Proceedings of the 2010 International Conference on Industrial Engineering and Operations Management
Dhaka, Bangladesh, January 9 – 10, 2010

Intermittent Demand Forecast and Inventory Reduction Using Bayesian ARIMA Approach

Mohammad Anwar Rahman
Industrial Engineering Technology
The University of Southern Mississippi
730 East Beach Blvd, Long Beach, MS 39560, USA

Bhaba R. Sarker
Department of Industrial Engineering
Louisiana State University
Baton Rouge, LA 70803, USA

Abstract

Natural calamities (e.g., hurricane, excessive ice-fall) may often impede the inventory replenishment during the peak sale season. Due to the extreme situations, sales may not occur and demand may not be recorded. This study focuses on forecasting of intermittent seasonal demand by taking random demand with a proportion of zero values in the peak sale season. Demand pattern for a regular time is identified using the seasonal ARIMA (*S-ARIMA*) model. The study proposes a Bayesian procedure to the ARIMA (*BS-ARIMA*) model to forecast the peak season demand which uses a dummy variable to account for the past years intermittent demand. To capture uncertainty in the *B-ARIMA* model, the non-informative prior distributions are assumed for each parameter. Bayesian updating is performed by Markov Chain Monte Carlo simulation through the Gibbs sampler algorithm. A dynamic programming algorithm under periodic review inventory policy is applied to derive the inventory costs. The model is tested using partial demand of seasonal apparel product in the US during 1996-05, collected from the US Census Bureau. Results showed that, for intermittent seasonal demand forecast, the *BS-ARIMA* model performs better and minimizes inventory costs than do *S-ARIMA* and modified Holt-Winters exponential smoothing method.

1. Introduction

The random occurrences of seasonal demand are studied to manage the inventory at minimum costs. The inventory replenishment and sales can be interrupted due to natural calamities. Demand of such products is often intermittent which may contain a proportion of zero values. Traditional forecasting methods can be inappropriate for forecasting such seasonal demand that increases during specific time period in a year. These forecast obtained by the autoregressive integrated moving average (ARIMA) models are found to be better compared to other ways of modeling. The focus of this paper is to demonstrate the inventory cost reductions through the application of an appropriate demand forecast of a seasonal product. A time series from January 1996 to June 2005 of a seasonal apparel demand in the U.S., collected from the U.S. Census Bureau, is selected. To exhibit the intermittent features in the selected time series, six arbitrarily chosen values during a peak season from July-December 2004 are considered unavailable. A seasonal ARIMA (*S-ARIMA*) model has been constructed to forecast the peak season demand from July-December 2005.

Time series forecasting models are increasingly applied to forecast seasonal demand and short-life products Makridakis *at. el*. (1998). Gardner and Diaz-Saiz (2002) analyzed forecasting implementation problems in inventory control systems (safety stock investment) for seasonal time series. Under an autoregressive moving average (ARMA) assumption, Kurawarwala and Matsuo (1998) estimated the seasonal variation of PC products demand using demand history of pre-season products. Miller and Williams (2004) incorporated seasonal factors in their model to improve forecasting accuracy while seasonal factors are estimated from multiplicative model. Hyndman (2004) extended Miller and Williams' (2004) work by applying various relationships between trend and seasonality under seasonal ARIMA procedure. The classical ARIMA approach becomes prohibitive, and in many cases it is impossible to determine a model, when seasonal adjustment order is high or seasonal adjustment diagnostics fails to indicate that time series is sufficiently stationary after seasonal adjustment. In such situations, the static parameters of the classical ARIMA model are considered the main restriction to forecasting high variable seasonal demand. Another restriction of the classical ARIMA approach is that it requires a large number of observations to determine the best fit model for a data series. In the ARIMA model, if the Bayesian approaches are used, the restriction of the static values of the parameters is relieved by imposing the probability distributions to represent the parameters. Although the practices of Bayesian ARIMA models for seasonal forecast are more realistic, but the literature on Bayesian methods applied to seasonal ARIMA time series is limited. In this study, after the classical ARIMA model is developed for the selected dataset, and the Bayesian method is applied to the *S-ARIMA* model. A Bayesian approach to the *S-ARIMA* model is studied with the special relevant to the problem that forecast should be performed from an incomplete time series. An adaptive approaches of Holt-Winters' (H-W) exponential smoothing technique is also studied to forecast the time series.

