UDC 332; DOI 10.18551/rjoas.2022-06.20

MULTICOLLINEARITY DIAGNOSTICS UPON COBB-DOUGLAS PRODUCTION FUNCTION FOR ESTIMATING RESOURCE USE EFFICIENCY OF TOMATO IN CHITWAN, NEPAL

Saugat Khanal*
Agriculture and Forestry University, Chitwan, Nepal

Sunita Lohani
Nepal Polytechnic Institute, Purbanchal University, Bharatpur, Chitwan, Nepal

Saurav Khanal
School of Business Studies, Tribhuvan University, Nepal

*E-mail: ksaugat506@gmail.com
ORCID: 0000-0002-9146-7428

ABSTRACT
Testing the multicollinearity problem upon the empirical estimation of the Cobb-Douglas production function forms a major part of the research analysis, but many analysts avoid it, which may cause a statistical error on the regression coefficients and correlation. Keeping this point in view, a study was carried out in Chitwan district in 2020 to analyze the profitability and resource use efficiency of tomato production with testing of collinearity problem on regression analysis. A total of 180 tomato growers were selected based on Cochran's formula of sample size determination, for interviews that used a pretested questionnaire. The benefit-cost ratio in the research area was 1.28 which indicates that tomato production is profitable and farmers of Chitwan get an additional 28 paisa with an investment of one rupee in tomato farming. The cost for tillage, potash, seed, and labor is overutilized, while for FYM, DAP, urea, and irrigation it is underutilized. The empirical findings of collinearity diagnostics revealed that there is no multicollinearity problem in the present regression analysis. Thus, the Cobb-Douglas production function was used without the need for any correction. The analysis suggests that the efficient utilization of resources is necessary to increase the profitability of tomato-growing farmers in the study area.

KEY WORDS
Profitability, variance inflation factors, tolerance, B/C ratio, regression.

Farming has been the way of living for mankind since the beginning of human civilization. Most of the developing countries largely depend on agriculture for their economy (Lawal, 2011; Mogues et al., 2012). Nepal is a developing country, with emerging agricultural economies. The population is predominantly rural, and agriculture remains the main occupation, with the majority of farmers following subsistence agriculture (Khanal et al., 2021). Tomato (Solanum lycopersicum L.), popularly known as 'Love apple', is one of the most cultivated vegetables in the tropical region and the second most produced vegetable next to potato, globally (FAOSTAT, 2018). The latest data of the Ministry of Agriculture and Livestock Development (MoALD, 2019), stated tomato as the third largest vegetable in Nepal given production, with a total production of 4.06 lakh tonnes and productivity of 18.01 tonnes ha^{-1}. The crop is highly important from the economic point of view because it largely contributes to the national Gross Domestic Product (GDP) of Nepal. Besides, it has health importance too. It is an excellent source of lycopene, the pigment that provides red color to tomatoes, which is believed to heal cancer patients (Giovannucci, 1999). The products of tomatoes have multifarious uses and are consumed in diverse ways. It is consumed raw in salads. Unripe green fruit of tomato can also be breaded and fried. The juice of tomato is often sold as a drink. The fruit is rich in vitamins (John et al., 2010), essential minerals, and lycopene, which function as an antioxidant (Osemwegi et al., 2010) that helps to alleviate the risk of prostate and breast cancer in women (Giovannucci, 1999).

In the Chitwan district, tomato production is a big part of the farming system. Tomato forms an important component of food consumed in Nepal and this is evident in the fact that many Nepalese food items have tomatoes as a component ingredient. However, tomato is a labor-intensive vegetable, usually wages to labor cover a major share of the total cost of production. The majority of the tomato produced in Chitwan is used for home consumption and only the surplus quantity is used for industrial purposes (Ghimire et al., 2018). As compared to other vegetables, tomato cultivation needs a high level of management, more numbers of labor, and capital inputs. Labor requirements for land preparation, harvesting, grading, packaging, and marketing are very intense (Ogunniyi et al., 2011). The allocation of resources varies with season and weather, which may result in unstable resource use efficiency. The revenue from tomato production is dependent on both increasing the productivity of tomatoes and extending the usage of tomato products as industrial inputs (Sharma, 2016). The limited availability of resources with the tomato producers and inefficient allocation of these resources has led to chronic inefficiency in tomato production (Shrestha et al., 2015). For optimum production and maximum benefit, the resources must be allocated in such a way that no inputs are neither underutilized nor over-utilized. It is thus felt necessary to determine whether the input resources have been efficiently utilized which can foster economic benefit of the tomato growers. Though many pieces of research were performed about resource use efficiency in tomatoes, no article has been written considering the testing of multicollinearity problems in Cobb-Douglas production function yet. Therefore, this study was designed to assess the resource use efficiency by implying the model Cobb-Douglas production along with collinearity testing.

