Forecasting Wheat Production in Pakistan

Falak Sher and Eatzaz Ahmad[*]

Abstract

This study analyzes the future prospects of wheat production in Pakistan. Parameters of the forecasting model are obtained by estimating a Cobb-Douglas production function for wheat, while future values of various inputs are obtained as dynamic forecasts on the basis of separate ARIMA estimates for each input and for each province. Input forecasts and parameters of the wheat production function are then used to generate wheat forecasts. The results of the study show that the most important variables for predicting wheat production per hectare (in order of importance) are: lagged output, labor force, use of tractors, and sum of the rainfall in the months of November to March. The null hypotheses of common coefficients across provinces for most of the variables cannot be rejected, implying that all variables play the same role in wheat production in all the four provinces. Forecasting performance of the model based on out-of-sample forecasts for the period 2005-06 is highly satisfactory with 1.81% mean absolute error. The future forecasts for the period of 2007-15 show steady growth of 1.6%, indicating that Pakistan will face a slight shortage of wheat output in the future.

JEL Classification: C13, C22, Q13, Q16

Keywords: Wheat, ARIMA, Production Function, Pakistan

1. Introduction

Wheat is the main staple food item in Pakistan. The share of this single item in total household consumption in Pakistan is about 9%. Among rural households wheat is the largest single consumption item, while among urban households it is the second largest consumption item following housing.[1] Over the years, factors including water shortages, increases in the input prices, extraordinary drought conditions, etc., have

---

[*]The authors are respectively Lecturer at Department of Economics, University of Sargodha, Pakistan and Professor of Economics, Quaid-i-Azam University, Islamabad, Pakistan.
[1] See Government of Pakistan (2005).

affected the wheat crop, although in the recent past Pakistan has experienced good wheat crops. Due to favorable conditions, it became possible to compile healthy strategic stocks.

Crops in general, and wheat in particular, provide leakages through which it is possible to provide stimulus to economic growth in other sectors of the economy. Wheat cultivation has been suffered from various problems, such as shortage of irrigation water, low yields, traditional methods of farming, increases in input prices, shortage of good quality key inputs, and low usage of modern technology. Pakistan has also experienced fluctuations in wheat production; farmers suffered heavy losses in the few years that the wheat crop was in surplus. Also, a negative relationship has been observed between flour prices and wheat production.

Several studies have been conducted to formulate a wheat-forecasting model and to make wheat forecasts in Pakistan. Most of these studies were conducted during the 1970s and 1980s when Pakistan faced shortages of wheat production and policy makers were concerned with food security, especially for staple foods. The most commonly used inputs in these models were rainfall, fertilizer, temperature, tractors and labor. The literature has proposed several alternative measures of rainfall. For example, Azhar et al (1972, 1974) used rainfall during the November to January period, Qureshi (1974) used three variables to capture rainfall: rainfall from July to September, rainfall from October to December and 'maximum effective' rainfall from January to March. Chaudhary and Kemal (1974) found that deviation of rainfall from normal levels during the period from July to January was the most appropriate rainfall variable for wheat production in irrigated areas of Pakistan. This study, along with another study by Griffiths *et al.* (1999), concluded that the choice between actual rainfall and the deviations from normal rainfall was a matter for empirical investigation and the results were not robust.

The empirical literature finds strong correlation of wheat production with fertilizer use. In the absence of any direct measure of fertilizer use for the wheat crop at the aggregate level, the literature adopts alternative procedures such as the purchase of fertilizer during the sowing season for wheat or a fertilizer consumption based on the share of the wheat crop in total cropped area (see Azhar *et al* (1972, 1974); Mukhtar and Mukhtar (1988); Saleem (1989)).

Other variables that had significant effects on wheat output include tractorization (Salam (1981)) and labor (Mukhtar and Mukhtar (2003); Salam (1981)).

Most of the studies cited above also derive wheat forecasts on the basis of the models estimated. A notable study is by Iqbal *et al.* (2005). According to this study, the estimated ARIMA model showed that production of wheat would grow to 29.77 million tones in the year 2022. The study concluded that the expected growth was low and that the scope for higher area and production laid in adequate government policies regarding wheat cultivation in the country.

Thus wheat forecasts have been made either for only one province or two but not for all the four provinces of Pakistan. As mentioned earlier, wheat forecasts that were made for individual provinces have a history in 1970s. A recent study by Iqbal *et al.* (2005) provided wheat forecasts at the aggregate level in Pakistan. Thus a study is needed that can provide updated estimates of wheat forecasts both for Pakistan as well as its four provinces that can predict future trends in the production of wheat over the next few years.

The present study determines the future prospects for wheat production in Pakistan as well as in its four provinces using the past trends. The objectives of the disaggregate analysis at the provincial level is two-fold. First, the province-wise disaggregated data increases the size of the sample four times as compared to the country level aggregate data, thereby increasing the reliability of parameter estimates of the forecasting model. Second, the province-wise data also allow the possibility of variations in parameters of the forecasting model across the four provinces through panel estimation and hence produce more accurate forecasts.

In order to assess the forecast accuracy of the estimated model, the study re-estimates the model after dropping the latest two years' observations (2005 and 2006), then makes wheat forecasts for these two years and compares them with actual production in the same years. Thus, the difference between the actual and predicted production is the wheat forecast error. A positive forecast error means that wheat production has been underestimated and vice versa. Thus, after determining the forecasting ability of the model and finding a basis for future wheat forecasting, the study makes wheat forecasts for the period 2007 to 2015.

The organization of the paper is as follows. In Section 2, the methodology is presented. Data and the estimation procedure are described in Section 3. Results are presented in Section 4. Finally, Section 5 concludes the study.

