Pak. J. Agri. Sci., Vol. 41(1-2), 2004

FORECAST MODELS FOR SUGARCANE IN PAKISTAN

M. Asif Masoood* and Malik Anver Javed
Biometrics Programme, Social Sciences Institute,
National Agricultural Research Centre, Park Road, Islamabad.

In this paper five separate models were developed for Pakistan, Punjab, Sindh and NWFP for Sugarcane area forecasting. One yield model was developed for Pakistan. Production was estimated by multiplying area and yield forecasts. The explanatory variables included in the area models were lagged area under sugarcane lagged by one year and sugarcane procurement price. The explanatory variables used in the yield models were fertilizer consumption of sugarcane (kgs/ha), total water availability at farm gate during the crop growth period. The analysis revealed that highly significant results for procurement price, lagged sugarcane area and water availability at farm gate during the crop growth period. Sugarcane area and yield forecast models are very efficient and can be used to predict future area and yield estimates with reasonable level of accuracy a couple of months before sowing and harvesting of sugarcane crop. By that time, secondary source data on all the used explanatory variables become available.

INTRODUCTION

The reliable estimate of total crop production well in advance before harvest is extremely important for determining the prices, export-import policies as also enabling the Government to take remedial measures on time in case of seeing any problems of scarcity or surplus in the national crop production. This is especially important for cash crops as the whole farm economy is at stake. Sugarcane is an important cash crop of Pakistan, which is mainly grown for sugar and *gur* production. Apart from being the major source of sugar used in human diet, recent use of fuel and fiber has been advocated (Deepchand, 1986) using its bye products. Sugarcane crop is an extremely water intensive and yet an important cash crop. Sugar production in the country mostly depends on this crop, though a small quantity of sugar is also produced from sugarbeet. Its shares in value added in agriculture and GDP are 6.2 percent and 1.5 percent respectively. Sugarcane was cultivated on an area of 1086 thousand hectares (Economic Survey 2002-03), showing an increase of 8.6 percent over the last year. The size of the sugarcane crop is provisionally estimated at 52049 thousand tons which is higher by 8.3 percent, as compared with last year. Sugarcane is mainly grown in Punjab, Sindh and N.W.F.P. The Punjab alone accounted for about 65.7 percent of the total area under sugarcane and produced 66.2 percent of the total sugarcane in the country during 2001-2002.

Many economists have accomplished substantial work on forecasting of area and yield of major crops. A number of forecasting models for projecting the crop have been formulated earlier. A few of them are by Azhar et al (1973), Malik (1983), Amir & Akhtar (1984), Noor Muhammad (2001) for wheat crop, Khan and Khan (1988) for rice crop and Naseer et al (1987) for sugarcane crop. The most important criterion in forecasting agricultural production is the theoretical consistency of the model (Gujrati 1978, Granger, 1980).

The present study focused on the developing of sugarcane area models for three major sugarcane producing provinces namely Punjab, Sindh, N.W.F.P. and Pakistan and yield model for Pakistan. Area multiplied by the yield then gave sugarcane production forecast model for Pakistan.

MATERIAL AND METHODS

The study was based on time series data for the period from 1975-76 to 2001-2002. The data on the response and explanatory variables were collected from various secondary sources to formulate the models. The main purpose of using secondary data was to save time and getting quick information about forecasting. The regression (Ordinary Least Squares) technique was employed for sugarcane area and yield forecast models. The method of the least squares has the two-fold merit of minimizing not only the sum of squares of deviations but also the variance of the estimators b_js. An econometric model is a system of one or more equations that describe the relationship among several economic and time series variables. Econometric models are probabilistic models and capitalize on the probabilistic relationship that exists between a dependent variable representing the time series and any of a number of independent variables. Economic models use auxiliary variables as predictors. Different explanatory variables were explored before settling down for the ones which could give best result in terms of economic logic and satisfy certain statistical criteria.

Two separate models were developed for forecasting of sugarcane area and yield, while production was estimated simply by multiplying area with yield forecasts. Four equations were estimated for sugarcane area for Pakistan and three provinces and one for sugarcane yield for Pakistan. Table 1&2 provide the description of the area and yield models used in this study.

