Alternative Forecasting Techniques for Vegetable Prices in Senegal

Dr Alioune Dieng[1]

Abstract

The objectives of this study are to investigate the performance of parametric models for forecasting selected vegetable prices and to make recommandations to potential users. Two forecasting approches are used. The forecasts obtained from the methods and evaluated using qualitative and quantitative criteria. The forecasting methods used consist of three alternative parametric models and a non parametric model. The parametric models consist of the naive model, the exponential smoothing and the Box and Jenkins autoregressive integrated moving average (ARIMA) model. The non parametric model uses the technique of spectral analysis. The findings of this study suggest that among parametric models Box and Jenkins'autogressive integrate moving average model will be a good technique to use in generating vegetable price forecasts for producers and consumers. However, additional research is needed to test the forecasting accuracy of parametric versus non parametric models with respect to other crops.

Keywords: Parametric and non parametric models-forecasting - vegetable prices - producers-consumers.

Résumé

Les objectifs de cette étude sont d'étudier la performance des modèles paramétriques pour prédire les prix des produits horticoles choisis et de faire des recommandations aux utilisateurs potentiels quelques modèles de prévisions. Deux approches de prévisions ont été utilisées pour produire des prévisions d'échantillons de prix de produits horticoles choisis. Les prévisions obtenues à partir des méthodes ont été alors évaluées en utilisant des critères qualitatifs et quantitatifs. Les méthodes de prévisions utilisées se composent de trois modèles paramétriques alternatifs et d'un modèle non paramétrique. Les modèles paramétriques sont le modèle naïf, le lissage exponentiel et le modèle intégré auto-régressif de la moyenne mobile de Box-Jenkins (ARIMA). Le modèle non paramétrique utilise la technique de l'analyse spectrale. Les résultats de cette étude suggèrent que le modèle paramétrique de Box-Jenkins est une bonne technique pour les prévisions des prix horticoles pour des producteurs et des consommateurs. Cependant, une recherche additionnelle est nécessaire pour évaluer la précision des prévisions des modèles paramétriques en comparaison aux modèles non paramétriques en ce qui concerne les autres produits.

Mots-clés : Modèles paramétriques et non paramétriques - prévision - prix des produits horticoles - producteurs - consommateurs.

[1] Institut sénégalais de recherches agricoles (ISRA),
Bureau d'analyses macro-économiques (BAME)
Bp : 3120
Tél. : (221) 33 859 17 59
E-mail : dieng alioune@hotmail.com

Introduction

In Senegal, vegetable production and consumption are being more and more popular. The reasons are related to health concerns and the search for alternative economic opportunities. The fast growing industry consists of a wide array of crops and products including potatoes, tomatoes, onions, various bean types and cabbage, etc. The cash receipts from all vegetables increased from 0.92 billion FCFA in 1980 to 5.8 billion FCFA in 2003 (DH, 2003; DPS, 2004). Together with fruits, vegetables account for about 15% of consumers's at home food budget today and represent 18% of food store sales. The per capita consumption of fresh market vegetable increased sharply from 12% during the 1980-1990 decade to 35% during the 1990-2003 decade. Vegetable produced in the country go either into the fresh market (about 65%) or export market/processing (35%).

The increase in the demand for vegetables will be driven mainly by increasing wealth and the need of consumers for fiber, low cholesterol, low fat and high vitamins A and C (Love, 1991). On the supply side, the desire for more profits will motivate producers to expand the supply of vegetables. Vegetable prices will play a major role in coordinating the supply and demand of these products. Hence, forecasts of vegetable prices will be useful to producers, consumers, processors, rural development planners and other people involved in the vegetable market.

Fildes and Lusk (1984) advise that forecasters should consider a range of methods and analyze their comparative performance over a random selection of series. Previous research efforts in forecasting economic variables have focused mainly on parametric methods. There is no reason why forecasters should not experiment with non parametric techniques. Therefore, the objectives of this study were to investigate the performance of parametric models for forecasting selected vegetable prices and then to make recommandations to potential users when they are to select models for forecasting specific vegetable prices.