Resources used in the production process are viewed as an input that operates production activity. Resources are supposed to be efficiently utilized when it is allocated in the best possible ways by optimizing the cost of production (Dhakal et al., 2015). The input resources required in tomato production include seeds, labor, fertilizers (FYM- Farm Yard Manure, DAP- Diammonium Phosphate, Potash, etc.), irrigation, tillage, mulching materials, and many others. The collinearity diagnostics and the Cobb-Douglas production model are used in the study to estimate the correlation effect between different input variables that might affect the resource use efficiency. The findings of the empirical studies of the model helped to figure out whether the farmers are making the rational allocation of the available resources or not. Most often, farmers might use inputs rationally, but not at the optimum level. The goal of every entrepreneur is to maximize profit while minimizing cost, so it is necessary to determine the efficiency of resource use, which is estimated in our research from the regression coefficients obtained from the Cobb-Douglas regression. Multicollinearity leaves a serious error in the statistical analysis. The output of any process depends on the inputs used. Several inputs combinely yield a particular output. Very often, in most production processes, these inputs might highly correlate with each other affecting the output. So, it is necessary to test whether there exists a problem of multicollinearity regression analysis or not. Multicollinearity makes it difficult to interpret the coefficients, and it reduces the power of the model to identify statistically significant independent variables. These are potentially serious problems. In this study we had focused on the detection of multicollinearity problems among the explanatory variables, however, no problem was found. Therefore, there was no need to correct the multicollinearity problem in the study.

MATERIALS AND METHODS OF RESEARCH

The research was conducted in 2020 in different villages of Bharatpur metropolitan city of Chitwan district of southern lowland of Nepal. Chitwan district was intentionally selected for the study because it is one of the major tomatoes producing districts with identifiable tomato-growing farmers. A total of 180 tomato growers were purposively selected from different local municipals that lie under Bharatpur metropolitan city. Interviews were conducted with the help of a questionnaire for both quantitative and qualitative data gathering. Primary data was collected through the key informant survey (KIS), questionnaire survey, focus group discussion (FGD), and an online survey with the farmers. The FGD involved gathering people from different backgrounds but of similar experiences together to talk about a specific topic of interest. Data thus obtained were analyzed through computer software packages like the Statistical Package for Social Science (SPSS) and Microsoft Excel.

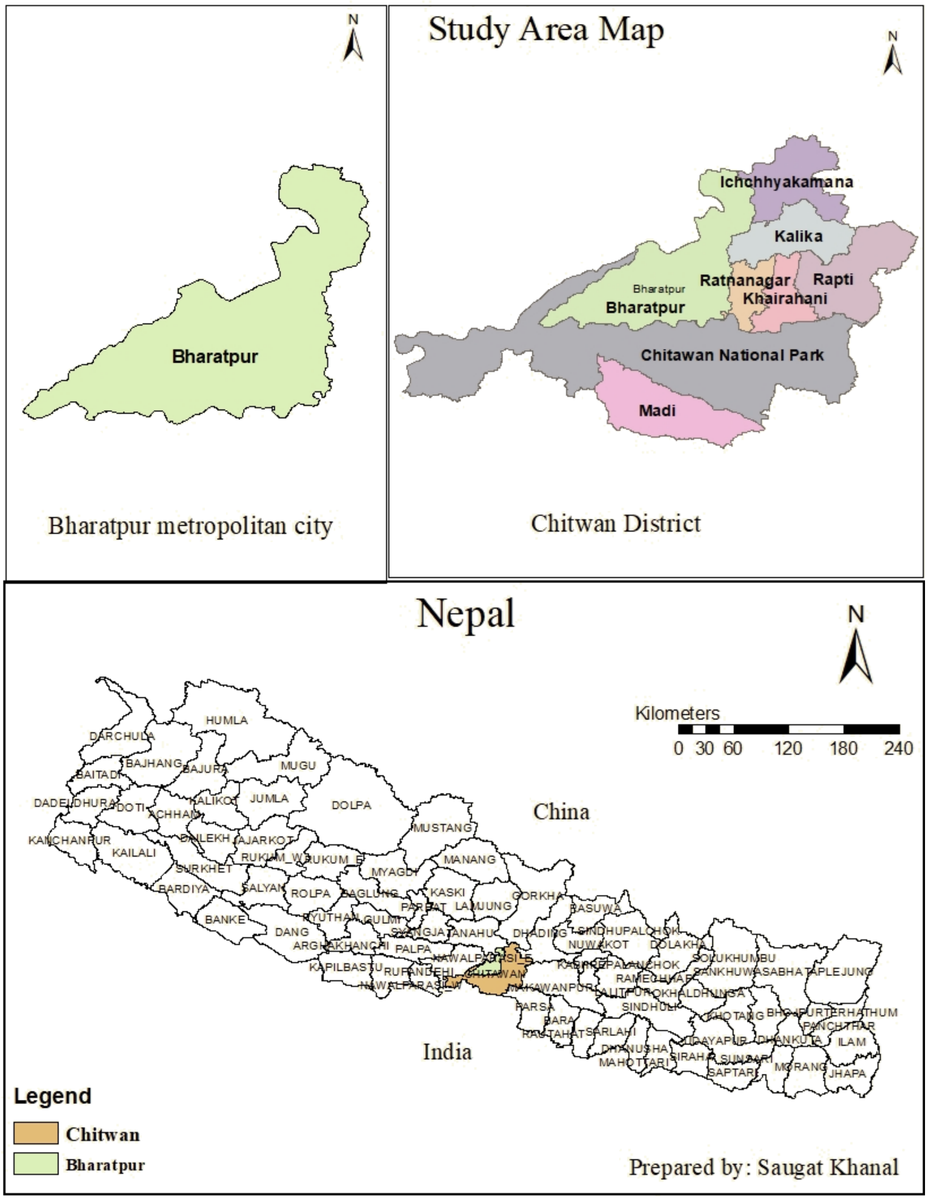

Figure 1 – Map of Nepal showing the study area

For large populations, Cochran (1963) developed Equation 1 to work out a representative sample for proportions.