Two criteria were used to evaluate the forecast performance of each of the above mentioned models. The first criterion used could measure actual forecast error but for the previous years for which data are available. The second criterion provided measurements of the model fit to the data available at a particular period of time. A brief description of the two criteria is as follows.

iv. The Durbin-Watson statistic as a test for serial correlation.
v. The size of regression coefficients, their significance and standard errors.
vi. The vector of residuals for the time series used in formulating the sugarcane area and yield forecast models.
vii. Study of the correlation matrices for the examination of multi-collinearity.

RESULTS AND DISCUSSION

The results of area and yield forecast models were studied considering their statistical properties and economic logic. The statistical properties are the coefficient of determination (R^2), statistical significance

Table 1. Specification of variables for sugarcane area forecast models

Province / Pakistan	Dependent variable	Explanatory Variables	
		1	2
Punjab	A_t	A_{t-1}	PRP
Sindh	A_t	A_{t-1}	PRP
NWFP	A_t	A_{t-1}	PRP
Pakistan	A_t	A_{t-1}	PRP

A_t = Area under sugarcane ('000' hectares) in the year t.
A_{t-1} = Area under sugarcane ('000' hectares) in the year t-1.
PRP = Procurement price of sugarcane (Rs. Per 40 kgs) in the year t.

Table 2. Specification of variables for sugarcane yield forecast model

	Dependent Variable	Explanatory Variables	
		1	2
Pakistan	Y_t	FCS	WA

Y_t = Yield (tons per hectare) of sugarcane in the year t.
FCS = Fertilizer consumption on sugarcane area (kilogram per hectare in the year t
WA = Water available at farm gate (MAF) in the year t

Type–I Criterion

The type-1 criterion to measure the forecast is the "Absolute percent deviation of the forecast from the estimate at maturity given by $|(P-A)/A| \times 100$, where P is the predicted or forecasted value for the year t, and A is the actual value for the year t.

Type–II Criterion

This criterion was used to provide a measure of model fit and its performance based on the following:

i. The absence of model assumption violation specifically the heteroscedasticity.
ii. The size of the residual mean square (RMS) from the multiple regression.
iii. The size of coefficient of determination (R^2).

of coefficients, Durban-Watson and F- Statistic. For a statistical model it was considered that economic logic prevails which means that the signs of coefficients are logical and consistent.

Sugarcane Area Forecast Model

Sugarcane area fluctuations are caused by a number of factors. Some of the most important factors are prices of sugarcane, gur, competing crops, availability of irrigation, rainfall and fertilizers etc. An explanatory variable of area under sugarcane was used on lag basis because farmers make crop acreage allocation decision on the basis of past experience. A lagged variable is an observation from a variable recorded from an earlier time period, such as A_{t-1} = Area under sugarcane in the year t-1. Therefore, the use of lagged

variable seemed necessary. Before fitting the models, it is very important to carefully identify these variables which will be used in the model. Keeping this viewpoint many explanatory variables were identified (i.e. Pakistan and province-wise area and yield of sugarcane, procurement price, prices of competing crops, gur prices, water availability, fertilizer consumption and rainfall). Area was considered as dependent variable in sugarcane area forecast model. For the selection of model variables, each independent variable was plotted against dependent variable and their behavior was studied. The variables which had linear relationship were kept while others were dropped. The multicollinearity problem in the explanatory variables was studied with the help of correlation matrix. After the statistical exercises and careful consideration of logical relationship of the variables, a set of variables was selected. The variables included in the sugarcane area models were lagged area under sugarcane lagged by one year and sugarcane procurement price. It is rational to consider procurement price of sugarcane crop as the major incentive which motivates the farmers to increase area under sugarcane. Timely announcement of procurement price by the government enhances the confidence of the farmer for the prospective profit from the crop and motivate him to allocate more or less area under a crop from his total area resources.

