Agricultural Sciences, 6(1-2):53-58 (2001)

Comparison of Agricultural Forecasts with Actual Data

Y.A. Alseleem*

Department of Agricultural Economics and Extension
College of Agriculture, King Saud University, Qassim, Saudi Arabia

مقارنة التوقعات الزراعية بالإحصاءات الفعلية

يوسف السليم

خلاصة: حتى الآن هناك القليل الذي كتب عن دقة التوقعات الموجودة لدى وزارة الزراعة والمياه. إن مقاييس الدقة تعتبر مهمة جدا مقارنة بطرق المعاينة والتحليل المختلفة، كما أن مقارنة التوقعات والبيانات الفعلية تزودنا بمقياس الدقة الذى تحصلنا عليه في السابق. في الحقيقة، أن التقييم الحالي لدقة توقعات المحصول ظهرت أهميته منذ قيام العملاء الحكوميين، وشركات الأعمال والزراع من خلال التعامل بملايين الريالات سنويا على أساس التوقعات، كما أن النقص في التوقعات من الممكن أن يتسبب في بعض الآثار الغير مرغوبة من ناحية الخطط المستقبلية وتوزيع الموارد. اختبر هذا البحث توقعات ٢٥٥ بيانا من إحصاءات وزارة الزراعة والمياه في المملكة العربية السعودية لكل من محصول القمح والشعير والطماطم والبطيخ والتمور والعنب والدجاج والأغنام والإبل، وذلك خلال الفترة من ١٤٠٠ - ١٤١٦ هـ (١٩٧٩-١٩٩٥م) وقد اختبرت الدراسة الفرق بين التقدير المتوقع والتقدير الفعلي. إن نتائج هذه الدراسة قد زودت وزارة الزراعة والمياه بالمعلومات المفيدة في كيفية اتخاذ القرارات المطلوبة تجاه طرق توقعات المحصول (الحيوان) لمقابلة احتياجات المستهلكين.

ABSTRACT: So far, little has been said regarding the accuracy of forecasting given by the Ministry of Agriculture and Water (MAW), Saudi Arabia. Measures of accuracy are quite useful in comparing several methods of sampling or analysis. A comparison of forecasts with actual data gives us a measure of accuracy. In fact, a current evaluation of the accuracy of crop forecasts appears useful since government agencies, agribusiness firms, and farmers make decisions involving millions of riyals annually on the basis of the forecast, and deficiencies in the forecasts may cause undesirable effects on plans and resource allocation. The present research examines the accuracy of 255 MAW crop area and production forecasts for wheat, barley, tomato, watermelons, palm dates, grapes, chicken, sheep, and camel for the period 1400-1416 H (i.e. 1979-1995G). The study tested the difference between actual and forecast estimates. The results of this study provide useful information about decision making in crop (animal) forecasting procedures to meet users requirements.

Keywords: agricultural forecasts, Saudi Arabia, agribusiness, crop production, decision-making.

The need for forecasting comes from the presence of uncertain phenomena or lack of knowledge of what is going to happen in the future. It is clear that if information in the future is adequate and certain, there would be no need for forecasting. When information is imperfect, forecasting should be done where the future may be represented by certain events that happened in the past.

Both agricultural data malfunction and non-authentication are important agricultural sector problems. No doubt, if the statistical data is insufficient or imperfect, agricultural plans, goals and ambitions, will be elusive, and the forecasts will be far from reality. Hence, economic goals and results will not be met. Then we can say that lack of accuracy of primary and secondary data on which the plans depend will lead to the total failure of agricultural objectives. Accordingly, it is clear that the importance of agricultural data collection on which the remaining stages of the statistical analysis depend, is crucial. Results will be of no value unless based on correct and accurate data. Also, in the absence of reliable data, the research recommendations will not be accurate.

This paper aims to evaluate the accuracy of agricultural statistical forecasts of the Ministry of Agriculture and Water (MAW) in the Kingdom of Saudi Arabia through testing the significant differences between the actual data and the forecast data for area, production of basic crops, and the number of animals in the Kingdom.