Methodology

Two forecasting approches were used to generate samples of forecast models of selected vegetable prices. The forecasts obtained from the methods were then evaluated using qualitative and quantitative criteria. The forecasting methods used consist of three alternative parametric models and a non parametric model. The parametric models consist of the naive model, the exponential smoothing and the Box and Jenkins autoregressive integrated moving average (ARIMA) model. The non parametric model used the technique of spectral analysis. The general formulation for each of these approaches is given below:

The Naive Model

This Naive Model assumes that the forecast for the price of crop i in period t + 1 is equivalent to its immediate price in period t. This model can be represented as follows:

$$X_{it+1} = VX_{it} \quad (1)$$

where

X_{it+1}: forecast price for i^{th} crop in time period t+1

X_{it} : Actual price for i^{th} crop in time period t

$V = 1$

Exponential Smoothing Model

The exponential smoothing model used in this study takes the form:

$$X_{it+1} = \alpha X_{it} + (1-\alpha) X_{it} - te_t \quad (2)$$

where

$0<\alpha<1$

where X_{it} represents the actual price of the i^{th} crop in time period t, X_{it} represents the forecast value of i^{th} crop in period t and X_{it+1} represents the forecast value in time t+1 and α is the smoothing parameter.

Box-Jenkins Arima Model

According to Box and Jenkins (1976), an economic variable, say X has a generating function which belongs to a class of autoregressive moving average models (ARIMA) which can be represented as:

$$\Phi_p (ß) (1-ß)^d (X_t) = \Theta_q (ß)e_t \quad (3)$$

where

ß : lag operator such that $ß_j X_t = X_{t-j}$

Φ_p : autoregressive parameter of order p

Θ : moving average parameter of order q

d : number of times (usually 0,1, or 2) that X_1 needs to be differenced to achieve stationarity.

Compared to spectral methods, Box-Jenkins forecasting approach involves an interactive process between the forecaster and the data in terms of using diagnostic statistics to select the appropriate models. In addition, the Box-Jenkins approach requires less data and has generally proved successful in practice.

Spectral Forecasting

Spectral forecasting involves the use of frequency domain techniques to fit a «generic» Box-Jenkins model to the data. It is an automatic procedure once the preliminary data preparation and transformation is completed. As pointed out by Doan (1992), it can produce good forecasts quickly and painlessly given adequate data. It has the potential to do better than Box-Jenkins model especially when the latter is over-parameterized.

Let X represent a time series whose forecasts are desired. As shown by Koopmans (1974), the spectral density of X can be written as the z-transform

$$f_x(z) = c(z)\, c(z^{-1})\, \sigma^2 \qquad (4)$$

Assuming reasonable regularity conditions on c, log f_x has a Laurent expansion

$$\log f_x(z) = d(z) + d(z^{-1}) + d\Theta \qquad (5)$$

where d is one sided polynomial in positive of z. Taking the exponent of 4 and combining it with equation 5 gives:

$$c(z)\, c(z^{-1})\, \sigma^2 = \exp d(z) \exp d(z^{-1}) \exp(d\Theta) \qquad (6).$$

Data, Estimation and Results

Monthly average consumer prices for tomato, potato and onion were obtained from the Direction de la Statistique et de la Prévision and the Direction de l'Horticulture. These crops were selected partly because of their relative importance in the vegetable industry and partly because of data availability. The price series were expressed in francs CFA per kilo and cover the time period 1980 to 2003. The observations for January and February 1987 were missing for tomato and potato. A simple 3 year moving average of the preceding observations for these crops were computed to fill in the missing observations. The 3 price series are characterized by fluctuations which is most noticeable in the case of tomato. The sharp fluctuations in tomato prices largely reflect the relatively high perishability of this commodity.

The naive model was estimated using EXCEL. Essentially, this estimation involved lagging each crop price series by one period such that the price forecast of a given crop in period t is taken to be its price in period t-1. The exponential smoothing model was estimated using Statistical Package for Social Sciences (SPSS). Alternative smothing parameters were tried between 0.01 and 0.50 following Bowerman and O'Connellius suggestion (1987). The optimal value was 0.30 based on minimum sum of squared error of forecast.

The ARIMA model presented in equation 3 was estimated, using the standard Box-Jenkins iterative three stages process of identification, estimation and diagnostic checking (Pindick and Rubinfiled, 1976). It was necessary to transform the original data to assure variance stabilization and stationarity. First differencing operations proved effective for this purpose for each of the three crops. The estimated results were presented in Table 1.

