FUTURE FORECASTING FOR CROPS PRODUCTIVITY (WHEAT, BARLEY AND RICE) IN SULAIMANI FROM (2009-2017)

Ezat Sabir Esmaeel

University of Pardubice, Faculty of Economics and Administration

Abstract: *With increasing the planted area and availability of enough amounts of rain, the crops production(Wheat, Barley and Rice) might raise, as there is proportional relation between the size of production and the factors that affect the production (the planted area and the amount of rain). The significant point in this research is the forecast about the production for the period of 2009-20017 to find out the amount of productions in the upcoming years. Our research concludes that, regarding to the statistical testing factors of production the area and rain have positive and significant relationships with the production of crops (Wheat, Barley and Rice).*

Keywords: *Forecasting for Crops Produktivity, Sulaimani*

1. Introduction

The corps product especially Wheat, Barely and Rice are one of the most significant products in any community. It is considered as a strategic commodity, that plant widely in a various different ways all over the Glob. It is also vital and directly related to what is known as food security in an economy, especially Third World economies that are characterized by volatility in food security for their population. These agricultural products are been used in different proportion for days food combination by individuals in different societies. It also plays a great deal in international trade. According to data of some international organizations, the scarcity of food at first instance is due to wheat shortages (Abdulkarim, 1985).

Wheat, barely and Rice are among necessary commodities in humans lives. Food productions, in general the foresaid products are on low elasticity. In other words, when the price of these products increase, it will not lead to a reduction in the quantity demanded, because these are necessary products and cannot be avoided. The problem this study aims at is; despite an increase in area seeded, and an increase in the level of rain, but these increases are not followed by an increase in productivity of these products. Further, it did show some reductions in productivity in some years. The importance of this study comes from the significance of the commodities themselves. As they are basic commodities to individual's life and cannot be avoided. This study assumes that with an increase in durable lands and the availability of its requirements of different factors like Rain will lead to a rise in the productivity of these products, as there is a hyperbolic relation between the size of production and the foresaid factors (Domenic, 1982).

The Econometric model used

The economic theory observed that some interrelations explain any change in the production is because of the earlier changes occurred in some independent variables (inputs), we regarded these variables (Rain & Land) as independent, and also they affect the production outcomes (Abdulhussain, 1992). The economic theory specifies that an increase in one or both independent variables will lead to an increase in the production. This means there is a positive relationship between independent and dependent variables. (**Y**): Represents quantity of production of agricultural products; (Wheat, Barely and Rice), which can be produced from two inputs {R (Rain), A (Area)} in a mathematical model as follow:

Y = F(R, A)

y = a + bR + cA

We can convert the function for econometric model, by entering a random variable to the function above as below:

y = a + bR + cA + U

After an introduction of the variables used in the sample, suitable data collected and created different combination to the observations of the inputs, and its relation to the outcomes. Computer programs used to set a regression, and to implements what is known as Ordinary Least Square (OLS). In doing so, the value of the coefficients (a, b and c) are estimated. A linear function has been used, and the variables are as follow (Milton and Arnold, 1995):

Y = Represents quantity of production (tones)

R = Quantity of Rains

A = the area seeded (Acre=2500 meter square)

The production function has been estimated by Multiple Regression model, using Ordinary Least Square (OLS), in a way which includes all estimations and necessary tests.

Statistical and econometric Tests for the estimated Functions

After specification and estimation stage in building econometric model, comes the testing stage for the coefficients. Therefore; there would be an examination to evaluate the accuracy of the variable's coefficient, using statistical and econometric methods. This is necessary to ensure that the values obtained through statistical and econometric methods, represents the real value in their community or not. There are two assumptions represent this evaluation, (Talb, 1991).

