.011	0.462	1.179
Watering Days	-0.002	0.288	1.227
Income	9.742E-5	0.011	1.145
Society	-0.008	0.660	1.097
Ln (Cultivated Quantity)	0.085	0.031	3.328
Ln (Subsidies Fertilizers)	-0.086	0.014	3.038
Ln (Subsidies Seed)	0.025	0.418	2.640
Ln (Subsidies Polythene)	0.018	0.558	1.739

Only the variables Seasonal Month, Income, log value of cultivated quantity and log value of Subsidies Fertilizers are significant variables at 5% significant level since they have p-values less than 0.05 and all other variables were non-significant because the p-value was larger than 0.05. From the VIF values clearly identify the collinearity of the variables. Also, from the coefficient values some were negative and some were positive and that means those negative variables were affected to depended variable negatively and positive variables were affected to depended variable positively.

For the above model, R-Squared and R-Adjusted values for the fitted regression model are shown in Table 3.

This regression model, intercept is 8.248 and regression coefficient value of “Seasonal Month” variable is -0.012, “Income” variable is 0.0001, “log value of cultivated Quantity” variable is 0.085 and “log value of Subsidies Fertilizers” variable is -0.086. Under multiple linear regression analysis consider log value of the big onion yield per acre.

Table 3. R - squared and R - adjusted values for fitted regression model

Model	R Value	R2 Value	Adjusted R2 value
1	0.894	0.799	0.784

In the Table 4, there represented the all significant variables and their P – values.

Table 4. P - values of regression variables for fitted model

Variable	Description	P - value
X4	Seasonal Month	0.003
X7	Income	0.011
X9	Cultivated Quantity	0.031
X10	Subsidies Fertilizers	0.014

Since the above displayed p-values < 0.05, the null hypothesis is rejected which indicates the variable is not significant. Therefore, each variable is significant in this model.

All the results that obtained from the fitted regression model has shown in the Table 5 and the p-value of this model shows that the fitted model is significant. Further describing the p-value which is less than 5% shows that the null hypothesis is rejected and hence the model is significant. Moreover, this model has an average adjusted R-squared value of 78.4%. Therefore, overall, the proposed model appears to have performed only with 4 variables and one interaction term. From this model 78.4% of log value of yield per acre was described by these 4 variables. Therefore, the model shows 78.4% of log value of yield per acre can be predicted using only 4 variables rather than monitoring many variables. Moreover, VIF = 8.654 which is less than 10 indicates that multicollinearity is not statistically significant.

Table 5. Results obtained from the fitted regression model

Residual standard error	R-squared	Adjusted R-squared	F-statistic	p-value
0.21458	0.799	0.784	53.96 on 2.4 of 12	< 0.0001

The model validation was done in order to check the validity of the fitted regression model. Since the estimated values from the fitted model agreed sufficiently well with the observed values, the model validation is successful. Therefore, our fitted model is a good model to predict log value of yield per acre. The observed values and estimated values are graphically shown in the Fig. 1.

As shown in the Fig. 2, observed value of log value of yield per acre fairly go with its predicted value through the identified regression model.

When considering MAPE value and SMAPE value, this model has the accuracy more than 98 percent. Accuracy values of each model and the error scores of the fitted model are represented in the Table 6.

Table 7 represent the average productivity comparison of Hambantota district with other cultivated areas.

According to the Table 7 (Average productivity of big onion in cultivated districts) Matale has the highest productivity and it is 20.00 MT/ha. Mahaweli H, Anuradhapura and Polonnaruwa has 18.00 MT/ha, 17.00 MT/ha and 14.00 MT/ha respectively. Those districts cultivated big onion during the peak season, but Hambantota district big onion cultivated during the off-season and 9.00 Mt/ha average productivity gave during that season. The average yield from off-season cultivation in Hambantota was as low as 3,387kgs per acre compared to the average yield from the Matale peak season 8,801kgs per acre. These data collected from Department of Census and Statistic in Sri Lanka.