A periodic review inventory model has been studied depending on the forecast obtained by the above models. The optimal inventory costs have been derived for each forecast via dynamic programming. The best forecast is considered as the one which provides the minimum inventory costs.

2. *S-ARIMA* Model Identification

In this paper, the *S-ARIMA* process begins by transforming the original time series into a stationary series by taking the mean difference of data. The stationarity of the time series has been achieved after the first and seasonal ($d_{1,12}$) difference of order 2. The autocorrelation function (ACF) and partial ACF (PACF) have been identified as decreasing, sinusoidal and alternate in sign and showed that the order of p and q for both AR(p) and MA(q) components for seasonal and non-seasonal series can at the most be one. Using goodness-of-fit statistics, e.g., (i) Akaike information criterion (AIC), (ii) Schwarz Bayesian information criterion (BIC) and the chi-square (χ^2) test, the best structure of *S-ARIMA* (p, d, q)×(P, D, Q)$_{12}$ has been identified as *S-ARIMA* (0,1,1)(1,1,0)$_{12}$. The estimated parameters of the *S-ARIMA* are reported in Table 1.

Table 1: Estimated values of the *S-ARIMA* parameters

Parameter	Estimate	Standard error	t-value	Pr>|t|	lag
MU	1226.80	7889.10	0.16	0.8700	0
MA1,1	0.74	0.07	10.59	0.0001	1
AR1,1	-0.35	0.09	-3.59	0.0003	12

Both the moving average "MA1,1" and the autoregressive AR1,1 parameters are 0.735 and -0.35, respectively with significant t values. The parameters are estimated using the maximum likelihood method by SAS software.

3. Point Forecast with *S-ARIMA* Model

The *S-ARIMA* model (0,1,1) (1,1,0)$_{12}$ has been identified best model for the time series and entailed to forecast the peak demand. In the *S-ARIMA* model, the autoregressive term p = 0, P = 1 (seasonal) [that is, $(1-0)(1-\phi_1 B^{12})$]; differencing term d = 1, D = 1 (seasonal difference) [that is, $(1-B)(1-B^{12})$] and moving average term is q = 1, Q = 0 (seasonal) [$(1-\theta B)(1-0)$] . The model has (1, 12) period differencing, the autoregressive factors is $(1-\phi B^{12})$ = (1 + 0.35096 B^{12}), the moving average factors is $(1-\theta B)$ = (1 - 0.73527 B), and the estimated mean, C = 1226.8. Transforming autoregressive terms and coefficient, the model is given by

$$y_t = \frac{C + (1-\theta B)e_t}{(1-\phi B^{12})\ (1-B)\ (1-B^{12})} . \tag{1}$$

After expanding, and transforming the back operator, Equation (1) can be simplified in the following form,

$$y_t = y_{t-1} + (1+\phi) y_{t-12} - (1+\phi) y_{t-13} - \phi y_{t-24} + \phi y_{t-25} + e_t - \theta e_{t-1} + C . \tag{2}$$

In order to forecast one period ahead, that is, y_{t+1}, the subscript of is increased by one unit, throughout, and using Using ϕ = -0.35 and θ = 0.735, the Equation (2) is given by

$$y_{t+1} = y_t + 0.65y_{t-11} - 0.65y_{t-12} + 0.35y_{t-23} - 0.35y_{t-24} + e_{t+1} - 0.735e_t + 1226.8 \quad (3)$$

In order to forecast for the period 115 (i.e., July 2005), Equation (3) is given by

$$y_{115} = y_{114} + 0.65y_{103} - 0.65y_{102} + 0.35y_{91} - 0.35y_{90} + \hat{e}_{115} - 0.735\hat{e}_{114} + 1226.8$$

The forecast results for stage-2 at 2005 using *S-ARIMA* is shown in Table 4.