$$N_0 = Z_{\alpha/2}^{2}\, p(1-p)/ e^2 \qquad (1)$$

Where: $Z_{\alpha/2}$ is the critical value of the Normal distribution at α/2, p is the estimated proportion of an attribute that is present in the population, and q is 1-p, and e is the desired level of precision (i.e., the margin of error). The value for Z is taken from statistical tables that contain the area under the normal curve. In this research, we are doing a study on the tomato producers of a large city, and we want to find the technical efficiency of tomato farming. We don't have adequate information on the subject, to begin with, so we assume p = 0.5, which gives the highest variability. At 95% level of confidence and ±6% precision level, the sample size was calculated. A 5% significance level gives us Z values of 1.96, per the normal tables. Using the equation and the data obtained, the sample size was estimated to be 180.

The socio-economic variables of age, gender, educational level, land size, economically active member, etc. were analyzed using descriptive statistics. Ordinary Least Squares (OLS) were used to estimate the production function. To make our empirical analysis achievable, we need to specify a functional form for the tomato production function in Eq. 2. We, therefore, assume a Cobb–Douglas production function of the form. The C-D production function is well fitted to empirical data in economics because it typically derives the fact that its variable follows a power-law distribution (Ishikawa et al., 2021). The power law, also known as scaling law, states that a relative change in the amount of one variable results in a proportional relative change in another (Corral and González, 2019).

The linear stochastic form of the specified Cobb – Douglas function is given as:

$$Y = f\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, \mu)$$
$$Y = \beta_0 X_1 \beta_1 X_2 \beta_2 X_3 \beta_3 X_4 \beta_4 X_5 \beta_5 X_6 \beta_6 X_7 \beta_7 X_8\ \beta_8 e^{\mu} \quad (2)$$

The above-mentioned equation 2 is linearized in logarithmic function as:

$$lnY = ln\beta_0 + \beta_1 lnX_1 + \beta_2 lnX_2 + \beta_3 lnX_3 + \beta_4 lnX_4 + \beta_5 lnX_5 + \beta_6 lnX_6 + \beta_7 lnX_7 + \beta_8 lnX_8 + \mu \quad (3)$$

Where, Y= Income from tomato production (NRs. /ha), X_1 is Tillage cost (per ha), X_2 is FYM cost per ha, X_3 is Urea cost per ha, X_4 is DAP cost per ha, X_5 is Potash cost per ha, X_6 is Seed cost per ha, X_7 is Irrigation cost per ha, X_8 is Labour cost per ha. β_0 is the intercept parameter (constant), β_1- β_8 are coefficient to be estimated, μ is the random error term, and ln is the natural logarithm. The value of 1 NRs. (Nepalese Rupee) is equal to 0.0085 US dollars.

$$RTS = \Sigma \beta i$$

Where: β_i = regression coefficient of i^{th} variables. The sum of β_i from the Cobb-Douglas production function shows the nature of a return to scale. RTS<1: Decreasing return to scale, RTS=1: Constant return to scale, RTS>1: Increasing return to scale.

The *a priori* expectation of the parameters is given as $\beta_i > 0$. The null hypothesis (H_0) and the alternative hypothesis (H_A) are:

- H_0: β_0, β_1, β_2, β_3, β_4, β_5, β_6, β_7, and $\beta_8 = 0$ (There is no relationship between the value of income from tomato production per hectare and the various inputs used);
- H_A: β_0, β_1, β_2, β_3, β_4, β_5, β_6, β_7, and $\beta_8 > 0$ (There is a significant relationship between the value of income from tomato production per hectare and the various inputs used).

To assure optimum profit and efficiency of resources, a farmer must allocate resources at the level where the marginal value product (MVP) of the resources is equal to their marginal factor cost (MFC) under perfect competition (Kabir Miah et al, 2006). The coefficients from Cobb-Douglas production are taken for the estimation of efficiency coefficients (Naqvi and Ashfaq, 2013). According to Goni et al. (2007), Fasasi (2006), and Stephen et al (2004), the efficiency ratio is calculated as:

$$r = \frac{MVP}{MFC} \quad (3)$$

Where: MVP= Marginal value product, MFC= Marginal Factor Cost, and r = efficiency ratio.

The marginal value product was computed by using formula:

$$MVPi = \left(\beta i \times \frac{Y}{Xi}\right) * Py \quad (4)$$

Where: βi = Estimated regression coefficients; Y and X_i are the values from the geometric mean, and Py is the unit price of output.

$$MFC = Pxi \quad (5)$$

Where: P_{xi} = Unit price of input X_i.

MFC of each input was however obtained from data collected on the unit market prices of the various inputs during the 2007 production season. The decision rule for the efficiency analysis is if:

- r = 1; efficient use of the resources;
- r >1; underutilization of resources and increased utilization will increase output;
- r <1; overutilization of resources and decrease in its usage would lead to maximization of profit.