The sugarcane area model for Pakistan showed highly significant results for lag area and significance of price coefficient in the model shows that price can influence the decision of the farmer for allocation of area for sugarcane. It may be mentioned that explanatory variables like gur deflated prices, rainfall in sowing season and water availability during sowing season were also tried as explanatory variables which were found to be non-significant and could not improve the overall performance of the model. The model is satisfactory with statistical properties i.e. high R^2, low value of Durbin-Watson Statistic and high F-Statistic (table 3).

Sugarcane area model for the Punjab showed significant results for procurement price whereas lag area was found highly significant. The measure of R^2 did not provide very high value. However, a highly significant F-Statistic, Durbin-Watson Statistic within the range of tolerance and logically acceptable signs of the regression coefficients indicated that the performance of the model is going to be satisfactory.

Table 3. Sugarcane area forecast model for Pakistan and provinces

Provinces/Pakistan	Constant	Coefficients		R^2	D.W.	F. Ratio
		A_{t-1}	PRP			
Punjab	276.83	0.47** (2.87)	2.28* (2.04)	0.54	1.60	13.83**
Sindh	19.79	0.94** (12.3)	0.14 (0.36)	0.93	2.00	161.20**
NWFP	49.61	0.47* (2.38)	0.16 (1.54)	0.56	1.65	15.08**
Pakistan	419.83	0.46** (2.89)	4.23* (2.59)	0.74	1.55	33.68**

Figures in parenthesis are "t" values, * Significant at 5 percent level. ** Highly significant

Sugarcane area model for the Sindh showed significant results for lag area at one percent level of significance. The procurement price variable did not show significant results (a regression coefficient of (-) 0.14 with a t-value of 0.36 was observed), however a highly significant lag area coefficient support that the farmers were more inclined to make decisions on the basis of what they observed in the previous year. The model gave high R^2 value, acceptable size of Durbin Watson Statistic and highly significant F-Statistic.

Sugarcane area model for NWFP province indicated significant results for lag area while non-significant for procurement price. The model has a medium R^2 and high F-Statistic. The size of Durbin-Watson Statistic is not problematic. The results presented in table 3 satisfy the properties of the selected area forecast models for type-II criterion.

Yield Forecast Model for Pakistan

Yield of sugarcane is largely dependant upon a number of factors, like fertilizer consumption, water availability during the growing season and rainfall etc. The use of fertilizer on sugarcane has strong influence on the sugarcane yield. If the farmer uses fertilizer on sugarcane as advised by the recommended package of crop production technology it is going to result in better yields. Therefore, the fertilizer consumption in kgs per hectare of sugarcane crop has been incorporated to see the clear impact of fertilizer on sugarcane crop (Pakistan Fertilizer Statistics, 1982).

Table 4. Sugarcane yield forecast model for Pakistan

	Constant	Coefficients		R^2	D.W.	F.Ratio
		FCS	WA			
Pakistan	21.79	0.023** (1.80)	0.13* (2.34)	0.73	1.24	33.07**

Figures in parenthesis are t values, * Significant at 5 percent level, ** Significant at 10 percent level

For yield forecast model for Pakistan, different combinations of these variables namely, fertilizer consumption, water availability, rainfall, yield risk (measured by the standard deviation of the yield of three preceding years) were used with a view to select the best subset. Yield risk and rainfall were found to be non-significant and could not show better performance for the yield model and were thus excluded from the final model. The variables used in the yield model were sugarcane yield as dependent variable whereas fertilizer consumption per kgs/ha and water availability at farm gate during the season as explanatory variables.

The National sugarcane area forecast for 2002-03 was 5.09 percent higher than the official estimates. The forecasted sugarcane area of Pakistan for 2002-03 and 2003-04 were 5.09 and 4.15 percent higher than the official estimates. The results revealed that the percent change for sugarcane area models for National as well as Provincial level was found within ±5 percent range.

The performance of sugarcane yield forecast model is given in table 6. When compared with official estimates of 45.40 tons per hectare yield in 2000-01; forecasted yield was only 1.23 percent higher than the official estimate and 3.47 percent lower than the official estimate of 2001-02 whereas 1.94 percent lower than official estimate of 2002-03.