4. Bayesian Sampling-based ARIMA Model (*BS-ARIMA*)

Bayesian methods have been widely applied in time series context. The Markov Chain Monte Carlo (MCMC) method is efficient and flexible algorithms for conducting posterior inference of Bayesian model through simulation. Depending on the time series data, a Markov chain can be constructed in various ways and the Gibbs sampler, a commonly used algorithm applied here to derive the posterior parameters of the forecasting model. The key advantage of developing *S-ARIMA* model from Bayesian perspectives is the capacity to forecast demand from an incomplete time series which contains both zero and non-zero data. In this model, the form of *S-ARIMA* $(0,1,1)(0,1,1)_{12}$ model in Equation (3) has been applied. The model can be expressed in the following,

$$y_t = y_{t-1} + \phi_1 y_{t-12} + \phi_2 y_{t-13} - \phi_3 y_{t-24} + \phi_4 y_{t-25} + e_t - \theta e_{t-1} + C$$

where $\phi_1 = (1+\phi)$, $\phi_2 = -(1+\phi)$, and $\phi_3 = \phi_4 = \phi$. It is noted that the demand for the stage-2 period from July-December 2004 are not available. A dummy variable w_t is added to Equation (7) to account the missing values of the data series. The form of the *BS-ARIMA* after adding dummy variable is given by

$$y_t = y_{t-1} + \phi_1 y_{t-12} + \phi_2 y_{t-13} - \phi_3 y_{t-24} + \phi_4 y_{t-25} + e_t - \theta e_{t-1} + C + w_t \quad . \quad (4)$$

The dummy variable w_t, $0 \le w_t \le 1$ is added to represent the status of past information, e.g., if the dummy variable is set to 'zero' when demand information of a period is known. A scaled value of w_t may be set (from 0 to 1) to reflect the partial demand of a period. The value of w_t as 1.0 when demand information for a period is missing, while the value 0.50 indicates incomplete demand information, i.e., approximately 50% demand was observed. For dummy variable w_t, the values placed at July to December 2004 are shown in Table 2.

Table 2: Values of dummy variables for July to December, 2004 (units in million)

Demand	Jul	Aug	Sep	Oct	Nov	Dec
Projected, y_p	2.54	3.81	4.14	4.05	2.41	1.75
Stage-1, y_{st-1}	1.55	1.55	1.55	1.55	1.55	1.55
Dummy variable, w_t	0.39	0.59	0.63	0.62	0.36	0.11

For the data series y_t, (t = 1, 2, …, n, n+1, …N), the y_t corresponds to the demand of a period t, where a vector time series from n+1 to N, $\{y_F = (N-n) \ge 1\}$ is the prediction periods. A Bayesian computation is carried out to predict the demand for (N-n) period through the use of sampling-based algorithm. The particular sampling-based approach used in this model is a Markov chain Monte Carlo method based on the Gibbs sampler algorithm. The likelihood function for n observation y_t, (y_1, y_2, …, y_n) is denoted by $f(y;\psi)$, where $\psi = (\phi_i, \theta, \beta, \tau)$ with $\phi_i = (\phi_1, \ldots, \phi_4)$. The conditional likelihood obtained from the factorization theorem (Zellner, 1996) is given by

$$f(y_t|\Psi) = f(y_1|\Psi) f(y_2|y_1,\Psi) \ldots f(y_n|y_1,\ldots,y_{n-1},\Psi).$$

Given the prior distribution for Ψ, $f(\Psi|y_t)$, the posterior density for Ψ is given by $f(\Psi|y_t) \propto f(y_t|\Psi).f(\Psi)$.

If $y_F = (y_{n+1},\ldots,y_N)$, for predicting ($N$ - n) = L period, the predictive density is given by

$$f(y_F|y_t) = \int (y_F|y_t,\Psi).f(\Psi)d\Psi , \quad (5)$$

where $\int (y_F|y_t,\Psi)$ is the density of the future data y_F. The L steps ahead forecast is then

$$f(y_F|y_t,\Psi) = \int (y_{n+1}|y_t,\Psi) \int (y_{n+2}|y_{n+1},y_t,\Psi) \ldots \int (y_{n+L}|y_{n+1},y_{n+L-1},y_t,\Psi)d\Psi .$$

To obtain a sample of predictions from the density function in Equation (9), for each Ψ_t, one needs to draw from $\int (y_F|y_t,\Psi)$. Following are the steps to predict the future values of the *BS-ARIMA* model,

Step 1: Data Definitions

y_t , {for t in (1: n) }

w_t , {Dummy (t), for t in 1: N}

Step 2: Model Description

$y_t \sim Normal(\mu_t, \tau)$ {for t in (2: n)}, where

$\mu_t = C + \phi_1 y_{t-12} + \phi_2 y_{t-13} + \phi_3 y_{t-24} + \phi_4 y_{t-25} + e_t + \theta_1 e_t + \beta w_t$, and $\tau = 1/\sigma^2$