*Corresponding author.

Methodology

The data used was taken from the Ministry of Agriculture and Water, Statistical Indicator and Agricultural Statistical Yearbook, published by Economic and Statistical Studies Department covering the period 1400-1416 H (i.e. 1979-1995G). The simple linear regression equation was used:

$$A_{it-1} = F_{it-1} + e_{it}$$

where: A_{it}, the actual crop statistics or the animal in the year t; F_{it-1}, the primary statistics carried by the Ministry of Agriculture and Water in the year t_{-1} forecasts to the year data to the crop or animal i; e_i, estimation errors; α and β, the estimated function parameters.

Tomek and Robinson (1981) said that the results of the previous estimated equation would be $R^2=1$, $\alpha=0$, $\beta=1$. Forecast errors can be measured by estimating variances from zero, β from one, and R^2 from 1 (Tomek and Robinson, 1981).

The following research assumption were tested:

H_0: $\alpha = 0$
H_a: $\alpha \neq 0$
H_0: $\beta = 1$
H_a: $\beta \neq 1$

By using (t- test) as follows:

$$t_c = \frac{a - 0}{\sqrt{\text{var}(a)}}$$

$$t_c = \frac{B - 1}{\sqrt{\text{var}(B)}}$$

In order to test the results by using another procedure, the following research assumptions were also tested:

$H_0 = F_{t-1} = A_t$
$H_a = F_{t-1} \neq A_t$

The value of t_c is determined by:

$$t_c = \frac{F_{t-1} - At}{\sqrt{\text{var}(F_{t-1})}}$$

If $t_c > t_{n-2}$, we reject the null hypothesis H_0, which says, there is no difference between the value of F_{t-1} of the dependent variable and the actual value A_t to that variable within certain limited period of time. Accordingly, we can say that the forecasting model is weak.

If $t_c < t_{n-2}$, we accept the null hypothesis, which says that there is no variance between the expected value and the dependent variable in the same period of time. Hence, we can say that the forecasting capability for the model is good and can be recommended (Awad, 1998).

This paper used the following models to estimate the relation between the primary data F_{t-1} and the actual data (A_t)

$F_{it-1} = \alpha + \beta A_{it} + e_{it}$. The first model.

$[F_{it-1} - A_{it}] = D_{it} = \alpha + \beta\,\text{Time} + e_{it}$. The second model.

$F_{it-1} = \alpha + \beta A_{it} + c\,\text{Time} + e_{it}$. The third model.

$F_{it-1} = \alpha + \beta A_{it-1} + e_{it}$. The fourth model.

where: D_{it} is the difference between the primary forecasts F_{it-1} and the actual data A_{it} which was found in the year t; Time denotes the year of the study; A_{it-1} is the actual data in the year t_{+1}.

The first model indicates the primary forecasts function estimate F_{it-1} valuated by actual data A_{it}. This model is the same as that of Tomek and Robinson (1981), but the dependent variable is now F_{it-1} and the independent variable is A_{it}, which is vice versa to what was mentioned by Tomek and Robinson. In the authors'opinion, this is due to the dependence of the agricultural data forecasters on previous actual data.

The second model is an equation giving a general trend about forecasting errors (D_{it}) and its relation with time. In this model, we test if $\alpha=0$, $\beta=0$ to know whether the forecasting errors are correlated with the time or not.

In the third model we incorporate the variables of both the first and the second model together because the author presumes that forecasting errors will be lessened with the passage of time because of user experience. Accordingly, we have to add the time variable to the first model to measure its effect. It should be mentioned here that the ministry forecasts are based on the general trend equation and the total production accountancy by multiplication of the average production by total area.

The fourth model is similar to the first model, except that the actual data is lagged behind by one year. This is because forecastors may depend on their estimates concerning the actual data based on the last two years and not only on the last year.

Review of Literature

In 1972, Gunnelson published a paper on analysis of forecast accuracy at USDA for barley, sorghum, oats, potatoes, soya beans and wheat in the period 1929-1970. This paper indicated the development of accuracy of agricultural forecasts for USDA with the passage of time. It also showed the lack of bias or systematic errors.