Table 5. Performance of sugarcane area forecast model

Province / Pakistan	2001-02			2002-03			2003-04		
	Official Est.	Forecast Est.	Percent Change (±)	Official Est.	Forecast Est.	Percent Change (±)	Official Est.	Forecast Est.	Percent Change (±)
	('000' hectares)								
Punjab	656.8	674.8	(+) 2.74	650.0	661.8	(+) 1.81	650.0	681.2	(+) 4.80
Sindh	240.7	231.5	(-) 3.82	241.0	238.2	(-)1.16	250.0	240.0	(-) 0.40
NWFP	101.5	105.1	(+) 3.54	100.0	104.9	(+) 4.90	100.0	102.9	(+) 2.90
Pakistan	999.7	1034.5	(+) 3.48	991.0	1041.5	(+) 5.09	1000.0	1041.5	(+) 4.15

Yield model for Pakistan gave significant results at 5 percent level for water availability. Fertilizer consumption for sugarcane (FCS), however, was significant only at 10 percent level of significance with a logical sign. A high R^2 as well as a highly significant F-Statistic showed significant joint contribution of explanatory variables and indicating a good explanatory power of the model. Durbin-Watson Statistic was a little lower than we would have expected.

The performance of sugarcane area forecast model was tested by comparing official estimates with the forecasts for a period of three years from 2000-02 to 2003-04 and is given in table 5.

Forecasted area for 2002-03 was only 1.81 and 4.90 percent higher than the official estimates of Punjab and NWFP respectively. In 2003-04, forecast was 4.80 and 2.90 percent higher for the Punjab and NWFP respectively whereas 1.16 and 0.40 percent was lesser for Sindh in 2002-03 and 2003-04 respectively.

Table 6 Performance of sugarcane yield forecast model for Pakistan

Year	Official Estimate	Forecast Estimate	Percent Change (±)
	(tons/ha)		
2000-01	45.40	45.96	1.23
2001-02	48.10	46.43	-3.47
2002-03	47.34	46.42	-1.94

Production of sugarcane was also forecasted by multiplying forecasted area with the forecasted yield of respective years. The performance was tested by comparing the forecasted production with the official estimates of the corresponding years. Comparing with the official estimate for the year 2001-02, the forecasted estimate is only 0.02 percent higher, while comparing the results of 2002-03, it is noticed that forecasted production is 7.12 percent lower than the official estimate. It indicates that the results of

forecasted production are satisfactory. These results reflect that the performance of area and yield forecast model is acceptable. Sugarcane area and yield forecast models are very efficient and can be used to predict future area and yield estimates with reasonable level of accuracy a couple of months before sowing and harvesting of sugarcane crop. By that time, secondary source data on all the used explanatory variables become available.

Table 7. Comparison between official and predicted estimates of sugarcane production in Pakistan, 2001-02 to 2002-03

Year	Official Estimate	Forecast Estimate	Percent Change (±)
	('000' tons)		
2001-02	48041.6	48031.8	(+) 0.02
2002-03	52055.8	48349.7	(-)7.12

Appendix–I

Comparison between official and predicted estimates of sugarcane area in Pakistan from 1975-76 to 2003-04

Year	Area (000 hectares)		Percentage Variation
	Official	Predicted	
1975-76	699.80	752.45	7.52
1976-77	787.80	764.79	-2.92
1977-78	822.50	804.90	-2.14
1978-79	752.50	820.72	9.07
1979-80	718.50	794.49	10.58
1980-81	824.70	788.10	-4.44
1981-82	946.70	836.51	-11.64
1982-83	911.70	892.13	-2.15
1983-84	896.50	876.18	-2.27
1984-85	903.60	869.25	-3.80
1985-86	779.80	872.48	11.89
1986-87	762.00	825.11	8.28
1987-88	841.60	817.00	-2.92
1988-89	876.90	856.85	-2.29
1989-90	854.30	877.86	2.76
1990-91	883.80	874.63	-1.04
1991-92	896.10	895.03	-0.12
1992-93	884.60	899.58	1.69
1993-94	962.80	899.93	-6.53
1994-95	1009.00	946.17	-6.23
1995-96	963.10	971.47	0.87
1996-97	964.50	961.65	-0.30
1997-98	1056.20	1009.96	-4.38
1998-99	1155.10	1051.76	-8.95
1999-00	1009.80	1096.85	8.62
2000-01	960.80	1030.61	7.27
2001-02	999.70	1034.5	+3.48
2002-03	991.00	1041.50	5.09
2003-04	1000.00	1041.50	4.15