The relative percentage change in the Marginal Value Product of each resource was calculated by using the following formula:

$$D = \left(1 - \frac{MFC}{MVP}\right) \times 100 \quad (6)$$

$$\text{Or, } D = (1-1/r) \times 100$$

Where: D= Absolute value of percentage change in MVP of each resource.

When a regressor is almost a linear combination of other regressors in the test, the affected estimates are unstable and have more standard errors. This problem is called collinearity. Multicollinearity describes a high degree of linear intercorrelation between independent variables in a multiple regression model, such as in the cobb-Douglas function, and results in incorrect results of regression analyses (Kim, 2019). The collinearity diagnostics indicate that there are serious problems with multicollinearity. If eigenvalues are close to 0, the predictors (dependent variables) are highly intercorrelated and even small changes in the data values may result in large changes in the estimates of the coefficients. The condition indices were estimated as the square roots of the ratios of the highest eigenvalue to each successive eigenvalue. If the value is greater than 15, it confirms a possible problem with collinearity, if greater than 30, indicates a serious problem. Variance Inflation Factor (VIF) was calculated as:

$$VIF_k = 1/(1-r_k^2) \quad (7)$$

Where: r_k^2 is the goodness of fit of the linear model for x_k based on all other variables. Thus, we estimate VIF to measure the severity of collinearity in regression analysis.

$$\text{Tolerance} = 1/VIF_k = (1-r_k^2) \quad (8)$$

Where: r_k^2 indicates the unadjusted coefficient of determination for regressing the i^{th} regressors on the remaining ones.

RESULTS AND DISCUSSION

Socio-Economic Characteristics of Tomato Farmers. The majority (82.22%) of the farmers was male (Table 1), which suggests that women's household responsibilities have limited their engagement in agriculture. The study areas are majorly patriarchal and the unrecognition of women's rights to land and property is viewed as the reason for women being less empowered for farm works (Khanal et al., 2021). Diversity in ethnicity was observed as a unique feature, dominated by Brahmin (86%), followed by Chhetri, Janajati, Dalit, and others. More than half (56.67%) of the individuals interviewed had received at least some secondary education; they were more educated than the average Nepali. 14.44% of the interviewed farmers were university graduates, which indicates the involvement of educated youth in farming. The nuclear family is dominant (71.67%).

Table 1 – Socio-demographic characteristics of the study area

Characteristics	Frequency (N=180)	Percentage	Mode category
Gender			
Male	148	82.22	Male
Female	32	17.78	
Ethnicity			
Brahmin	86	47.78	
Chhetri	14	7.78	
Janajati	31	17.22	Brahmin
Dalit	16	8.89	
Others	33	18.33	
Educational level			
Illiterate	12	6.67	
Primary level	40	22.22	Secondary level
Secondary level	102	56.67	
University level	26	14.44	
Family type			
Nuclear	129	71.67	Nuclear
Joint	51	28.33	

The average age of farmers was 44.317 (Table 2), which implies higher productivity and technical efficiency among tomato farmers in the area because most of the farmers are still in their active and productive age. Usman and Bakar (2013) in their research had found that tomato production by farmers of the age range of 40-60 years yield higher technical efficiency since they can actively work in the field and obviate the need for additional labor hiring. The mean farm size was 2.67 hectares. A very few parts of the total were under tomato cultivation. This signifies that tomato production in the study areas is still at the subsistence level. The less area of farmland may also be attributed to the perishable nature of tomatoes. The average number of economically active members per household was 2.533 and the agriculture engaged population was 2.383 per household of the interviewed farmers. Economically active population refers to the age group of 16-59.

Table 2 – Demographic characteristics of the sampled household

Variables	Mean (N = 180)	Std. Deviation	Kurtosis	Skewness
Age of the respondents	44.317 ± 0.661	8.867	-0.686 ± 0.36	0.068 ± 0.181
Income from tomato cultivation (NRs. /ha/ annum)	1189211.66 ± 141471.01	1898032.8	11.875 ± 0.36	3.34 ± 0.181
Total Land Size (ha)	2.668 ± 0.27	3.622	7.894 ± 0.36	2.765 ± 0.181
Economically active members	2.533 ± 0.086	1.15	0.809 ± 0.36	1.121 ± 0.181
Agriculture engaged population	2.383 ± 0.089	1.188	0.768 ± 0.36	1.1 ± 0.181

Note: Standard error of respective statistic in the parenthesis.

Economic Analysis of Tomato Production. The cost incurred during tomato production in the study areas largely varied due to variations in the amounts of inputs used by the growers. The total fixed cost incurred during the production period was NRs. 21473.67 per hectare, which includes land rent, equipment cost, and the cost of nursery tray. The variable cost includes the cost of seeds, labor, fertilizers, and pesticides, management costs, and others. As inferred from (table 3), the majority (26.76%) of the cost incurred was calculated from labor, which was estimated in terms of man-days and transformed into a monetary value of NRs. 31833.72 / ha. The result revealed that the total cost of production per hectare was NRs. 118939.645. Seed costs incurred in tomato production accounted for NRs. 3091.42 ha^{-1}, fertilizer cost was NRs. 8859.46 ha^{-1}, and pesticides cost were NRs. 13494.67 ha^{-1}. Tillage, irrigation, and mulching cost tallied for NRs. 12614.49 ha^{-1}. The costs for staking, packaging, and plastic rope accounted for NRs. 25023.37 ha^{-1}.