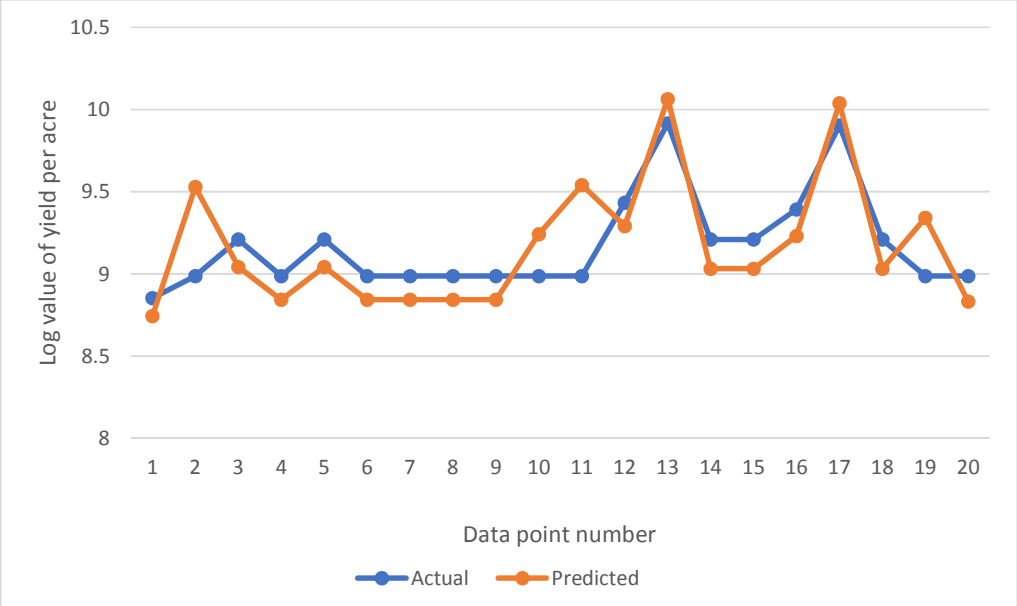

Fig. 1. The plot of fitted values vs. actual values

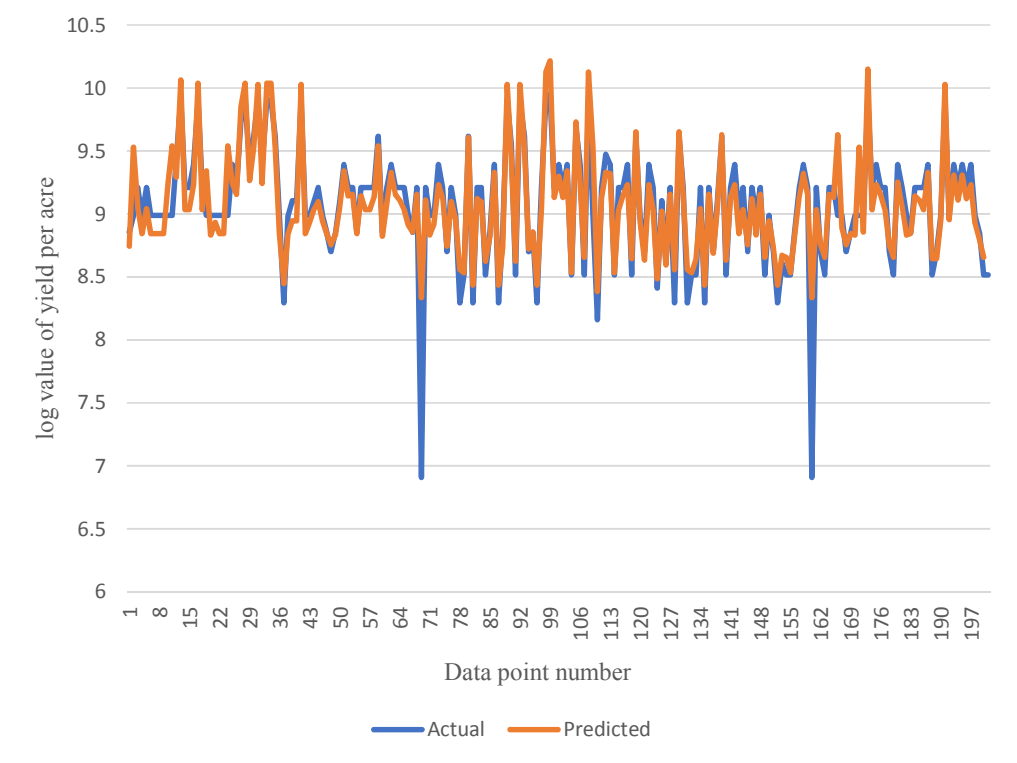

Fig. 2. Plot of Actual values and forecast values of yield per acre log value

Table 6. Accuracy method and error scores

Accuracy method	Error score	Accuracy percentage
MAPE	3.05/201*100=1.52%	(100 – 1.52) % = 98.48%
SMAPE	1.51/201*200=1.502%	(100 – 1.502) % = 98.498%