2. Methodology

One way to determine which inputs are crucial in the production of wheat output is to calculate elasticities of wheat output with respect to these inputs. These elasticities can be found by estimating a production function with an appropriate functional form. For this purpose, the current study estimates a Cobb-Douglas production function. This specification has the advantage that it yields direct estimates of production elasticities with respect to various inputs. For this reason, several previous studies on wheat production have assumed the Cobb-Douglas specification.[2] Thus the regression equation specified is linear in the natural logs of the variables.

Wheat output per hectare is assumed to be a function of labor force per hectare, number of tractors per hectare, fertilizer use per hectare, rainfall in the months of November, December, January, February and March, rainfall in the month of April, weighted standard deviation of rainfall in the months of November, December, January, February and March, and lagged wheat output per hectare. The inclusions of the inputs labor, tractors and fertilizer are easily justified. Rainfall in the November to March period is expected to contribute to wheat output because the supply of canal and tube-well water is limited, especially in NWFP, Baluchistan and northern parts of Punjab. The standard deviation of rainfall is included to consider the potential unfavorable effects of volatile rains as compared to consistent rains. Since April is the wheat-harvesting season in most parts of the country, rainfall during this period is considered harmful. Also note that the use of canal and tube-well water is not included in the production function. The reason is that data on water use is available only on an aggregate basis and it is understood that the water use on wheat can vary to a great extent from the overall water use because: a) different crops have different needs of water and b) water availability during the wheat season varies a lot from the water supply in other seasons like July and August. Thus in the case of water inputs, it is more unrealistic to adopt any procedure for estimating the input use for wheat crop based on the given aggregate data as we have done in case of other inputs like fertilizer.

---

[2] See, for example, Qureshi (1974) and Khan et al. (2003).

The study proposes the following specification of the production function.

$$\ln Y_t = \beta_1 + \beta_2 \ln L_t + \beta_3 T_t + \beta_4 \ln F_t + \beta_5 \ln R1_t + \beta_6 \ln R2_t + \beta_7 \ln SR_t + \beta_8 \ln Y_{t-1} + u_i \quad (1)$$

where

Y_t = Wheat output per hectare

L_t = Labor force per hectare

T_t = Number of tractors per hectare

F_t = Fertilizer use per hectare

$R1_t$ = Weighted average of rainfall in the Months of November, December, January, February and March

$R2_t$ = Weighted average of rainfall in the Month of April

SR_t = Weighted standard deviation of rainfall in the Months of November, December, January, February, and March.

Y_{t-1} = Lagged output per hectare

u = stochastic disturbance term

The future values of various inputs were obtained by estimating a separate Auto Regressive Integrated Moving Average (ARIMA) model for each input for each province. Dynamic forecasts for the required forecasting period are then made using this model. This exercise is performed for each input and for each province separately using the time series data from 1979 to 2006.

To evaluate the ability of the model to accurately forecast wheat output, the entire analysis was conducted using the time period 1979 to 2004, making forecasts for the years 2005 and 2006. These forecasts were then compared with the actual values of wheat output realized in 2005 and 2006 to assess the quality of forecasts.

3. Data and Estimation

3.1 Data

The study uses province-level data for the period 1979 to 2006. Data on wheat output and area under wheat cultivation are taken from the *Agricultural Statistics of Pakistan and Area Production* (By Districts), Ministry of Food, Agriculture and Live-Stock (MINFAL), Government of Pakistan.

The data on the total number of tractors in Pakistan from 1980 to 2006 have been taken from *World Development Indicators*, World Bank, Washington D.C. *Census of Agriculture Machinery* (1975, 1984, 1994 and 2004), and *Census of Agriculture* (1980, 1990 and 2000), Agriculture Census Organization (ACO), Federal Bureau of Statistics, Government of Pakistan.

To calculate the total labor force used in the production of wheat, data on the total population of Pakistan and its four provinces have been taken from the *Economic Survey,* Ministry of Finance, Government of Pakistan. Labor force participation rates and percentages of persons employed in the agricultural sector of Pakistan and its four provinces have been taken from *Labor Force Survey,* Federal Bureau of Statistics, Government of Pakistan.

Data on total fertilizer consumption have been taken from the *Agricultural Statistics of Pakistan,* Ministry of Food, Agriculture and Livestock (MINFAL), Government of Pakistan. Data on the percentage of fertilizer consumption used for the wheat crop have been taken from the *Fifth and Sixth Five Year Plans* and from the *Fertilizer Use Survey*, National Fertilizer Development Centre, Islamabad.

Data on the monthly average rainfall are taken from the *Agricultural Statistics of Pakistan,* Ministry of Food, Agriculture and Livestock (MINFAL), Government of Pakistan.

3.2 Construction of the Variables

Before the construction of various variables used in this study is explained, it may be noted at the very outset that for quite a few variables, specifically fertilizers, tractors and labor, no direct data could be obtained from any source. These variables are constructed on the basis of certain specific assumptions that are explained where used. Thus the data on such variables is likely to suffer from measurement error and hence may introduce bias in the parameter estimates.

Wheat Output Per Hectare

Wheat output per hectare for each province is found by dividing total wheat output in thousand tones in each province by total acreage in thousand hectares.

Fertilizer Per Hectare

To find the consumption of fertilizer for the wheat crop, first province-wise total consumption of three types of fertilizers, nitrogen, potash and phosphate, on all the crops is obtained. Then, their consumption for the wheat crop for each province is calculated according to the assumptions about its use during different Five Year Plans and according to the percentage allocated to wheat by the National Fertilizer Development Centre, Islamabad (NFDC) in different *Fertilizer Use Surveys* during different time periods.