The principal used to determine the deviation value of coefficients from its original value is ordinary least square (OLS), which uses partial derivation to differentiate between estimated values, also equalizing the results to zero. In doing so, the least square of summed deviation for estimated and real value can be obtained. The variation can be obtained as below, Wooldridge(2003) :

$$Var\ (\hat{b}) = S^2 (X^{/} X)^{-1}$$

From above we obtain Standard Error of Estimation of the equation, via dividing the square of summed deviation by numbers of degree of freedom as follow:

$$S^2 = \frac{\sum_{i=1}^{n} e_i^{\,2}}{n-k}$$

Where:

n : Represents the size of sample.

k : Represents the number of the variables in the model

The partial derivative for standard error of each coefficient will be taken, as below:

$$S\hat{b}_R = \sqrt{\frac{S^2}{\sum_{i=1}^{n} R^2}} \quad , \quad S\hat{b}_A = \sqrt{\frac{S^2}{\sum_{i=1}^{n} A_i}}$$

From this other statistical testing can be done.

T – Test

The production function coefficient that has been estimated by using econometric functions means the elasticity of production in relation to the variables used which are level of rain, area seeded. By using T-test the statistical credibility of each coefficient can be informed singularly, In other words; knowing the statistical significance of each independent variable on dependent variable. By testing two important hypotheses (Dominic, 1982):

A: Null – Hypotheses: Ho: b = 0

This assumes no relationship between dependent and independent variables.

B: Alternative – Hypotheses: H1: b=/= 0

The t value can be obtained as follow:

$$t = \frac{\hat{b}}{S\hat{b}}$$

Through the number of degree of freedom, we derive schedule (t), and we compared with accounted (t). If the value of accounted (t) is bigger than scheduled (t), we deny null – hypotheses and accept the alternative-hypotheses. If the value of accounted (t) is smaller than scheduled (t), then we accept null-hypotheses and refuse the model. In other words, as the volume of standard error decreases, the accounted (t) value should increase, Studemanmund (2006).

Coefficient of determination – testing R^2

This test is used to distinguish the important explanatory variables from those of little significance, such as variables with sudden effect on the dependent variable. The coefficient of determination value is lying between zero and one ($0 \leq R^2 \leq 1$).

If $R^2 = 1$, this means that the independent variables explain and illustrate all changes happened in dependent variables but this is a very rare case. And if the value of $R^2 = 0$ this indicates that the independent variable does not explain and has no effect on the changes in the dependent variable, this is rare too. In general, the highest the value of (R^2) or the closer to one (1), the stronger the explanatory power of the estimated function is, and vise versa. The deviation between the real value of the samples and its maiden is called total deviation, and by summing them we can derive the sum square total of the deviation, (Abdulkarim, 1985).

$$SST = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \qquad \text{[SST] (Sum Square Total)}$$

The variation equation will show the variation between the real value of the samples and estimated value, called sum square of the unexplained variation.

$$SSU = \sum_{i=1}^{n} (Y_i - \hat{Y})^2$$

But the variation between estimated value and its maiden (after been summed and powered by two), called the sum of explained variation.

$$SSE = \sum_{i=1}^{2} (\overline{Y}_i - \hat{Y})^2$$

$$SST = SSE + SSU$$

We conclude that:

SST = SSE + SSU

By dividing both sides by SST:

$$1 = R^2 + \frac{SSU}{SST} \rightarrow \qquad R^2 = 1 - \frac{SSU}{SST}$$

Taking degree of freedom into account, the number of degree of freedom decline as we add more independent variables into the model, then we get the adjusted coefficient of determination.

$$\overline{R}^2 = R^2 = \frac{n-1}{n-k}(1 - R^2)$$

This demonstrates what the added variables supplements of changes will be larger than decline of the degree of freedom. In a way, these extra variables will be significance and not excessive.

F – Test

This test will compare between the explanatory variation and non-explanatory variation James and Mark (2006).

$$D.W = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^{\ 2}}$$

This test is used to know the significance of estimated function, also it can be used to test two hypotheses; null-hypotheses, which illustrates the real value of coefficients which are equivalent and equal to zero. In other words, these independent variables have no significant effect on dependent variable. Thus the F – test is used to examine coefficient of determination (R^2), in null-hypotheses ($R^2 = 0$). But the alternative hypotheses refers that the real value of the coefficients are not equal to zero, or the independent variables together have a significance effect on dependent variables. This means R^2 =/= 0. The scheduled F value can be obtained throughout special tables depending on degree of freedom (k – 1), (n – k), then we compare between the accounted (F) and scheduled (F), here; if the value of accounted F is larger than scheduled F, then we accept alternative hypotheses and refuse null-hypotheses, and vise versa.