The average total return was calculated NRs. 118939.65 ha^{-1} (NRs. = 0.0085 USD); average price of tomato per kg was NRs. 53. The Benefit-cost ratio was calculated simply as the ratio of total benefit which accounts for gross revenue to the total cost. On that account, tomato production was considered to be a profitable enterprise in the study area; with an investment of Rs. 1 (0.0085 US$) farmers got an additional 28 paisa.

Table 3 – Cost of production for tomato farming in the study area

Type	Items	NRs. / Ha	% of the total cost
Fixed Cost	Land	9912.72	8.33
	Equipment	4300.88	3.62
	Nursery tray	7260.05	6.1
Variable cost	Seed cost	3091.42	2.6
	Labor cost	31833.72	26.76
	Cost of Fertilizers	8859.46	7.45
	Biopesticides cost	13494.67	11.35
	Cost of Tillage, Irrigation, Mulching	12614.49	10.61
	Cost of staking, packing, plastic rope,	25023.37	21.04
	Others	2548.81	2.14
Total	Total fixed cost	21473.65	18.05
	Total variable cost	97465.94	81.95
	Total cost	118939.59	100

Note: 1 NRs. = 0.0085 US $.

Table 4 – Profitability of tomato farming in Chitwan district

Items	Mean
Total cost of production (NRs. / ha)	118939.59
Average price of tomato (NRs. /kg)	53
Gross revenue (NRs. / ha)	543706.39
Net profit (NRs. / ha)	424766.39
B/C ratio	*1.28*

Empirical findings of regression. From the regression results (Table 5), The R^2 value was 0.838, which indicates that 83.8% of the variation in income of tomato was explained by the independent variables included in the model. Cost of seed and Irrigation is statistically significant at 1% level of significance. The cost of tillage, FYM, potash, and labor is significant at 5% level of significance. The other variables were significant at 10% level of significance. Potash and labor cost was negatively significant to the total outcome. 10% increase in labor cost and potash cost resulted in 3.45% and 1.9% decrease in income respectively. To contrary, the other variables were positively significant to the outcome from tomato production. 10% increase in the cost of tillage, urea, DAP, seed, and irrigation resulted in a respective increase in income by 1.3%, 2.2%, 1.2%, 1.85%, 3.19%, and 4.9%. The sum of regression coefficients of different variables was 0.945 which is less than 1; indicating a decreasing return to scale. It signifies that a 100% increase in all the factors of production included in this model would result in 94.5% increase in income from tomato production.

Moreover, the adjusted R^2 value was 0.817, which implies that 81.7% of the variation in the income from tomato production is explained by the explanatory variables described in the model while accounting degree of freedom. The F-stat value was 38.935 which is highly significant at 1% level of significance, which indicates that the input variables fitted in the regression analysis were significant for explaining the variation in income from tomato production in the study area.

Table 6 is a correlation matrix that shows correlation coefficients between sets of variables. This coefficient is the covariance of the two sets of variables divided by the product of their standard deviations. The correlation between a variable and itself is always 1. All the variables in the study model were significantly correlated with each other, except the correlation between urea and irrigation. The estimated correlation coefficient between urea cost ($X_{3)}$ and seed cost (X_6) was negative (-0.023), which describes the extent to which two variables move in opposite directions.

Table 5 – Linear regression for the estimation of resource use efficiency in tomato production

Variables	Coefficient	Std. error	t-stat	prob.	Collinearity Statistics	
					Tolerance	VIF
Constant	6.468735744	0.709	9.128	0.000	-	-
ln (Tillage)	0.130407002	0.058	2.230**	0.029	0.405329388	2.46712928
ln (FYM)	0.2252565	0.092	2.442**	0.018	0.25376271	3.940689316
ln (Urea)	0.128447952	0.066	1.936*	0.057	0.336888348	2.968342497
ln (DAP)	0.185602163	0.098	1.898*	0.062	0.161409639	6.195416885
ln (Potash)	-0.191517987	0.080	-2.386**	0.020	0.169797607	5.889364519
ln (Seed)	0.319745041	0.062	5.143***	0.000	0.562546021	1.77763234
ln (Irrigation)	0.492644711	0.101	4.853***	0.000	0.179449794	5.572589286
ln (Labour)	-0.34599704	0.145	-2.39**	0.020	0.12429073	8.045652303
R^2	0.838484741					
Adjusted R^2	0.816949373					
F-value	38.935					
RTS	0.945					

*Note: '***', '**' and '*' indicates statistically significant at 1%, 5%, and 10% level of significance, respectively.*