Another study by Gorham (1978), about government and private agriculture data sources, indicated that government sources have great effect upon agricultural markets compared with private sources. The study also showed that development of government agricultural data accuracy with time was comparable with private data. This may be true for two reasons. The first is that government sources deal with no profit in their transactions. Therefore, they develop their personal career for data collection and forecasting, and they increase the sample size. The second reason is that the private sector is confined to a more limited budget. The results of this paper indicate that the forecasts of USDA are getting better compared to the private sector outspread forecasts.

A study carried out by Crowder (1972) pointed to the importance of modernization of used forecasts in the presence of new data. He also indicated that forecast accuracy depends on commodity importance, and on the nature and importance of decisions taken about forecast results.

Paulino (1980) in his study compared FAO and USDA outspread data for production, area and foreign trade for important crops in different areas of the world. He found that there are some differences in the outspread data. This study recommends that the FAO and USDA should work as a group to lessen the variances in the outspread data. The study also recommends that more attention should be drawn towards developing country' s agricultural data, which creates a major share of the differences in the outspread data.

Smith (1992), studied forecast bias and accuracy for eight models used in estimation of population forecasts. He used population data for some states in USA in different time series data. He concluded that there is no evidence that the complicated mathematical models are more accurate than the simpler ones.

Wilkie and Pollock (1996), studied the quality judgement upon the accuracy of exchange rate forecasts where they used models which depend on time series for exchange rate. In their study, forecast errors were calculated and used for judgement of forecast quality.

Madura *et al.* (1998) evaluated bias and accuracy in three models for exchange rate forecasts. They employed 12 exchange rates. The study showed that every model indicates a bias in forecast and differs from the other model.

Results

Comparing the results of the estimated four models[1] for both area and production of crops, it is clear that the third model is the best (Tables 1-8).

This model estimates the forecasts closer to accuracy than the other models. But we must indicate that the third model will be illusive in the case of wheat and barley area estimation, and wheat, barley, dates and grapes production (Table 9).

The results of null hypothesis test, indicates ($\alpha = 0$, $\beta = 1$) our acceptance of this hypothesis at the level of 5% for the first, the third and the fourth models, except the following:

In the fourth model for area and for production of watermelon, dates area (Tables 1,3,4,5,7,8), forecasts of data are not accurate (not ideal forecasts), where the estimated and actual data for the previous period are not identical. This occurs in the first model for the production of barley, tomato and watermelon, in third model for the area of wheat and barley, for production of wheat, barley, dates, and grapes.

The results of test hypothesis for the second model indicate ($\alpha=0$, $\beta=0$) that there is no relation between the forecast errors (forecast variance from the actual) and the time, except for the production of wheat and barley forecasts (Table 6). We should note that the test of hypothesis for the second model will be set according to the search for the relationship between the forecast

TABLE 1

Regression analysis results and the test of hypothesis for the first model for the area of important agricultural crops in the Kingdom.

Area	α	β	R^2	Test of Hypothesis Results
Wheat	78.17 (202.26)	0.99 (0.03)	0.98	$\alpha = 0$ $\beta = 1$
Barley	260.99 (499.27)	0.93 (0.54)	0.25	$\alpha = 0$ $\beta = 1$
Tomato	406.19 (363.92)	-0.75 (1.51)	0.03	$\alpha = 0$ $\beta = 1$
Watermelon	36.49 (49.81)	0.79 (0.24)	0.55	$\alpha = 0$ $\beta = 1$
Dates	-0.71 (57.05)	0.99 (0.07)	0.94	$\alpha = 0$ $\beta = 1$
Grapes	1.99 (10.52)	0.97 (0.15	0.83	$\alpha = 0$ $\beta = 1$

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

[1]Comparison was done according to the Coefficient of Determination R^2 .

TABLE 2

Regression analysis results and the test of hypothesis for the second model for the area of important agricultural crops in the Kingdom.