These Testing come first to explain and illustrate the range of dependency for model's estimated coefficients statistically. And the econometric theory will illustrate for us other testing of second degree to distinguish the majority hypotheses of econometric model, is it accomplished or not? Then we use it to reveal the probability of existence of economic measures problem, from the probability of not existence, in the study which is:

The (D.W) Test: Durbin Watson – test

This test is used to inform the existence of autocorrelation problem or not existence, among random variables on primary degree. Again by this test, the two hypotheses will be examined. The null-hypotheses which inform no relationship between (et-1, et), in reverse to alternative hypotheses which shows:

$$e_t = f(e_t - 1)$$

To test these two hypotheses, we calculate (D.W) as follow:

$$D.W = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^{\,2}}$$

After calculating the value of D.W, we will compare it with (du, dl) scheduled, to judge on the existence or not existence of autocorrelation problem, the (dl) would be the lowest value, and (du) is the highest as follow:

If: D. W < dl → positive autocorrelation

If: dl ≤ D. W ≤ du → test not definitive

If: du ≤ D.W ≤ 4 – du → no autocorrelation

If: 4 – du ≤ D.W ≤ 4- dl → test not definitive

If: 4 ≤ D. W ≤ 4 - dl → negative autocorrelation

The values will be between (0 ≤ D.W ≤ 4).

2. The Results and Debate

In this part, the data of the productions wheat, Barley and Rice have been used from 1986 – 2008 in Sulaimani provinces. Some applications have been done by using instant statistical program (Minitab 11 for Windows), and a special program has been prepared, on this program the prediction of production is calculated:

First- Wheat productivity

Tab.1: The area, rain and production of Wheat

Years	Rain(mm)	Area(Acre)	Production(tones)	One Acre Productivity(tones)
1986-1987	566.2	522300	116734.05	0.22
1987-1988	781.7	565000	129441.5	0.23
1988-1989	972.8	464900	110177.58	0.25
1989-1990	484.4	671512	98040.752	0.15
1990-1991	710	742709	185677.25	0.25
1991-1992	720.5	428720	95604.56	0.27
1992-1993	729.3	240227	63900.382	0.27
1993-1994	748.9	153999	40963.734	0.27
1994-1995	903.2	136990	27808.97	0.20
1995-1996	498.5	293651	92500.065	0.31
1996-1997	941	130648	18029.424	0.14
1997-1998	930.6	595250	89287.5	0.15
1998-1999	1007.5	795343	136003.653	0.17
1999-2000	873.7	840506	245427.752	0.29
2000-2001	952.8	881850	291010.5	0.33
2001-2002	659.1	657532	83506.564	0.13
2002-2003	790.1	565508	120453.204	0.21
2003-2004	854.8	531727	132564.8584	0.25
2004-2005	623.6	410184	47909.4912	0.12
2005-2006	339.4	522447	96600.4503	0.18
2006-2007	499	517902	140299.6518	0.27
2007-2008	512.8	653300	195010.05	0.30

Source: Sulaimani Statistical office

Tab.2: ANOVA Table

S.O.V.	D.F.	S.S	M.S.	F test	P-Value
Model	2	6.53808E10	3.26904E10	21.91	0.0000
Residual	19	2.83536E10	1.4923E9		
Total	21				

Tab.3: Estimation of parameters and statistical tests

Parameter	Estimation	Durbin Watson	R-Square	R-Square (adjusted for d.f)
Constant	-29592.2	1.93079	69.7511	66.567
Rain	22.2337			
Area	0.252426			

Linear trend model

$\hat{Y}$ =-29592.2 + 22.2337*Area + 0.252426 * Rain

t= (-0.77391) (0.499192) (6.57041)

Tab.4: future forecasting for wheat

ID	Period	Forecasting value
1	2009	140036
2	2010	142106
3	2011	144177
4	2012	146247
5	2013	148318
6	2014	150388
7	2015	152459
8	2016	154529
9	2017	156600
MAD = 47470		

According to the data in table (4), for prediction of future production of Wheat for the years 2009 - 2017, it discerns that the result is coinciding with the economic theory. The prediction results for future years are bigger than previous percentage with small disparity for year 2007-2008. This means that the prediction results would not be affected by one rate, but it will be affected by all rates for all years. With an increase in area planted and an increase in the level of rain, the productivity of this product will increase.