Appendix – II

Comparison between official and predicted estimates of sugarcane yield in Pakistan from 1975-76 to 2002-03

Year	Yield (Tons/ha)		Percent Variation
	Official	Predicted	
1975-76	36.50	35.17	-3.64
1976-77	37.50	35.15	-6.27
1977-78	36.60	35.92	-1.86
1978-79	36.30	36.92	1.71
1979-80	38.30	37.90	-1.04
1980-81	39.20	37.18	-5.15
1981-82	38.60	37.14	-3.78
1982-83	35.70	38.35	7.42
1983-84	38.20	40.50	6.02
1984-85	35.60	38.24	7.42
1985-86	35.70	39.40	10.36
1986-87	39.60	40.95	3.41
1987-88	39.20	40.60	3.57
1988-89	42.20	42.20	0.00
1989-90	41.50	43.12	3.90
1990-91	40.70	43.27	6.31
1991-92	43.40	43.49	0.21
1992-93	43.00	44.68	3.91
1993-94	46.10	44.61	-3.23
1994-95	46.70	44.66	-4.37
1995-96	47.00	45.97	-2.19
1996-97	43.50	44.16	1.52
1997-98	50.30	42.85	-14.81
1998-99	47.80	45.21	-5.42
1999-00	45.90	46.53	1.37
2000-01	45.40	45.96	1.23
2001-02	48.10	46.43	-3.47
2002-03	47.34	46.42	-1.94

REFERENCES

Amir, P. and M.R. Akhtar. 1984. "A model for forecasting wheat production in the four provinces of Pakistan". Social Sciences Division, PARC.

Azhar, B.A., G.M. Chaudhry and M.A. Shafique. 1973. A model for forecasting wheat production in the Punjab" Q.J. PIDE XIV(1), PP 407-414.

Deepchand, K. 1986. Economics of electricity production from sugarcane tops and leaves – a preliminary study. Ind. Sugar J. 88: 210-216.

Naseer, A., M. Khan, Inayat Khan & Qazi Tauqir Azam. 1987. "Factors responsible for sugarcane area and yield fluctuations in Pakistan" Pakistan J. Agric. Res. Vol.8, No.2, 1987, 150-154.

Government of Pakistan, Agricultural Statistics of Pakistan', Ministry of Food, Agriculture and Livestock, Food and Agriculture Division, Planning Unit, Islamabad, 1999-00.

Government of Pakistan, Pakistan Economics Survey. Finance Division Economic Advisor Wing, Islamabad (2002).

Granger, C.W.J. "Forecasting in Business and Economics" Academic Press, Inc. 1980.

Gujrati, Damodar. 1978. "Basic Economics" McGraw Hill, Inc.

Khan, M.I. & N.A. Khan. 1988. "Rice forecast model of Pakistan" Gomal Univ. J. Res. Vol. 8, No. 1 & 2, 1988, 95-105.

Malik, M.A. 1983. "Supply and demand production of wheat in Pakistan" University of Agriculture, Faisalabad.

National Fertilizer Development Centre. (1982). "Pakistan Fertilizer Statistics". National Fertilizer Development Centre, Planning and Development Division, Government of Pakistan, Islamabad. Statistical Bulletin No. 2.

Noor Muhammad. 2001. "Factors affecting area and yield of wheat crop- A case study" Pak. J. of Agricultural Economics. Vol. 4, No. 2, 2001, 66-76. Agricultural prices commission of Pakistan, MINFAL, GOP, Islamabad.