Area	α	β	R^2	Test of Hypothesis Results
Wheat	-323.2 (25844.1)	22.95 (18.35)	0.09	α = 0 β = 0
Barley	-172261 (124273.34)	122.32 (88.14)	0.18	α = 0 β = 0
Tomato	-10042 (9570.84)	7.12 (6.79)	0.14	α = 0 β = 0
Watermelon	1210.00 (2658.78)	-0.86 (1.89)	0.02	α = 0 β = 0
Dates	-1385.99 (3815.89)	0.98 (2.70)	0.01	α = 0 β = 0
Grapes	-341.50 (1024.59)	0.24 (0.73)	0.01	α = 0 β = 0

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

errors and the time (α=0, β=0) where the test of hypothesis for the rest of models should be done to know whether the forecast is ideal or not (α=0, β=1).

The estimation (not calculation) of the actual data from the general trend equation and an estimation of forecasts will lead to a symmetry between the actual and the forecasted data. This gives an impression that the data estimate was accurate, and this is what was obtained in this study. In general, the actual data should be calculated from the real figures every year. If this is costly, then it should be done according to short periods. The forecasts of the coming years can

TABLE 3

Regression analysis results and the test of hypothesis for the third model for the area of important agricultural crops in the Kingdom.

Area	α	β	C	R^2	Test of Hypothesis Results
Wheat	-53526 (30093.66)	38.21 (21.45)	0.95 (0.04)	0.98	α ≠ 0 β = 1
Barley	-441612 (159143.83)	314.07 (113.12)	-0.39 (0.63)	0.62	α ≠ 0 β ≠ 1
Tomato	-12979 (8865.76)	9.59 (6.35)	-1.28 (1.44)	0.30	α = 0 β = 1
Water-melon	-3253.53 (5394.28)	2.37 (3.89)	0.52 (0.51)	0.57	α = 0 β = 1
Dates	-20866 (12224.29)	14.99 (8.79)	0.61 (0.24)	0.96	α = 0 β = 1
Grapes	-2950.04 (2404.81)	2.11 (1.72)	0.57 (0.36)	0.86	α = 0 β = 1

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

TABLE 4

Regression analysis results and the test of hypothesis for the fourth model for the area of important agricultural crops in the Kingdom.

Area	α	β	R^2	Test of Hypothesis Results
Wheat	-120.44 (775.30)	0.98 (0.13)	0.80	α = 0 β = 1
Barley	447.96 (477.51)	0.35 (0.49)	0.06	α = 0 β = 1
Tomato	435.93 (400.65)	-0.89 (1.68)	0.05	α = 0 β = 1
Watermelon	115.98 (68.70)	0.38 (0.32)	0.15	α = 0 β ≠ 1
Dates	71.93 (51.47)	0.85 (0.06)	0.95	α = 0 β ≠ 1
Grapes	17.56 (16.11)	0.68 (0.23)	0.55	α = 0 β = 1

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

TABLE 5

Regression analysis results and the test of hypothesis for the first model for the production of important agricultural crops and number of animals in the Kingdom.

Production	α	β	R^2	Test of Hypothesis Results
Wheat	124.68 (171.30)	0.98 (0.07)	0.94	α = 0 β = 1
Barley	-36.62 (255.85)	1.50 (0.04)	0.64	α = 0 β ≠ 1
Tomato	-84.75 (202.75)	1.20 (0.05)	0.39	α = 0 β ≠ 1
Watermelon	-570.03 (241.09)	-0.41	0.05	α ≠ 0 β ≠ 1
Dates	36.88 (67.50)	0.93 (0.13)	0.86	α = 0 β = 1
Grapes	-1.34 (16.29)	1.03 (0.16)	0.83	α = 0 β = 1
Sheeps	455.91 (726.20)	0.92 (0.11)	0.87	α = 0 β = 1
Camel	188.22 (232.36)	0.52 (0.58)	0.08	α = 0 β = 1
Poultry	-6137.97 (28644.51)	0.92 (0.17)	0.77	α = 0 β = 1

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

TABLE 6

Regression analysis results and the test of hypothesis for the second model for the production of important agricultural crops and number of animals in the Kingdom.