Weighted Average of Rainfall

Initially, average monthly rainfall in millimeters for the months of November, December, January, February, March and April, for the available stations in each province has been taken. The sum of the weighted average of rainfall of the months of November, December, January, February and March has been used as one explanatory variable, while the weighted average of April has been used as a separate explanatory variable. Weighted average of rainfall has been calculated with the following formula.

$$\text{Weighted Average of Rainfall in each month: } \overline{X} = \frac{\sum_{i} W_i R_i}{\sum_{i} W_i}$$

Where R_i and W_i are the rainfall in millimeters and wheat output in thousand tones in each district of each province, respectively. Thus, the districts with more wheat output get a higher weight. For the wheat growing districts for which data on rainfall are not available, the rainfall of the nearest station (nearest in distance) having data on rainfall has been used as a proxy. Note that in the calculation of the weighted average and weighted standard deviation of rainfall we have used wheat output, rather than the area under wheat cultivation as the weight. The weight based on area assigns more importance to the rainfall that falls on a larger area whereas the weights based on wheat output assign more weight to the rainfall that falls on areas producing more wheat and we have preferred the latter weighting scheme. Note that the two weighting schemes produce different results to the extent that the yield per hectare varies across districts.

Weighted Standard Deviation of Rainfall

The weighted standard deviation of rainfall for the months of November, December, January, February and March has been calculated for

each province. First, the weighted standard deviations of rainfall for these months are calculated and then the resulting figures are summed to use it as an explanatory variable. Weighted standard deviation of rainfall for each month for each province has been calculated with the help of the following formula.

Weighted Standard Deviation of Rainfall:

$$S = \sqrt{\frac{\sum_i W_i \left(X_i - \overline{X}\right)^2}{\sum_i W_i}}$$

Where W_i is wheat output in thousand tones in each district of each province; X_i is the rainfall of a particular district or the district that was used as a proxy in each province and $\overline{X}$ is the weighted average of rainfall in each month in each province (as calculated above).

Tractors Per Hectare

Data on the total number of tractors in Pakistan from 1980 to 2004 has been obtained from the *Word Development Indicators 2006*. From the *Census of Agriculture Machinery* and *Census of Agriculture*, conducted by the Agriculture Census Organization (ACO), Federal Bureau of Statistics, the share of agricultural tractors in total tractors for each province can be roughly calculated. The *Census of Agriculture Machinery* was conducted in the years 1975, 1984, 1994 and 2004 while the *Census of Agriculture* was conducted in the years 1980, 1990, and 2000. The number of agricultural tractors for the years between the census years has been obtained by exponential interpolation.

Labor Force Per Hectare

The labor force used in the production of wheat, on which no direct data are available from any source, has been calculated as follows. First of all, figures on the total population of Pakistan have been obtained from various issues of the *Economic Survey*. Exact population figures for Pakistan and its four provinces were only available in the census years. These censuses were conducted in 1972, 1981 and 1998. Figures on Pakistan's population for the years that lie between different censuses years are obtained by exponential interpolation.

From the *Labor Force Survey*, conducted by the Federal Bureau of Statistics in different years, data is taken on the total labor force participation rates for each province of Pakistan. From the *Labor Force Survey*, the percentage of employed persons for each province in the agricultural sector can be known. The total labor force employed in agriculture has been calculated by using the labor force participation rates for each province and then applying the percentage of employed persons in the agriculture sector.

From the *Agricultural Statistics of Pakistan*, published by the Ministry of Food, Agriculture and Livestock (MINFAL), total cropped area and area under wheat cultivation are obtained. Fiinally, an estimate for the labor force engaged in growing wheat is obtained by multiplying the proportion of cropped area allocated to the wheat crop by the total labor force in agriculture.

3.3 Estimation Procedure

Equation (1) is initially estimated for each province separately, yielding a total of 32 parameter estimates. The Wald coefficient test is then applied to find common coefficients across provinces. The coefficient of that variable is taken as common for which the null hypothesis of equal coefficients is accepted with the maximum probability value among the variables. The study again applies the Wald test and notes the probability values for the remaining null hypotheses. Again, the coefficient of that particular variable is taken as common among the remaining ones for which null hypothesis is accepted with the maximum probability value. This procedure is continued until null hypotheses for all the remaining variables are rejected.

The study then estimates ARIMA equations for all the independent variables using the data from 1980 to 2004 and makes forecasts for 2005-06. On the basis of these predicted inputs along with the estimated coefficients of the production function, forecasts for wheat output per hectare are made for the years 2005 and 2006. In the specification of the ARIMA model, we must first choose the order of integration (I), and the orders of autoregressive (AR) and moving average (MA) terms. The order of integration is based on the standard ADF test for unit roots. The orders of AR and MA terms are determined on the basis of careful application of the standard diagnostic procedures that include: a) the study of correlogram along with Q-statistics to determine potential orders; b) estimation of the potential model and consideration of t-statistics of parameters associated with various AR and MA terms and re-estimation of the model after excluding redundant variables; and c) application of AIC and SBC criteria in case two competing specifications seem equally good on the basis of correlogram, Q statistics and t-statistics.[3]

---

[3] See Enders (2004) for more details on the subject.

After fine-tuning the forecasting model and assessing its quality, the full sample (1979 to 2006) is used to forecast wheat output for the period of 2007-15.

4. Results and Discussions

4.1 Elasticities of Wheat Output

Table-1 shows the elasticities of dependent variable (wheat output per hectare) with respect to the independent variables.