Table 6 – Pearson correlation matrix of the dependent and independent variables in the production function for tomato

n/n	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	lnY	lnX_1	lnX_2	lnX_3	lnX_4	lnX_5	lnX_6	lnX_7	lnX_8
Y	1																	
X_1	.643**	1																
X_2	.413**	.317**	1															
X_3	.426**	.432**	.299**	1														
X_4	.447**	.465**	.481**	.638**	1													
X_5	.467**	.385**	.356**	.772**	.791**	1												
X_6	.317**	.379**	.358**	-0.023	.369**	0.058	1											
X_7	.622**	.542**	.519**	.662**	.417**	.583**	0.086	1										
X_8	.692**	.775**	.502**	.576**	.483**	.519**	.347**	.780**	1									
lnY	.895**	.555**	.407**	.390**	.478**	.488**	.338**	.568**	.606**	1								
lnX_1	.605**	.786**	.403**	.265**	.452**	.360**	.361**	.464**	.607**	.634**	1							
lnX_2	.536**	.442**	.823**	.297**	.526**	.434**	.362**	.515**	.543**	.622**	.599**	1						
lnX_3	.687**	.435**	.542**	.892**	.714**	.763**	.313**	.664**	.564**	.689**	.394**	.546**	1					
lnX_4	.488**	.423**	.515**	.484**	.873**	.719**	.305**	.448**	.467**	.565**	.450**	.592**	.778**	1				
lnX_5	.524**	.383**	.383**	.531**	.669**	.856**	.162*	.551**	.485**	.623**	.440**	.563**	.727**	.804**	1			
lnX_6	.504**	.499**	.399**	0.088	.380**	.230**	.744**	.313**	.478**	.583**	.557**	.445**	.413**	.418**	.381**	1		
lnX_7	.588**	.498**	.487**	.482**	.464**	.539**	.244**	.823**	.660**	.664**	.566**	.657**	.705**	.547**	.629**	.344**	1	
lnX_8	.635**	.594**	.477**	.438**	.468**	.484**	.356**	.639**	.816**	.702**	.691**	.643**	.659**	.512**	.555**	.551**	.724**	1

*Note: **. Correlation is significant at 1% level (2-tailed); *. Correlation is significant at 5% level (2-tailed).*

Multicollinearity diagnostics test in the C-D production function. The estimation of Cobb -Douglas production functions (CDPF) encounter many statistical problems. The multicollinearity problem seen in the least-squares estimation of the CDPF is not new. It is an error that emerged with the model itself (Enaami et al., 2011). The last two columns in table 5 describe collinearity statistics of Tolerance and Variance Inflation Factors (VIF). VIF above 10 were considered as the indicators for problems with multicollinearity (Adnan et al., 2006). The column of Tolerance value is not important because it doesn't give us any new information. Eigenvalues close to 0 imply the problem of multicollinearity, in which explanatory variables are largely intercorrelated and even small changes in the data resulted in large changes in regression coefficient estimates. The highest condition index is termed the condition number. A condition number between 15 and 30 signifies the presence of a collinearity problem and when a value is greater than 30, the multicollinearity is regarded as strong (Kim, 2019).

In this section, multicollinearity is tested from variance inflation factors, condition indices, eigenvalues, and the variance proportions, using calculated data (table 5 and table 6). The results of regression analysis showed all the 8 predictors (X_1-X_8) with VIF less than 10. It means the regression doesn't have a serious problem with multicollinearity. Table 7 showed 8 predictors and 9 dimensions. As no variables were observed to have a VIF value of more than 10, all the dimensions contributed significantly to the predictor data set. The eigenvalues can also be used to measure the presence of multicollinearity. If one or more of the eigenvalues are small (close to zero) and the corresponding condition number is large, then it indicates multicollinearity.

In table 7, the last part of the matrix is the variance proportions. Each column in the variance proportions summed up to 1 or 100%. For those dimensions, with a high Condition Index (greater than 15), we checked for variance proportions greater than 0.9. The row containing at least two variance proportions above 0.9 was assumed to have the predictor's collinearity problem. In our regression analysis, no rows contain at least two variance proportions about 0.9, so there doesn't exist a collinearity problem in this regression analysis of the Cobb-Douglas production function. The fact of not having a collinearity problem in this study was also confirmed from the assessment of the VIF value because all explanatory variables had a value less than 10.

Table 7 – Multicollinearity testing of the predictors

Collinearity diagnostics											
Dimension	Eigen Value	Condition Index	Variance proportion								
			Intercept	lnX_1	lnX_2	lnX_3	lnX_4	lnX_5	lnX_6	lnX_7	lnX_8
1	8.928	1.000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.033	16.350	0.02	0.02	0.00	0.05	0.01	0.07	0.02	0.00	0.00
3	0.012	27.379	0.02	0.24	0.02	0.00	0.00	0.00	0.23	0.00	0.00
4	0.011	29.159	0.02	0.01	0.01	0.59	0.01	0.15	0.01	0.00	0.00
5	0.007	36.270	0.21	0.03	0.01	0.13	0.02	0.03	0.22	0.09	0.00
6	0.004	44.712	0.00	0.45	0.09	0.01	0.09	0.01	0.11	0.23	0.00
7	0.003	56.620	0.01	0.00	0.16	0.21	0.46	0.28	0.13	0.25	0.00
8	0.002	68.218	0.36	0.14	0.47	0.01	0.40	0.43	0.13	0.08	0.00
9	0.001	124.736	0.37	0.10	0.24	0.00	0.01	0.03	0.16	0.34	0.99