__Tab.5: The area, rain and production of Barely__

Years	Rain(mm)	Area(Acre)	Production(tones)	One Acre Productivity(tones)
1986-1987	566.2	188696	37739.0113	0.20
1987-1988	781.7	192218	36521.03556	0.19
1988-1989	972.8	222027	41297.022	0.19
1989-1990	484.4	245702	14742.12	0.16
1990-1991	710	284946	63542.958	0.22
1991-1992	720.5	144473	21382.004	0.15
1992-1993	729.3	43050	8523.9861	0.20
1993-1994	748.9	17572	3478.99242	0.19
1994-1995	903.2	34536	7044.99864	0.20
1995-1996	498.5	55760	15054.97696	0.27
1996-1997	941	23908	2056.088	0.09
1997-1998	930.6	82530	8253	0.10
1998-1999	1007.5	166485	30633.07352	0.18
1999-2000	873.7	184307	22683.0311	0.12
2000-2001	952.8	212300	35197.0047	0.17
2001-2002	659.1	69550	11823.01315	0.17
2002-2003	790.1	146056	22492.624	0.15
2003-2004	854.8	247545	49583.2635	0.20
2004-2005	623.6	231191	23119.1	0.10
2005-2006	339.4	330197	48496.03339	0.15
2006-2007	499	430220	103209.778	0.24
2007-2008	512.8	489109	143162.2043	0.29

Source: Sulaimani Statistical office

Tab.6: ANOVA Table

S.O.V.	D.F.	S.S	M.S.	F-test	P-Value
Model	2	1.97777E10	9.88885E9	45.93	0.0000
Residual	19	4.09076E9	2.15303E8		
Total	21				

Tab.7: Estimation of parameters and statistical tests

Parameter	Estimation	Durbin Watson	R-Square	R-Square (adjusted for d.f)
Constant	-17842.1	1.2059	82.8612	81.0571
Rain	0.248162			
Area	8.65827			

General Linear trend model

$\hat{Y}$ =-17842.1 + 0.248162*Area + 8.65827 * Rain

t= (-1.03361) (8.71288) (0.45508)

Tab.8: future forecasting for barley

ID	Period	Forecasting value
1	2009	59693.8
2	2010	61919.9
3	2011	64146.1
4	2012	66372.3
5	2013	68598.5
6	2014	70824.7
7	2015	73050.9
8	2016	75277.1
9	2017	77503.3
	MAD=23212	

According to the data in table (8) for prediction of future production of Barely for the years 2009 - 2017, it discerns that the result is coinciding with the economic theory. In comparison, it appears that in general the planted area and the level of rain have been increased in recent years. It also appears that predicted value increases year after year. With an increase in the area and the rain the productivity of this product will increase.

Tab.9: The area, rain and production of Rice

Years	Rain(mm)	Area(Acre)	Production(tones)	One Acre Productivity(tones)
1986-1987	566.2	3508	2431.044	0.69
1987-1988	781.7	3074	1847.474	0.60
1988-1989	972.8	3292	2469	0.75
1989-1990	484.4	1403	1050.847	0.74
1990-1991	710	1750	1368.5	0.78
1991-1992	720.5	1525	1067.5	0.70
1992-1993	729.3	154	100.1	0.65
1993-1994	748.9	63	44.1	0.70
1994-1995	903.2	75	45	0.60
1995-1996	498.5	324	277.344	0.85
1996-1997	941	495	346.5	0.70
1997-1998	930.6	3950	2765	0.70
1998-1999	1007.5	16765	6991.005	0.47
1999-2000	873.7	28760	12654.4	0.44
2000-2001	952.8	33051	17252.622	0.52
2001-2002	659.1	32822	17428.482	0.53
2002-2003	790.1	11000	5554.12	0.50
2003-2004	854.8	8545	5203.13595	0.61
2004-2005	623.6	3333	160.03398	0.50
2005-2006	339.4	1844	723.60404	0.39
2006-2007	499	6845	3454.1239	0.50
2007-2008	512.8	4523	2663.18763	0.59