Table 1: Estimated ARIMA Results for Selected Vegetable Crop Prices

Crop	Model	Estimation
Tomato	(1,1,2)	$(1 + .37B)\,(1\text{-}B)\,X_t\;(1 - .47B - 0.47B^2)$ (2.87) (-3.96) (-4.58)
Potato	(1,1,3)	$(1 + .66B)\,(1\text{-}B)\,X_t\;(1 - .22B - .41B^2 - .23B^3)$ (6.23) (-1.76) (-4.84) (-2.37)
Onio	(2,1,0)	$(1 + 0.40B - 0.15B^2)\;(1\text{-}B)\,X_t$ (4.66) (-1.76)

X_t = Transformed price series.

For all models presented in Table1, the estimated coefficients were statistically significant. In addition, the Ljunk-Box Q statistics associated with each model fall with in acceptable region when compared with the appropriate X^2 distribution. Each estimated model was therefore used to generate a sample of price forecasts for the relevant crop.

Doan (1992) provides some insight for estimating with a spectral model. For equation 4 above the log spectral density for X is first computed. Next is to mask the negative and zero frequencies to get the Fourier transform (FT) of d (one-sided polynomial in positive power of z). Then the exponent of the FT is taken to get the FT of c. The latter is divided into the FT of X to get the FT of ε process. To get the desired forecast, set ε to zero outside of the range of actual data and ε by c in the frequency domains. Initially each of the series was transformed to achieve stationarity as was done in the case or ARIMA application. Then using SPSS program spectral forecasting was performed for each price series. The forecasts generated from the program were recorded and compared with those produced by the other models.

Evaluation of Forecasts

The forecasts generated by each model were evaluated against the actual values. The evaluation was based on qualitative as well as quantitave criteria. The qualitative evaluation was done by applying the 4X4 turning points (TP) contingency table suggested by Naik and Leuthold (1986). The quantitative evaluation was based on the percentage root mean squared error (PRMSE).

The 4X4 contingency table for qualitative evaluation led to the computation of 3 ratios. These ratios include the ratio of accurate forecasts (RAF) defined as the number of perfect model forecasts devided by the total number of forecasts; the ratio of worst forecasts (RWF) defined as the number of model forecasts which were opposite the direction of actual movement; and the ratio of inaccurate forecasts (RIF) which is defined as the number of TPs inaccurately by the model (excluding the worst cases). Results of the evaluation of the forecasts of the competing models are presented in Tableau 2.

The results in Table 2 are mixed in terms of the relative performance of parametric models versus the non-parametric models. Based on the RAF and the RIF values, all the parametric models (Naive, ES and ARIMA) clearly outperformed the non-parametric (spectral model only in the case of potato price. For tomato and onion, there is a very close competition between the parametric and the non parametric models. For example, the spectral model outperformes the Naive and the ES models for the two crops but was outperformed by the ARIMA model in the case of onion. Considering the turning point predictions for tomato, the ARIMA and the spectral model produced equivalent results.

Table 2: Turning Point Evaluation of Alternative models in Forecasting Senegalese Selected Vegetable Prices

Commodity	**Model**	**Ratios**		
		RAF (a)	RWF (b)	RIF (c)
Potato:	Naive	0.50	------	0.50
	ES	0.50	0.17	0.35
	ARIMA	0.33	0.25	0.42
	Spectral	0.17	------	0.83
Tomato:	Naive	0.42	0.16	0.42
	ES	0.42	0.25	0.33
	ARIMA	0.50	0.08	0.42
	Spectral	0.50	0.08	0.42
Onion:	Naive	0.42	0.16	0.42
	ES	0.50	0.33	0.17
	ARIMA	0.67	0.08	0.25
	Spectral	0.58	0.08	0.34

Note: a) Ratio of accurate forecasts
b) Ratio of worst forecasts
c) Ratio of inaccurate forecasts.

The highest RAF value (0.67) occurred with respect to onion price forecast from the ARIMA model. This value implies that 67% of the time, the ARIMA-based forecast of onion price accurately predicted the movements in the actual price of onion. The lowest RAF value, 17% was produced by the spectral model in forecasting potato price.

The results are also mixed in terms of intra-parametric model performance. While the ARIMA model outperformed the naive and the exponential smoothing in predicting turning points for tomato and onion, it is outperformed by these 2 models when the commodity is potato.

In general, it appears that the relative performance of parametric versus non-parametric model may be commodity specific with respect to turning point prediction. For potato prices the parametric models proved superior to the non-parametric model. For tomato and onion however, there exists some «gray areas» between the two groups of models.