Table-1: Elasticities of Wheat Output With Respect to Inputs

Variable	Baluchistan	NWFP	Punjab	Sindh
Acreage	1.45	1.32	1.50	1.75
	(2.66)*	(2.47)**	(2.62)*	(3.15)*
Labor Per Hectare	0.11	0.11	0.11	0.11
	(2.41)**	(2.41)**	(2.41)**	(2.41)**
Tractors Per Hectare	0.078	0.078	0.078	0.078
	(1.89)***	(1.89)***	(1.89)***	(1.89)***
Fertilizer Per Hectare	0.016	0.016	0.016	0.016
	(0.39)	(0.39)	(0.39)	(0.39)
Mean Rainfall Nov-Mar	0.044	0.044	0.044	0.044
	(4.32)*	(4.32)*	(4.32)*	(4.32)*
Mean Rainfall April	-0.23	-0.230	-0.23	-0.23
	(-2.49)**	(-2.49)**	(-2.49)**	(-2.49)**
SD Rainfall Nov-Mar	-0.013	-0.013	-0.013	-0.013
	(-1.04)	(-1.04)	(-1.04)	(-1.04)
Lagged Output Per Hectare	0.7405	0.3709	0.6664	0.4936
	(8.36)*	(2.53)**	(6.60)*	(4.06)*
System R-Squared	0.90			
System Adjusted R-Squared	0.89			
System Durban Watson	2.37			

Notes:

1. The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

2. Since all the parameters are estimated from one equation allowing for changes in parameters across provinces using a Fixed-Effects Model, there is one value for each of the R-square, adjusted R-square and Durbin-Watson statistics.

The results show that the values of the intercept for the four provinces of Pakistan: Baluchistan, NWFP, Punjab and Sindh, are 1.45, 1.32, 1.50 and 1.75, respectively. These intercept values show variations in the level of output per hectare in the four provinces of Pakistan. The reason for these variations is differences in the climate, nature of soil, variation in the distribution of rainfall not captured by the mean and standard deviation, and the temperature, etc. It is also apparent from the table that a one percent increase in labor force per hectare leads to a 0.11 percent increase in wheat output per hectare in all the four provinces of Pakistan. The results also show that there is a positive relationship between the number of tractors and wheat output per hectare. A one percent increase in the total number of tractors leads to a 0.078 percent increase in wheat output per hectare.

Similarly, a one percent increase in the application of fertilizer per hectare leads to a 0.016 percent increase in wheat output per hectare in all the four provinces of Pakistan. However, this relationship is not statistically significant. A possible reason for this may be that farmers may not know at what stage and in how much quantity a particular type of fertilizer should be applied for a good crop. The results further show that a one percent increase in rainfall over the November to March period leads to an increase of 0.04 percent in wheat output per hectare. However a one percent increase in average monthly rainfall in the month of April (during which harvesting as well as threshing takes place) leads to a 0.23 percent decrease in wheat output per hectare.

The results also show that a one percent increase in the level of rainfall deviations, as measured by the sum of the standard deviations in rainfall from November to March leads to a 0.013 percent reduction in wheat output per hectare. However, this relationship is not statistically significant. This means that it is the level of rainfall, rather that its fluctuations, which play an important role in determining wheat output per hectare.

Finally, the results show that wheat output per hectare depends significantly on the output level in the previous year. There may be different reasons for this. Firstly, if the farmers enjoy good harvest this year, their income level increases and now they are in a position to spend more on the next crop as the income from one crop is used for the expenditures on the next one or two crops. If this is true, then due to investment on the wheat crop next year, farmers are expected to enjoy a good wheat harvest next year. Secondly, it is possible that farmers spent more time in looking after the crop in the form of better irrigation by private sources and in the forms of application of pesticides etc. as they expect to receive a high support

price for their crop. Another interpretation of the significance of lagged output is that it is capturing technology improvements in the inputs, better quality tractors, fertilizers, general production techniques, etc.[4]

It is also apparent from the results that, except the intercepts and coefficient of lag output per hectare, all the other coefficients are common for the four provinces. The value of R-squared is 0.90 showing that 90 percent of the variation in the dependent variable is explained by the included explanatory variables. The value of the Durban Watson statistic is 2.37, which falls within acceptable limits.

4.2 Estimates of ARIMA Models

First of all we determine the order of integration of all the variables. The application of ADF tests indicates that the dependent variable and included explanatory variables are non-stationary. Furthermore, the first differences of all the variables are stationary. In other words, all the variables are integrated of order one.

The future values for the various inputs are obtained by estimating a separate Auto Regressive Integrated Moving Average (ARIMA) model for each input. Dynamic forecasts for the required forecasting period are then made using this model. This exercise is performed for each input and for each province separately, first using the time series data from 1979 to 2004 to get the future values of various inputs for the period of 2005-06 and then for the period 1979 to 2006 to make forecasts for the period of 2007-15. The reason for estimating the ARIMA model for various inputs for the shorter time period is that the estimates using the 1979-2004 data help in evaluating the ability of the model to forecast wheat output. These forecasts are then compared with actual values of wheat output realized to assess the quality of forecasts. Finally, the full sample (1979 to 2006) is used to forecast wheat output for the period 2007-15.

The results of the parameter estimates of the ARIMA model equations are shown in Tables 2A - 2D for the period 1979 to 2004 and Tables 3A - 3D for the period 1979 to 2006. One can see that out of 152 parameters, 114 are statistically significant. Further scrutiny establishes that eighty parameters are significant at the 1% level, twenty-three parameters are significant at the 5% level, and eleven are significant at the 10% level of significance. Thus the statistical performance of all the estimated models appears quite impressive.

---

[4] The authors are thankful to one of the referees of this paper for suggesting this interpretation.