Source: Sulaimani Statistical office

Tab.10: ANOVA Table

S.O.V.	D.F.	S.S	M.S.	F-test	P-Value
Model	2	5.61894E8	2.80947E8	597.05	0.000
Residual	19	8.94063E6	470559.0		
Total	21				

Tab.11: Estimation of parameters and statistical tests

Parameter	Estimation	Durbin Watson	R-Square	R-Square (adjusted for d.f)
Constant	414.782	1.73135	98.4338	98.2689
Rain	0.491224			
Area	0.236781			

General Linear trend model

$\hat{Y}$ =414.782 + 0.491224*Area – 0.236781 * Rain

t= (0.691206) (33.0908) (-0.286264)

From the model we see the negative sign and this will prove a reality that the Rice does not need rain. And in most years the increase in rain did not lead to an increase in production of this product.

Tab.12: future forecasting for Rice

ID	Period	Forecasting value
1	2009	6849.82
2	2010	7105.94
3	2011	7362.06
4	2012	7618.18
5	2013	7874.31
6	2014	8130.43
7	2015	8386.55
8	2016	8642.67
9	2017	8898.79
	3474= MAD	

According to the data in this table for prediction of future production of Rice for the years 2009 to 2017, it discerns that the result is coinciding with the economic theory. Through comparison, it shows that in general the area and amount of rain have been increased in recent years, and we found predicted value has increased year by year. Through an increase in area seeded the production of this product will increase.

3. Conclusion

1. In the production of Wheat, considering (Y) as dependant variables. A combination is produced, also the result and statistical tests (F, R^2 and T) and standard test (D. W.) that has been explained in previous chapter, is broadcasted. Then after the independent variables has been viewed one after another.

This combination has passed statistical tests (F, R^2, T), the explanatory value (R^2) of this combination was (69.75%) which means a substantial change in dependent variable (yearly production of Wheat) because of changes in the two independent variables (Area, Rain). Beside this, there might be other variables affecting dependent variable which are not taken into account.

The test is passed F – test too, where if its accounted value (21.91) larger than its scheduled value (3.52) by (5%), then we should accept the model and refuse null-hypotheses, which specifies that all real coefficient values are equivalent and equal to zero. Or not the independent variables together have significance effect on dependent variable. Regarding econometric testing, the model has passed D. W – test in the area where autocorrelation dose not existed. This indicates no autocorrelation problems between the variables in first degree. Or there is no relationship between (et-1, et). Therefore, we accept null-hypotheses in this model in terms of economic theory. As described in equation below:

$$Y = -29592.2 + 22.2337 * Rain + 0.252426 * Area$$

It's clear from above that the function is agreed with economic theory, which clarifies positive relationship between dependent and independent variables with an increase in the area devoted for planting Wheat, also an increase in Rain will lead to an increase in yearly production of Wheat. In other words, this will lead to an increase of productivity of one Acre of land seeded. As long as the results is positive, it will prove the validity of the relationship between the two variables. The coefficient of constant value came negative in this model; this can be returned to the political circumstances of that period for example the expatriation of Kurdish people in year 1991, leaving lands without sowing. This can be interpreted as impossibility of production process without using inputs. Finally, if the value of coefficient of constant value was too large this is an indication of the size of externality that can not be explained by eliminated variables from the model.

2. In the production of Barely, we assume (Y) as dependent variables, we also produce a combination. The combination has been tested and passed the statistical tests (F, R^2, T). The combination's explanatory power (R^2) has reached (82.16%), indicating that the significant changes in dependent variables (Y, or yearly production of Barely) is due to changes in independent variables (Rain, Land). The other variables that has not been taken into account have their effect on dependent variable, as the model has passed the (F) test, its accounted value is (45.93) larger than its scheduled value (3.52) by standard measure of (5%). Encouraging us to accept the model and refuse null-hypotheses, which refers the fact that the real value of coefficients are equivalent and equal to zero, i.e. the independent variables together, have no effect on dependent variable.