Table 7. Average productivity of big onion in other districts

District	Average productivity (MT/ha)
Anuradhapura	17.00
Matale	20.00
Mahaweli H	18.00
Polonnaruwa	14.00
Hambantota (off-season)	9.00

4 Conclusion

In this study, a model for predicting the log value of yield per acre was built using only 4 variables namely Seasonal Month, Income, Cultivated Quantity and Subsidies Fertilizers.

The regression model which was found is describing almost all the variation of the dataset since it has a 78.4% of adjusted R-squared value. The model developed through the research would be helpful to achieve the 75% goal of Yield Per Acre, basically focusing on only these for variables also for more than 90% goal we have to analysis more reasons that affecting to the yield such as land preparation for the cultivation, variety of seeds used for cultivation, variety of watering methods etc.

The results of data validation show the model predict well within its tolerance. Therefore, below model can be used to predict the log value of yield per acre that the farmers can get from their cultivation.

$$\ln Y = 8.248 - 0.012X_4 + 0.0001X_7 + 0.085lnX_9 - 0.086lnX_{10}$$

where Y = Yield per acre, X_4 = Seasonal Month, X_7 = Income, X_9 = Cultivated Quantity, X_{10} = Subsidies Fertilizers.

The results from the average productivity comparison, Hambantota district average productivity was less than other areas. Big onion productivity of Matale is more than 2 times greater than big onion productivity of Hambantota. Off season big onion cultivation in Hambantota district is not very effective because of the average productivity is less than other areas in Sri Lanka.

The Government should improve current subsidies quantity of seeds, polythene and fertilizers and dissemination of information and increase its current level of extension services to provide better awareness on proper cultural practices, control of pest and diseases and about suitable water amount to increase both quality and the quantity of output.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Rambukwella R, Hathurusinghe C, Vidanapathirana R, Somaratne T. Production and marketing of other FieldCrops: A review. Hector Kobbekaduwa Agrarian Research and Training Institute, Colombo; 2012.

[2] Pattie PS, Wickramasinghe YM. Present status and future prospects of onion production in Sri Lanka. 1993;45.

[3] P. T. F. O. N. F. Production. "www.agrimin.gov.lk," Divisions - Ministry of Agriculture - Sri Lanka, 2016 - 2018.
Available:http://www.agrimin.gov.lk/web/index.php/en/downloads/circularseng

[4] F. P. a. A. D. Marketing, "www.harti.gov.lk," Hector Kobbekaduwa Agrarian Research and Training Institute, January 2015.
Available:http://www.harti.gov.lk/images/download/market_information/monthly_price_list/january_2015.pdf

[5] C. A. Businesses, Cost of Production of selected Vegitables, Grains and Pulses, CIC Institute of Agri Businesses; 2010.

[6] Matsuyama K. Agricultural productivity, comparative advantage, and economic growth. Journal of Economic Theory. 1992;58(2):317-334.

[7] Thennakoon M, Silva D. Market window analysis: A case of tobacco paddy and big oinin farmers in Galewela, Sri Lanka. Sabaragamuwa University Journal. 2012;11(1):95-108.

[8] Jayathissa R, Wickramasinghe W, Piyasena C. Food consumption patterns in Sri Lanka. Hector Kobbekaduwa Agrarian Research and Training Institute, Colombo; 2014.

[9] Wijesinghe R, Wijesinghe I, Perera A. Socio economic factors affecting the productivity of green gram. Hector Kobbekaduwa Agrarian Research and Training Institute, Colombo; 2015.

[10] Jin S, Huang J, Hu R, Rozelle S. The creation and spread of technology and total factor productivity in China's agriculture. American Journal of Agricultural Economics. 2002;84(4):916-930.

[11] Gunawardena S. An overview of onion production in Sri Lanka, Department of Agriculture, Peradeniya.

[12] Lesly W. Agronomic research on big onion, Field Crops Research and Development Institute, Mahailluppallama.

© 2019 Sulochana and Nawarathna; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/52735