Table-2 A: Estimates of ARIMA Models for 1979-2004 (Baluchistan)

Variable	Intercept	AR(1)	MA(1)	MA(2)	D.W
Acreage	0.0292 (1.39)	-0.5368 (-3.05)*			2.11
Labor Per hectare	0.0303 (1.68)***	-0.7509 (-5.20)*			1.87
Tractors Per Hectare	0.0623 (5.62)*				1.04
Fertilizer Per hectare	0.0600 (14.70)*		-0.7367 (-5.60)*		2.14
Mean Rainfall Nov-Mar	-0.0053 (-0.09)			-0.5480 (-2.92)*	2.07
Mean Rainfall April	0.0009 (0.14)	-0.5517 (3.00)*			2.26
SD Rainfall Nov-Mar	0.0012 (0.03)		-0.3267 (-1.58)	-0.5570 (-2.55)*	1.90

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

Table-2 B: Estimates of ARIMA Models for 1979-2004 (NWFP)

Variable	Intercept	AR(1)	MA(1)	MA(2)	D.W
Acreage	0.0024 (0.29)		-0.0004 (-0.002)		1.99
Labor Per Hectare	0.0291 (2.55)**	-0.7939 (-5.29)*			1.95
Tractors Per Hectare	0.0525 (4.54)*				0.92
Fertilizer Per Hectare	0.0448 (13.26)*			-0.9791 (-9.05)*	2.50
Mean Rainfall Nov-Mar	-0.0149 (-0.72)	-0.7566 (-5.10)*		-0.9072 (-11.03)*	1.83
Mean Rainfall April	0.0041 (1.81)***		-0.9897 (-2673)*		1.70
SD Rainfall Nov-Mar	-0.0010 (-0.07)		-0.9659 (-14.03)*		1.85

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

Table-2 C: Estimates of ARIMA Models for 1979-2004 (Punjab)

Variable	Intercept	AR(1)	MA(1)	MA(2)	D.W
Acreage	0.0091 (6.04)*	-0.7886 (-5.76)*		0.9075 (-4.68)*	
Labor Per Hectare	0.0268 (20.97)*	0.3250 (1.45)	-0.9509 (-17.28)*		
Tractors Per Hectare	0.0659 (2.01)***	0.6320 (2.53)**			
Fertilizer Per Hectare	0.0485 (2.75)**				
Mean Rainfall Nov-Mar	-0.0001 (-0.009)			-0.9791 (-8.07)*	
Mean Rainfall April	0.0065 (2.87)*		-0.9460 (-10.03)*		
SD Rainfall Nov-Mar	0.0134 (0.727)	-0.6349 (-3.67)*		-0.9392 (-11.40)*	

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

Table-2 D: Estimates of ARIMA Models for 1979-2004 (Sindh)

Variable	Intercept	AR(1)	AR(2)	MA(1)	MA(2)	Dummy	D.W
Acreage	0.0125 (2.12)**					-0.3571 (-11.87)*	1.78
Labor Per Hectare	0.0172 (1.69)***	-0.5928 (-3.29)*					1.60
Tractors Per Hectare	0.0483 (4.13)*						0.90
Fertilizer Per Hectare	0.0340 (13.55)*			-0.9324 -(14.99)*			1.27
Mean Rainfall Nov-Mar	0.0040 (0.11)	-0.5046 (-2.62)**			-0.9537 (-23.87)*		2.02
Mean Rainfall April	0.0013 (0.16)	-0.4767 (-2.499)**					2.23
SD Rainfall Nov-Mar	-0.0007 (-0.07)	-0.4948 (-2.52)*	-0.6500 (-3.45)*	-0.9894 (-2164)*			1.76

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

Table-3 A: Estimates of ARIMA Models for 1979-2006 (Baluchistan)

Variable	Intercept	AR (1)	MA (1)	MA (2)	D.W
Acreage	0.0294 (1.25)	-0.5317 (-3.10)*			2.06
Labor Per Hectare	0.0296 (1.75)***	-0.7273 (-5.52)*			1.94
Tractors Per Hectare	0.0621 (6.10)*				1.63
Fertilizer Per Hectare	0.0599 (15.75)*		-0.7316 (-5.77)*		2.15
Mean Rainfall Nov-Mar	-0.0182 (-0.28)			-0.4849 (-2.18)**	2.05
Mean Rainfall April	0.0002 (0.03)	-0.560 (-3.21)*			2.26
SD Rainfall Nov-Mar	0.0147 (-0.37)		-0.6930 (-4.27)*		1.72

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

Falak Sher and Eatzaz Ahmad

Table-3 B: Estimates of ARIMA Models for 1979-2006 (NWFP)

Variable	Intercept	AR (1)	MA (1)	MA (2)	D.W
Acreage	0.0012 (0.16)				1.97
Labor Per Hectare	0.0256 (2.22)**	-0.7281 (-4.96)*			1.97
Tractors Per Hectare	0.0514 (4.83)*				1.47
Fertilizer Per Hectare	0.0447 (10.63)*			-0.9192 (-31.99)*	2.34
Mean Rainfall Nov-Mar	-0.0101 (-0.53)	-0.8326 (-5.75)*		-0.8817 (-14.07)*	1.80
Mean Rainfall April	0.0039 (1.68)***		-0.9538 (-12.05)*		1.70
SD Rainfall Nov-Mar	-0.0820 (-2.58)**		-1.4956 (-4.97)*		1.73

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

Table-3 C: Estimates of ARIMA Models for 1979-2006 (Punjab)