Production	α	β	R^2	Test of Hypothesis Results
Wheat	-49666 (22929.05)	35.36 (16.29)	0.27	$\alpha \neq 0$ $\beta \neq 0$
Barley	-162166 (52006.43)	115.112 (36.87)	0.55	$\alpha \neq 0$ $\beta \neq 0$
Tomato	-11091 (8569.77)	7.86 (6.08)	0.16	$\alpha = 0$ $\beta = 0$
Watermelon	1041.53 (8267.77)	-0.73 (5.87)	0.0019	$\alpha = 0$ $\beta = 0$
Dates	-965.39 (2511.95)	0.68 (1.78)	0.02	$\alpha = 0$ $\beta = 0$
Grapes	-820.58 (1042.50)	0.58 (0.74)	0.06	$\alpha = 0$ $\beta = 0$
Sheeps	32918 (41318.96)	-23.37 (29.29)	0.06	$\alpha = 0$ $\beta = 0$
Camel	1396.64 (3633.43)	-0.99 (2.58)	0.01	$\alpha = 0$ $\beta = 0$
Poultry	240062 (9086282.70)	-181.54 (6798.76)	0.0001	$\alpha = 0$ $\beta = 1$

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

TABLE 7

Regression analysis results and the test of hypothesis for the third model for the production of important agricultural crops and number of animals in the Kingdom.

Production	α	β	C	R^2	Test of Hypothesis Results
Wheat	-116766 (22418.47)	83.40 (15.99)	0.76	0.98	$\alpha \neq 0$ $\beta \neq 1$
Barley	-215343 (79275.19)	152.97 (56.32)	0.60 (0.45)	0.83	$\alpha \neq 0$ $\beta \neq 1$
Tomato	-21214 (13765.41)	15.25 (9.93)	0.28 (0.76)	0.53	$\alpha = 0$ $\beta = 1$
Watermelon	-6602.14 (7037.53)	5.18 (5.08)	-0.73 (0.068)	0.18	$\alpha = 0$ $\beta = 1$
Dates	-16632 (3413.82)	12.13 (2.48)	0.09 (0.18)	0.97	$\alpha \neq 0$ $\beta \neq 1$
Grapes	-8132.24 (2851.77)	5.85 (2.05)	-0.13 (0.43)	0.91	$\alpha \neq 0$ $\beta \neq 1$
Sheeps	39506 (107105.86)	-28.13 (77.15)	1.02 (0.30)	0.87	$\alpha = 0$ $\beta = 1$
Camel	-9489.28 (9192.58)	7.31 (6.94)	-1.04 (1.60)	0.18	$\alpha = 0$ $\beta = 1$
Poultry	-21004341 (21565614.26)	14920 (15322.81)	0.58 (0.38)	0.80	$\alpha = 0$ $\beta = 1$

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

TABLE 8

Regression analysis results and the test of hypothesis for the fourth model for the production of important agricultural crops and number of animals in the Kingdom.

Production	α	β	R^2	Test of Hypothesis Results
Wheat	-303.39 (260.66)	1.03 (0.09)	0.96	$\alpha = 0$ $\beta = 1$
Barley	135.76 (316.92)	0.79 (0.49)	0.24	$\alpha = 0$ $\beta = 1$
Tomato	-81.68 (206.13)	1.19 (0.52)	0.40	$\alpha = 0$ $\beta = 1$
Watermelon	697.45 (289.08)	-0.72 (0.72)	0.13 (0.6)	$\alpha \neq 0$ $\beta \neq 1$
Dates	30.14 (108.00)	0.92 (0.21)	0.74	$\alpha = 0$ $\beta = 1$
Grapes	8.22 (16.76)	0.88 (0.16)	0.79	$\alpha = 0$ $\beta = 1$
Sheeps	1267.01 (1196.55)	0.77 (0.18)	0.66	$\alpha = 0$ $\beta = 1$
Camel	77.56 (296.81)	0.79 (0.73)	0.11	$\alpha = 0$ $\beta = 1$
Poultry	-10663 (39784.21)	0.70 (0.22)	0.56	$\alpha = 0$ $\beta = 1$

Source: Regression results executed by the author.
Figures between the brackets denote the standard error.
At level of significance = 5%.