Table 8 – Estimation of allocative efficiency of tomato production

Expenses (NRs. /ha)	GM	Coefficient	MVP	MFC	r	D	Efficiency
Tillage	5303.684227	0.130407002	7.619488027	130	0.058611446	-1606.15138	Overutilized
FYM	6858.625783	0.2252565	10.17753872	5	2.035507744	50.87220851	Underutilized
Urea	398.7207906	0.128447952	99.82994723	11	9.075449749	88.98126233	Underutilized
DAP	1320.985011	0.185602163	43.53992558	24	1.814163566	44.87817864	Underutilized
Potash	503.4837347	-0.191517987	-117.876345	30	-3.92921151	125.4503988	Overutilized
Seed	3804.909113	0.319745041	26.04126455	104	0.250396774	-299.366167	Overutilized
Irrigation	1380.272046	0.492644711	110.6042133	32	3.456381666	71.06800994	Underutilized
Labor cost	44244.16757	-0.34599704	-2.42336535	109	-0.02223271	4597.877299	Overutilized

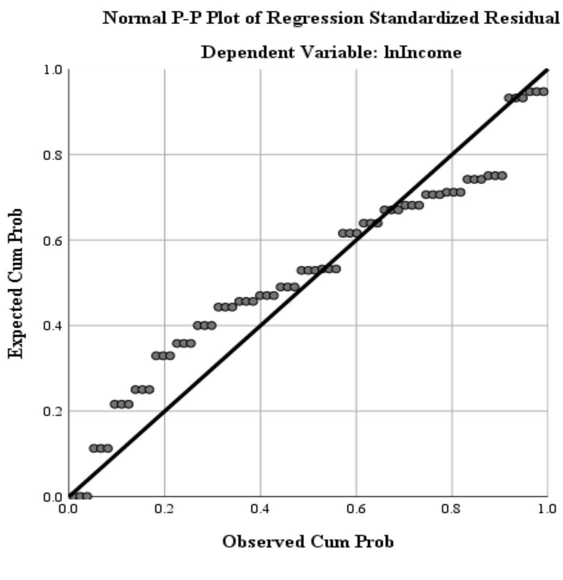

Figure 2 – Normal plot of Regression Standardized Residual

Resource Use Efficiency. The respective regression coefficients of input variables estimated from the Cobb-Douglas production function were used for the calculation of resource use efficiency. From table 8, it is revealed that the allocation of expenditure for tillage, potash, seed, and labor was over-utilized. However, the allocation of expenditure for FYM, DAP, urea, and irrigation was under-utilized, since their efficiency coefficient is greater than one. The data showed that resources allocated such as labor and pesticides used are underutilized in tomato production in Ghana (Tambo & Gbemu, 2010) and Nigeria (Saleh et al., 2016). The cost for tillage operation and labor hiring was found to have a negative efficiency coefficient. This indicates overuse of the expenditure for labor and tillage which, in turn, resulted in the reduction in gross revenue obtained. The adjustment in the MVPs for optimal resource use is shown in Table 8, which indicates that for optimal allocation of resources, expenditure on FYM, urea, DAP, and irrigation needed to be increased by 50.87%, 88.98%, 44.87%, and 71.06% respectively.

CONCLUSION AND POLICY IMPLICATIONS

This study is more empirical and uses a collinearity diagnostics test to check whether there is a multicollinearity problem in the explanatory variables of the Cobb-Douglas production function or not. The results revealed that all the predictors significantly contributed to the income from tomato production. However, expenditure on labor and potash showed a negative significant relationship with output. The output from agriculture is found to be less efficient than economic inputs i.e., for every 100% increase in the cost of inputs, there is an increment of 94.5% in revenue from tomato production. There is disequilibrium in resource use efficiency in tomato production in the study areas. The cost for tillage, potash, seed, and labor is overutilized. While the cost for FYM, DAP, urea, and irrigation is underutilized. The study suggests that the profitability of farmers can be increased through efficient allocation and utilization of resources. From collinearity diagnostics, it is confirmed that there is no multicollinearity problem in the performed regression analysis and the Cobb-Douglas production function is therefore applied without any correction. The conclusion is that introducing the Cobb-Douglas function into resource use efficiency calculation provides hints to the reader and farmers as well to further focus on the optimal allocations of different inputs during the production process. Considering these factors, producers can determine their priority to invest more whether on labor, fertilizers, intercultural operations, seeds, or equipment to achieve the maximum profit at minimum cost.

CONFLICT OF INTERESTS

The authors declare no conflict of interests.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Adnan, N., Ahmad, M.H., Adnan, R. (2006). A comparative study on some methods for handling multicollinearity problems. Mathematika, 22: 109-119. https://doi.org/10.11113/matematika.v22.n.179.
2. Cochran, W. G. (1963). Sampling Techniques, 2nd Ed., New York: John Wiley and Sons, Inc.
3. Corral, Álvaro and González, Álvaro. (2019). Power Law Size Distributions in Geoscience Revisited. Earth and Space Science, 6(5): 673-697. https://doi.org/10.1029/2018EA000479.
4. Dhakal SC, Regmi PP, Thapa RB, Sah SK. and Khatri-Chhetri DB. (2015). Productivity and profitability of maize-pumpkin mix cropping in Chitwan, Nepal. Journal of Maize Research and Development, 1(1): 112-122. DOI:10.3126/jmrd.v1i1.14249.
5. Enaami, M., Ghani, SA., Mohamed, Z. (2011). Multicollinearity Problem in Cobb-Douglas Production Function. Journal of Applied Sciences, 11: 3015-3021. DOI: 10.3923/jas.2011.3015.3021.