Concerning the econometric testing, the model has passed (D.W) test, where its value is laid in the area where no autocorrelation existed. This means that there are no autocorrelation problems between variables in first degree. Or there is no relationship between (et-1, et), therefore we accept null-hypotheses. The model in economics view is shown below:

$$Y = -17842.1 + 0.248162 * Rain + 8.65827 * Area$$

From the above, it is clear that the equation is in agreement with the economic theory, which specifies affirmative relationship between dependent and independent variables. As the area planted increased, and the level of Rain increases, the yearly production of Barely increases, the productivity of a hectare of seeded land will increase. Also the positive sign of independent variable's coefficient is prove of the affirmative relationship between dependent and independent variables.

The coefficient of constant value again appeared in negative sign in this linear model, which can be explained by the existence of some abnormal data in the time series: in 1996 due to oil-food exchange agreement, that led to a reduction in the production of Barely and the area seeded by Barely. Where in 1994 – 1995 an area of (212300) Mile square were seeded by Barely and this figure decreased to only (69550 M2) in 1996. It can also be described as unfeasibility of production when eliminating the inputs. Finally, if the value of coefficient of contingence is high, this is an indication of the size externals that can only be explained through the eliminated variables form the model.

3. After obtaining a combination as a dependent variable (Y), the data of Rice has passed the statistical tests (F, R^2, T). The explanatory power of the combination (R^2) has reached (98.43%), meaning that the high part of the changes in dependent variables (yearly production of Rice), can be backed to the changes in independent variables (Area, Rain), along with the effect of other external variables which are not taken into account. But the proportion of these externals is small and has reached (1.57%). The model also passed F – test, where its accounted value is (597.05) bigger than its scheduled value of (3.52), which leads us to accept the model and refuse null-hypotheses, that confirms no significance effect for the independent variables on dependent.

The model is also passed the econometric tests, it passed D.W test, where its value laid in the area of no autocorrelation, the value was (1.73) close to (2), the median of the area that autocorrelation do not exist. This implements no autocorrelation problems between variables in first degree, i.e. no relationship between (et-1, et), with acceptance of null-hypotheses in economics point of view, as illustrated below:

$$Y = 414.782 + 0.491224 * Area – 0.236781 * Rain$$

It appears from the equation that, the model coincides with the economic theory which states that there is a positive relationship between independent variable A (Area), holding that an increase in the area planted will lead to an increase in the yearly production of Rice. Here, the level of rain is not agreed with the economic theory; therefore, the coefficient of this variable showed a negative sign, but this can be returned to the production conditions of this product. This product can only be planted in places and surfaces covered by water. Thus, it dose not need further amounts of rain. The constant value coefficient showed a positive sign in this linear function. This proves of none production in case of removal of factor inputs especially, the area planted. The tiny value of contingence in compare to two previous equations will prove the smallness of externals that has not been explained by independent variables of the model.

References:

[1] ABDULHUSSAIN, A. The contribution of technology spillover in industrial sector in Iraq, university of mustansria, college of administration and economics, 1992.

[2] ABDULKARIM, M. Econometric method for analyzing cement for the period of (1965-1983), Journal of economics, review 26, Baghdad, 1985.

[3] DOMENIC, S. Statistics and econometrics, McGraw hill pub, 1982.

[4] JAMES, H. and MARK, W. Introduction to econometrics, brief edition, Pearson International edition, 2006.

[5] MILTON, J. and ARNOLD, C. Introduction to Probability and Statistics, 3rd edition, McGraw-Hill Book Company, Singapore, 1995.

[6] STUDEMANMUND, A. Using econometrics, fifth edition, Pearson International edition, 2006.

[7] TALIB, H. Introduction to econometrics, university of mousl, 1991.

[8] WOOLDRIDGE, J. Introductory Econometrics: A Modern Approach. 2nd Ed, Thomson, South- Western, 2003.

Acknowledgement:

Financial support from grant n. 137 319/2006-ODEV; MZV Studium v doktorském studijním programu pro irácké odborníky v oblasti průmyslu, obchodu a služeb.

Contact address:

MSc. EZAT SABIR ESMAEEL
University of Pardubice
Faculty of economics and Administration
E-mail: azizsabr@yahoo.com
Phone Number: 009647701366093