Variable	Intercept	AR (1)	MA (1)	MA (2)	D.W
Acreage	0.0092 (19.54)*	-0.8638 (-11.06)*		0.9796 (-4732)*	1.79
Labor Per Hectare	0.0256 (9.60)*	0.4346 (2.29)**	-1.76 (-4.78)*		2.44
Tractors Per Hectare	0.0522 (3.11)*	0.3302 (1.64)			2.13
Fertilizer Per Hectare	0.0533 (3.12)*				2.29
Mean Rainfall Nov-Mar	-0.0033 (0.15)	0.2901 (1.55)	-0.9618 (-20.67)*		1.72
Mean Rainfall April	0.0057 (0.57)		-0.9895 (-6.90)*		1.86
SD Rainfall Nov-Mar	0.0043 (0.28)			-0.9377 (-23.27)*	1.88

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

Table-3 D: Estimates of ARIMA Models for 1979-2006 (Sindh)

Variable	Intercept	AR (1)	AR (2)	MA (1)	MA (2)	Dummy	D.W
Acreage	0.0138 (2.46)**					-0.3584 (-12.03)*	1.73
Labor Per Hectare	0.0171 (1.47)	-0.9321 (-7.64)*		0.5993 (2.28)**			1.97
Tractors Per Hectare	0.0470 (4.37)*						1.43
Fertilizer Per Hectare	0.0364 (10.54)*	0.3001 (1.32)		-0.9510 -(21.55)*			1.82
Mean Rainfall Nov-Mar	0.0313 (0.65)	-0.4258 (-2.05)**			-0.8732 (-6.74)*		1.87
Mean Rainfall April	0.0026 (2.06)***			-0.9895 (-4492)*			1.86
SD Rainfall Nov-Mar	-0.0065 (0.42)	-0.4138 (-2.09)**	-0.7015 (-3.56)*	-0.91 (-6.56)*			1.59

Note: The t-values, shown in parentheses, are marked by *, ** and *** if these are statistically significant at 1%, 5% and 10% levels of significance respectively.

The presence of autoregressive trends as shown by the ARIMA equation results implies that in about half of the cases (26 of 56) there exists a strong autoregressive process of order one, that is, an AR(1) process. This means that the turbulence experienced throughout the time period under consideration is significantly related to the occurrences in the previous period. The AR(1) process has been justified on the basis of geometric decline in the autocorrelation function (ACF). This means that the shocks in output per hectare experienced during a period have a rigid relationship with future output. This effect declines in severity with the passage of time. One can also see that in the province of Sindh, AR(2) is present in the weighted standard deviation of rainfall in the months of November to March for both the periods i.e. in 1979-2004 and 1979-2006.

The moving average (MA) or temporary disturbance terms are also present in most cases. The order of the MA process determines the nature of the one-off relationship between the current and past fluctuations in wheat output. For example, with an MA(1) process, a shock occurring in one period will have an effect on the wheat output per hectare in the next

period. This shock is, however, eliminated from the system within one period. The results show that 20 out of 56 cases experience an MA(1) process while in 14 out of 56 cases an MA(2) process exists.

The results also showed that the dummy variable used in the acreage ARIMA equation for the province of Sindh (representing some shock), is significant at the 1% level of significance. In 2000, wheat output in the province of Sindh declined from 1144.2 thousand tones to 810 thousand tones in 2001. So the use of the dummy variable is justified.

It is also apparent from the results that the intercepts of the estimated ARIMA equations are significant in 32 out of 56 cases. Since the intercept measures the systematic component, it follows from a non-zero intercept that the average growth rate of a particular independent variable is non-zero. Out of 56 intercept estimates, 10 have a negative sign and one (the weighted standard deviation of rainfall from November to March in NWFP) is statistically significant, implying that the average growth rate of this independent variable is negative and significant. On the other hand, the estimates of 46 out of 56 intercepts are positive and 31 are statistically significant, indicating that the average growth rate of these independent variables is positive and significant.

4.3 Results of Forecast Errors for Pakistan and its Four Provinces

Tables 4A – 4E show the results of predicted output, actual output, forecast error (the difference of actual and predicted output) in thousand tones and percent forecast error for Pakistan and its four provinces for 2005, 2006 and combined forecast error (average of 2005 and 2006 error) for 2005-06. As is apparent from the results presented in Tables 4A – 4D, forecast errors as well as percent forecast errors are positive for Baluchistan province, while negative for NWFP and Sindh provinces for the years 2005 and 2006. Combined forecast error for the year 2005-06 is also positive for Baluchistan province and negative for NWFP and Sindh provinces. Forecast error is negative in year 2005 and positive in year 2006 for Punjab province. Combined forecast error for the period 2005-06 for the Punjab province is negative. As far as overall Pakistan is concerned, forecast error is negative in 2005 and positive in 2006. Combined forecast error for the period 2005-06 for overall Pakistan is negative. As forecast error is the difference between actual and predicted wheat output, a positive forecast error means that actual wheat output is greater than its predicted value: therefore wheat output is underestimated and vice versa. In any case, the overall forecast error for Pakistan is small; as a result the model proposed and estimated in this paper performs satisfactorily.