6. FAOSTAT. (2018). Food and Agricultural Organization (FAO). Food and Agricultural Organization. Retrieved fromhttp://www.fao.org.
7. Fasasi AR (2006). Resource Use Efficiency in Yam Production in Ondo State, Nigeria. Agricultural Journal, 1 (2): 36-40. URL: URL: https://medwelljournals.com
8. Ghimire, N., Kandel, M., Aryal, M., & Bhattarai, D. (2018). Assessment of tomato consumption and demand in Nepal. Journal of Agriculture and Environment, 18, 83-94. DOI: https://doi.org/10.3126/aej.v18i0.19893.
9. Giovannucci, E. (1999). Tomatoes, Tomato-Based Products, Lycopene, and Cancer: Review of the Epidemiologic Literature. Journal of the National Cancer Institute, 9(4): 317-331. https://doi.org/10.1093/jnci/91.4.317.
10. Goni M, Mohammed S, Baba BA (2007). Analysis of Resource-Use Efficiency in Rice Production in the Lake Chad Area of Borno State, Nigeria. Journal of Sustainable Development in Agriculture & Environment, 3:31-37.
11. Ishikawa, A., Fujimoto, S. & Mizuno, T. (2021). Why does production function take the Cobb-Douglas form? Evolut Inst Econ Rev., 18, 79-102. https://doi.org/10.1007/s40844-020-00180-3.
12. John D, Suthin RT, Usha RS, Udhayakumar R. (2010). Role of defense enzymes activity in tomato as induced by Trichoderma virens against Fusarium wilt caused by Fusarium oxysporum F. Sp Lycopersici. Journal of Biopesticide, 3:158-162.
13. Kabir Miah MA, Ashraful Alam AKM, Rahman AHMA (2006). Impact of Agricultural Credit on MV Boro Rice Cultivation in Bangladesh. Journal of Agriculture & Rural Development, 4(1&2):161-168. https://doi.org/10.3329/jard.v4i1.784.
14. Khanal, S., Dhital, P., & Christian, S. (2021). Farming the future: Youth enthusiasm and transforming Nepal's economy through agriculture. Journal of Agriculture, Food Systems, and Community Development, 10(2), 359-372.
15. Kim J. H. (2019). Multicollinearity and misleading statistical results. Korean Journal of anesthesiology, 72(6), 558-569. https://doi.org/10.4097/kja.19087.
16. Lawal, W. A. (2011). An analysis of government spending on agriculture sector and its contribution to GDP in Nigeria. International Journal of Business and Social Science, 2(20), 244-250.
17. MoALD. (2019). Statistical Information on Nepalese Agriculture, 2017/18. Agribusiness Promotion and Market Development Directorate, DOA, Kathmandu, Nepal.
18. Mogues, T., Yu, B., Fan, S., & McBride, L. (2012). The impacts of public investment in and for agriculture: Synthesis of the existing evidence (ESA Working Paper No. 12-07). Rome: Food and Agriculture Organization of the UN, Agricultural Development Economics Division. Retrieved from http://www.fao.org/3/a-ap108e.pdf.
19. Naqvi, S. & Ashfaq, M. (2013). Technical efficiency analysis of hybrid maize production using translog model case study in district Chiniot, Punjab (Pakistan). Agricultural Sciences, 4, 536-540. Retrieved from http://dx.doi.org/10.4236/as.2013.410072.
20. Ogunniyi LT, Oladejo JA. (2011). Technical efficiency of tomato production in Oyo State, Nigeria. Agricultural Science Research Journal, 1(4):84-91.
21. Osemwegi OO, Oghenekaro AO, Owolo LO. (2010). Effect of pulverized Ganoderma spp., on Sclerotium rolfsii Sacc and Post-harvest tomato fruit preservation. Journal of Applied Science and Research, 6: 1794-1800.
22. Sharma Wagle, T. (2017). Spatial Analysis of Cob-Douglas Production Function in Agriculture Sector of Nepal: An Empirical Analysis. Journal of Advanced Academic Research, 3(2), 101-114. https://doi.org/10.3126/jaar.v3i2.16759.
23. Shrestha, R., Huang, W., & Pradhan, U. (2015). Resource use efficiency in vegetable production in the high hills of eastern Nepal. Nepalese J. of Agr. Economics, 2-3, 79-87.
24. Stephen J, Mshelia S I, Kwaga BT (2004). Resource-use Efficiency in Cowpea Production in the North-Eastern zone of Adamawa state, Nigeria; Department of Forestry and Wildlife Management, Federal University of Technology. Yola, Nigeria.
25. Usman, J., and Bakar, UM. (2013). Profitability of small-scale dry season tomato (Lycopersicon esculentum Mill.) production in Adamawa States, Nigeria. Journals of Science and Technology, 3(6):25-39.