The forecasting accuracy of the model used in this study was also evaluated quantitatively, based on the percentage root mean squared error (PRMSE). The results of this evaluation are reported in Table 3.

For each crop in Table 3 the PRMSE value associated with the non-parametric spectral model are substantially higher than those associated with any of the parametric models. Therefore, the superiority of the parametric model over the non-parametric model in forecasting the price of each crop is unquestionable. The forecast performance of the spectral model is worst for potato (69% error) and best for tomato (44% error).

Table 3: Quantitative Forecast Evaluation of Alternative Models for Selected Crops

Model	PRMSE*		
	Potato	Tomato	Onion
Naive	0.07	0.21	0.11
ES	0.24	0.23	0.25
ARIMA	0.07	0.16	0.09
Spectral	0.69	0.44	0.58

*PRMSE = Percentage Root Mean Squared Error

Among the parametric models, the ARIMA model appears to the best forecaster of the prices of the three crops. The Naive model ranks second to the ARIMA model, except in the case of potato price where the two models performed equally well.

The best price forecasts were obtained for potato in which the ARIMA and the naive models each produced a PRMSE value of 0.07. In other words, the forecast error of potato price generated by the ARIMA and the naive models was 7%.

Conclusions

Chances are that the Senegalese vegetable industry will continue to gain increasing importance due to dietary, health and other concerns. Price will play a major role in the allocation of the resources between production and consumption. Various participants in the vegetable markets (farmers, producers, buyers, etc.) will be in constant need of price forecasts to facilitate accurate decision making.

Previous research work on forecasting agricultural commodity prices have focused mainly on using parametric methods with little attention to non parametric methods. Under certain circumstances, the latter may yield better forecasts than the former. In this study both parametric methods and non parametric methods were used to generate forecasts of selected Senegalese vegetable crop prices (potato, tomato and onion).

The results indicate that with respect to turning point predictions the relative performance of parametric and non parametric models was commodity specific. For potato prices, all the parametric and non parametric models (naive, exponential smoothing and the ARIMA) were clearly superior to the non parametric spectral model. In the case of the other two crops (tomato and onion) the non parametric spectral model outperformed only of the parametric models (naive and exponential smoothing).

When evaluation was based on the quantitative criterion of the percentage root mean squared error the parametric models substantially outperformed the non parametric spectral model in terms of forecasting acuracy.

Based only on the results of this study, parametic models would be recommended for forecasting Senegalese vegetable prices. Among these models, Box and Jenkins time series models should receive high priority. The fact that these models are amenable to the use of diagnostic statistics for selecting the appropriate models explains their tendency to be superior, hence their desirability. The economic well beings of the small scale farmers in the rural areas depends to a large extent, on how efficiently they allocate their scarce resources. Agricultural professionals must continue to provide assistance to these farmers in terms of entreprise choice and technical support. The findings of this study suggest that among parametric models Box and Jenkins' autogressive integrated moving average model will be a good technique to use in generating vegetable price forecasts for producers and consumers. Howerver, additional research is needed to test the forecasting accuracy of parametric versus non parametric models with respect to other crops.

References

Box, G.E.P. and G.M.Jenkins. «Time Series Analysis: Forecasting and Control». San Francisco. Holden-Day, 1976.

Bowerman B.L. and R.T. O'Connel. «Time Series Forecasting: Unified Concepts and Computer Implementation». Boston: Duxburg Press, 1987.

Direction de la statistique et de la prévision (DSP). Situation économique du Sénégal, 2004.

Direction de l'horticulture du Sénégal (DH). Rapport annuel, 2003.

Doan, T.A. «Regression Analysis of Time Series, User's Manual Version 4». Evanston: Estima, 1992.

Fildes, R. and E.J., Lusk. «The Choice of a Forecasting Model». Omega, 12(5) (1984), 427-35.

Koopmans L. H. «The Spectra Analysis of Time Series». New York: Academic Press, 1974.

Love, M.J. «The Produce Industry's Challenge to Double Demand for Fresh Vegetables» in Vegetables and Specialities: Situation Outlook Yearbook. Washington: USDA/ERS, Dec 1991.

Naik, G. And R.M. Leuthold. «A note on Qualitative Forecast Evaluation». American Journal of Agricultural Economics. 68 (1986): 721-726.

Pyndick R.S. and D.L. Runbinfed. «Econometric Models and Economic Forecasts». New York: McGraw-Hill, Inc. 1976.

Vient de paraître