Table-4 A: Wheat Forecast in Thousand Tones (2005-06 Baluchistan)

Years	Actual Output	Predicted Output	Forecast Error	Percent Forecast Error
2005	738.01	637.60	100.41	15.75
2006	715.01	649.90	65.11	10.02
2005-06	1453.02	1287.50	165.52	12.68

Table-4 B: Wheat Forecast in Thousand Tones (2005-06 NWFP)

Years	Actual Output	Predicted Output	Forecast Error	Percent Forecast Error
2005	1070.84	1091.10	-20.26	-1.86
2006	1093.82	1100.60	-6.78	-0.62
2005-06	2164.66	2191.70	-27.04	-1.23

Table-4 C: Wheat Forecast in Thousand Tones (2005-06 Punjab)

Years	Actual Output	Predicted Output	Forecast Error	Percent Forecast Error
2005	16730.01	17375.00	-644.99	-3.71
2006	17065.47	16776.00	289.47	1.73
2005-06	33795.48	34151.00	-355.52	-1.04

Table-4 D: Wheat Forecast in Thousand Tones (2005-06 Sindh)

Years	Actual Output	Predicted Output	Forecast Error	Percent Forecast Error
2005	2243.21	2508.60	-265.39	-10.58
2006	2455.52	2750.30	-294.78	-10.72
2005-06	4698.73	5258.90	-560.17	-10.65

Table-4 E: Wheat Forecast in Thousand Tones (2005-06 Pakistan)

Years	Actual Output	Predicted Output	Forecast Error	Percent Forecast Error
2005	20782.07	21612.3	-830.23	-3.99
2006	21329.82	21276.8	53.02	0.25
2005-06	42111.89	42889.1	-777.21	-1.85

The forecast errors are the largest for the province of Baluchistan followed by Sindh. A possible reason for the large forecasting error for Baluchistan is that this province has been subject to greater volatility in weather conditions, which are not entirely captured by our rainfall variables. Another reason could be that Baluchistan is area-wise the largest province but population-wise (by far) the smallest. Wheat is grown on small farms scattered across rugged lands and, therefore, it is difficult to collect accurate data on inputs and output of wheat. As far as the large forecasting error in Sindh is concerned, there is no obvious interpretation that one can provide. One possible reason could be that wheat growing in this province relies heavily on floodwaters in river Indus, for which we do not have any variable in our model. These floods normally occur in the months of July, August and September and provide a new and fertile layer of earth and the needed moisture for the wheat growing that takes place in September and October.

4.4 Wheat Forecasts for the Period 2007-15

Forecasts for wheat output are made for Pakistan and its four provinces for the period of 2007 to 2015 and are reported in Table-5. The predicted wheat output has a positive trend over the period 2007-2015 for Pakistan and its four provinces.

Table 5: Wheat Forecasts in Thousand Tones (2007-15)

Year	Baluchistan	NWFP	Punjab	Sindh	Pakistan
2007	734	1080	16436	2716	20967
2008	765	1095	16433	2707	21000
2009	833	1098	16512	2689	21132
2010	874	1117	16732	2747	21470
2011	928	1126	16977	2826	21856
2012	972	1144	17306	2855	22276
2013	1020	1155	17638	2913	22725
2014	1065	1172	18026	2992	23255
2015	1112	1184	18410	3078	23785

The growth rate of actual as well as predicted wheat output for Pakistan and its four provinces for the period 2001 to 2015 are reported in Table-6. The table shows actual annual growth rates for the period 2001 to 2006 and predicted growth rates for the period 2007 to 2015. The growth rate of wheat output is positive, ranging from 1.3% in 2004 to 16% in 2001, for the province of Baluchistan for the whole period except in 2005 when it is negative (-3.9%). Wheat output growth rates (actual and predicted) are positive for the province of NWFP, ranging from 0.3% in 2009 to 19.5% in 2003, but were negative in the years 2001, 2004, and 2007.

Table 6: Growth Rates of Wheat Output (Actual and Predicted)

Years	Baluchistan	NWFP	Punjab	Sindh	Pakistan
2001	16.0	-28.5	-6.4	-25.8	-9.7
2002	4.3	16.6	-5.3	-5.6	-4.2
2003	2.2	19.5	5.2	0.4	5.2
2004	1.3	-3.7	1.8	3.0	1.6
2005	-3.9	6.4	11.1	15.5	6.6
2006	1.9	0.9	-3.4	9.6	2.6
2007	13.0	-1.9	-2.0	-1.2	-1.7
2008	4.2	1.4	0.0	-0.3	0.2
2009	8.9	0.3	0.5	-0.7	0.6
2010	5.0	1.7	1.3	2.2	1.6
2011	6.1	0.8	1.5	2.9	1.8
2012	4.7	1.6	1.9	1.0	1.9
2013	5.0	1.0	1.9	2.0	2.0
2014	4.4	1.5	2.2	2.7	2.3
2015	4.4	1.0	2.1	2.9	2.3

Similarly, the growth rate of wheat output was negative in years 2001, 2002, 2006 and 2007 but predicted to be positive for the remaining years for the province of Punjab. The actual and predicted growth rates of wheat output were also negative for the province of Sindh in 2001, 2002, 2007, 2008 and 2009 but is positive for years. As far as Pakistan is concerned, wheat output growth rate is positive except for the three years i.e. 2001, 2002 and 2007 when growth rate was negative.

5. Conclusion

The study finds that lagged output per hectare is the most important factor in determining the current output per hectare in all the four provinces of Pakistan. The other two important variables in determining the current outputs are labor force per hectare and tractors per hectare. The sum of the rainfall in the months of November, December, January, February and March is another important variable in the determination of output.

Wheat forecasts are made for Pakistan and its four provinces using the ARIMA forecasting models for all the inputs for the period of 2005 and 2006. Wheat forecasts errors are negative for the provinces of NWFP, Punjab and Sindh, and positive for the province of Baluchistan, in 2005. These results imply that in 2005, wheat output is overestimated for NWFP, Punjab and Sindh provinces and for overall Pakistan, and underestimated for the Baluchistan province. On the other hand, in 2006, forecast errors are negative for the provinces of NWFP and Sindh, and positive for the provinces of Baluchistan and Punjab. In 2006, wheat output is overestimated for the NWFP and Sindh provinces and underestimated for the Baluchistan and Punjab provinces and for overall Pakistan. As far as overall Pakistan is concerned, wheat output is overestimated and underestimated in 2005 and 2006, respectively. The combined forecast error for overall Pakistan, for the years of 2005 and 2006, is only 1.85%, which implies that on average, wheat output is slightly underestimated for the combined period 2005-2006.