TABLE 9

Comparison of test of hypothesis among different models for crops and animals in study.

First Model		Second Model		Third Model		Fourth Model	
Area	Production	Area	Production	Area	Production	Area	Production
	Barley $\beta \neq 1$		Wheat $\alpha \neq 0$ $\beta \neq 0$	Wheat $\alpha \neq 0$	Wheat $\alpha \neq 0$ $\beta \neq 1$	Water-melon $\beta \neq 1$	
	Tomato $\beta \neq 1$		Barley $\alpha \neq 0$ $\beta \neq 0$	Barley $\alpha \neq 0$ $\beta \neq 1$	Barley $\alpha \neq 0$ $\beta \neq 1$	Dates $\beta \neq 1$	Water-melon $\alpha \neq 0$ $\beta \neq 1$
	Water-melon $\alpha \neq 0$ $\beta \neq 1$				Dates $\alpha \neq 0$ $\beta \neq 1$		
					Grapes $\alpha \neq 0$ $\beta \neq 1$		

be estimated according to a mathematical model[1] intermediated by parameters not of area and time[2], such as expected atmospherical and environmental factors, the change in agricultural policies and technical progress.

(1) Like simulation model.

(2) Ministry of Agriculture and Water used the general trend equation in most of its estimates, it also uses the area in production estimates.

Summary and Recommendations

Forecast accuracy measurement is useful for comparing different ways for sampling and analysis. It is beneficial in forecast model comparison.

The comparison of forecasts with actual statistics gives an accurate measure. In the evaluation of agricultural crop forecasts, accuracy is useful for governmental corporations and agricultural companies that make large contracts on an annual basis, in part, on these forecasts. Therefore, the inefficiency of forecasts may have undesirable effects in development and resource allocation.

By comparing the four models, it is clear that the third model is the best. It is possible to say that, this model can estimate forecasts very close to accuracy compared with other models, but we have to say that, this model can be illusive in the case of estimation of wheat and barley area, and wheat, barley, dates and grapes production.

This study recommends the following: (1) the agricultural data should be modernized and available and may be outspread at the end of every agricultural season; (2) there is no need to depend on a fixed and limited equation (like general trend or link the production by the area only) for every crop in estimation of actual data or forecasts; (3) we must calculate the actual data from real figures after some limited time (10 years for example); and (4) we must estimate the forecast for years to come according to a mathematical model with parameters, not only with respect to time and area, but also with respect to the expected environmental and atmospherical factors, the changes in agricultural policies, and technical progress.

References

Crowder, R.T. 1972. Statistical vs. judgement and audience considerations in the formulation and use of econometric models. *American Journal of Agricultural Economics* 54 (5):779-783.

Gorham, M. 1978. Public and private commodity markets. Federal Reserve Bank of San Francisco. *Economic Review* Spring, pp. 30-38.

Gunnelson G., W. D. Dobson, and S. Pompefin. 1972. Analysis of the accuracy of USDA Crop Forecasts. *American Journal of Agricultural Economics*. 54(4): 639 - 645.

Madura, A., D. Martin, and M. Wiley. 1998. Forecast bias and accuracy of exchange rates in emerging markets. *Journal of Multinational Financial Management* 9(1): 27-43.

Paulino, L. and S.S. Tseng. 1980. A comparative study of FAO and USDA data on production, area, and trade of major food staples. *International Food Policy Research Institute,* Research Report No. 19.

Smith, S.K. 1992. Evaluating the forecast accuracy and bias of alternative population projections for states. *International Journal of Forecasting*. 8(3):495-508.

Tomek, W.G. and K.L. Robinson. 1982. *Agricultural Product Prices,* 2nd edition, Cornell University Press, pp. 352-359.

Wilkie, M.E. and A.C. Pollock. 1996. An application of probability judgement accuracy to currency forecasting. *International Journal of Forecasting* 12(1):25-40.

Received 31 October 1999.
Accepted 15 November 2000.