An important conclusion is that wheat forecast errors for Pakistan are mainly determined by the direction and size of forecast error in Punjab, by far the largest contributor to wheat supply in Pakistan. Thus, it is of utmost important to expedite research efforts on making reliable forecasts, especially in the province of Punjab.

Wheat forecasts show positive growth for the period 2007 to 2015, after forecasting for the period of 2005 and 2006. In 2007, the growth rate of total wheat output is negative for the provinces of NWFP and Punjab and for overall Pakistan while it is negative for the province of Sindh in 2008 and 2009.

The forecasts show that in the next eight years wheat output will on average grow at a rate of 1.6%, which is slightly less than the expected growth rate of population. So Pakistan is likely to see a slight shortfall in the supply of wheat, which can be overcome by taking appropriate steps like timely import of wheat, maintenance of strategic wheat reserves, bringing the wheat price for farmers closer to the world price and removal of other distortions.[5]

The study finds that the expected growth in wheat output is the highest in Baluchistan and lowest in NWPF. The growth in Punjab is also expected to be on the lower side. Thus, if the population and income levels continue to grow at the existing pace, Punjab will no longer be able to provide sufficient surplus output to feed the smaller provinces.

---

[5] Since the present study is mainly confined to forecasting based on the production function, it cannot be used to make precise policy prescriptions for which other studies exist.

References

Azhar, B. A., Chaudhary, M. G. and M. Shafique (1972), "A Model for Forecasting Wheat Production in the Punjab," *Pakistan Development Review,* 11(2): 407-415.

Azhar, B. A., M. G. Chaudhary and M. Shafique (1974), "A Forecast of Wheat Production in the Punjab for 1973-74," *Pakistan Development Review,* 13(1): 106-112.

Box, G. E. P. and G. M. Jenkins (1976), *Time Series Analysis, Forecasting and Control,* Holden-Day, San Francisco.

Chaudhary, Ghaffar M. and A. R. Kemal (1974), "Wheat Production under Alternative Production Functions," *Pakistan Development Review,* 13(2): 407-415.

Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimation for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association,* 74: 427-431.

Dickey, D. A. and W. A. Fuller (1981), "Likelihood Ratio Statistics For Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association,* 49: 1057-1072.

Enders, W. (2004), *Applied Econometric Time Series (Second Edition),* John Wiley & Sons (Noida, U.P. India).

Government of Pakistan (Various Issues), *Agricultural Statistics of Pakistan,* Ministry of Food, Agriculture and Livestock, Islamabad.

Government of Pakistan (Various Issues), *Census of Agricultural Machinery,* Agricultural Census Organization, Federal Bureau of Statistics, Statistics Division, Islamabad.

Government of Pakistan (Various Issues), *Census of Agriculture,* Agricultural Census Organization, Federal Bureau of Statistics, Statistics Division, Islamabad.

Government of Pakistan (2005,2006), *Crops Area Production (By Districts),* Ministry of Food, Agriculture and Livestock, Islamabad.

Government of Pakistan (Various Issues), *Economic Survey,* Ministry of Finance, Islamabad.

Government of Pakistan (Various Issues), *Fertilizer Use Survey*, National Fertilizer Development Centre, Planning & Development Division, Islamabad.

Government of Pakistan (Various Issues), *Fifth Five-Year Plan*, Planning and Development Division, Islamabad.

Government of Pakistan (2005), *Household Integrated Economic Survey*, Federal Bureau of Statistics, Islamabad.

Government of Pakistan (Various Issues), *Pakistan Statistical Year Book*, Federal Bureau of Statistics, Statistics Division, Islamabad.

Government of Pakistan (1988), *Sixth Five-Year Plan*, Planning & Development Division, Islamabad.

Green, W. H. (2003). *Econometric Analysis (Fifth Edition)*, Pearson Education: Delhi, India.

Griffiths, W., Thomson, G., and T. Coelli (1999). "Predicting Output from Seemingly Unrelated Area and Yield Equations." Working Paper Series in Agricultural and Resource Economics (99-9). Graduate School of Agricultural and Resources Economics, University of New England.

Gujrati, D. N. (2003), *Basic Econometrics* (Fourth Edition), McGraw-Hill Higher Education: New Delhi, India.

Iqbal, N. B. K., Maqbool Asif and Shohab Abid Ahmad (2005), "Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan," *Journal of Agriculture and Social Sciences,* 1(2): 120-122.

Khan, N. Z., M. Ahmad and A. Rasheed (2003), "Wheat Production in Pakistan: Saga of Policy Disincentive," *Pakistan Development Review*, 42(1): 1-27.

Mukhtar, M. E. and H. Mukhtar (1988), "Input Use and Productivity across Farm Sizes: A Comparison of the Two Punjabs," *Pakistan Development Review*, 27(4): 595-604.

Qureshi, S. K. (1974). "Rainfall, Acreage, and Wheat Production in West Pakistan: A Statistical Analysis," *Pakistan Development Review*, 13(3): 566-593.

Salam, A. (1981), "Farm Tractorization, Fertilizer Use and Productivity of Mexican Wheat in Pakistan," *Pakistan Development Review*, 20(3): 323-345.

Saleem, M. A. (1989), "Farmers' Attitude Towards Risk in Dryland Wheat Production Areas in Jordan," University of Jordan, *Dirasat-Agricultural Science* 16(7): 28-52.

World Bank, (2006), *World Development Indicators 2006*, Washington, DC.