Step 3: Assigning Priors

μ ~ *Normal* (0, 0.001), ϕ_i ~ Normal(0, 0.001), θ_i ~ Normal(0, 0.001),

β ~ Normal(0, 0.001), τ ~ Chi-sq (1)

Step 4: Forecasts Period {$t = n$ +1 … N}

$y_{(new)t} \sim Normal[\mu_{(new)t}, \tau]$

$\mu_{(new)t} = C + \phi_1 y_{t-12} + \phi_2 y_{t-13} + \phi_3 y_{t-24} + \phi_4 y_{t-25} + e_t + \theta_1 e_t + \beta w_t$ {for t in (n +1: N)}

Carlin and Gelfand (1990) illustrated that the point estimates arising from $\int (y_F \mid y_t, \Psi)$ perform well if the variance of the predictive distribution remain small. In '*Step* 3', the following non-informative prior distribution has been used for each parameter. The prior distributions *Normal*(0, 0.001) are assumed for coefficient ϕ and θ. Parameter β is expected to follow a relatively informative prior distribution as *Normal*(1.0, 0.1). The precision (a reciprocal of variance), τ follows a chi squared distribution with one degree of freedom. The choice of prior distribution is followed by (Gelman *et. al.*, 2004; Congdon, 2003). In the model, the estimated parameters of Bayesian ARIMA model are shown in Table 3.

Table 3: Estimated parameters of Bayesian ARIMA model (units in million)

node	mean	St.dev	2.50%	median	97.50%	node	mean	St.dev	2.50%	median	97.50%
alpha1	0.15	0.16	-0.17	0.15	0.48	alpha4	0.03	0.22	-0.39	0.03	0.44
alpha2	0.23	0.15	-0.073	0.23	0.53	theta1	-0.01	0.67	-1.45	-0.008	1.38
alpha3	0.68	0.19	0.29	0.68	1.05	beta	1.00	3.22	-5.49	1.05	7.30

Actual demand, the simulated results of the *BS-ARIMA* and *S-ARIMA* results for stage-2 period of 2005 are shown in Table 4. The posterior models are derived using MCMC approach through Bayesian inference using Gibbs Sampling (BUGS) package

Table 4: Demand Forecast by *BS-ARIMA* model (units in million)

Month	*BS-ARIMA* model						*S-ARIMA* model			
	Actual demand	Estimate	Std error	2.5%	Med	97.5%	Estimate	Std error	2.5%	97.5%
Jul	2.83	2.47	0.71	1.07	2.46	3.88	2.74	0.38	0.38	3.50
Aug	3.33	3.62	0.73	2.19	3.61	5.07	3.65	0.40	0.40	4.34
Sep	4.10	4.77	0.74	3.32	4.76	6.22	4.92	0.41	0.41	5.96
Oct	5.31	5.29	0.73	3.89	5.28	6.76	5.03	0.42	0.42	6.91
Nov	3.46	4.28	0.77	2.75	4.28	5.76	2.69	0.44	0.44	4.31
Dec	2.28	2.97	0.77	1.53	2.96	4.50	2.70	0.45	0.45	3.58

Adaptive exponential smoothing forecasting (Holt-Winter) technique is widely spread in practice, has been used to compare the proposed model. Multiplicative exponential smoothing (*M-ES*) model, defined as $d_{t+T} = \left(\bar{R}_{t-1} + T\,\bar{G}_{t-1}\right)\bar{S}_{t+T-L}$ has been applied to compare the forecast for next T periods, where $\bar{R}_t$ is the estimate of level index, $\bar{G}_t$ is the estimate of trend, and $\bar{S}_t$ is the estimate of seasonal component (seasonal index). The initial values of the parameters are determined using the data from July 2002 to December 2002 and the values are modified in subsequent years. The data series from January 2003 to June 2005 are used to adjust the weight of the smoothing parameters and demand forecast is performed for the stage-2 (July to December) in 2005.

5. Inventory Cost Compare to Evaluate the Best Forecast

In this section, the cost saving approach in the inventory of the seasonal product is presented. A monthly review plan is considered for periodic inventory replenishment. There are t (t = 1, 2, …, n) forecasting periods at stage-2 and the demand forecast at any period t is y_t, while the actual demand for any period is x_t. Shortages may occur when $x_t > y_t$. The shortage cost is π_t dollars per period. To place an order for procuring y_t items, the fixed ordering cost is A dollars, unit purchasing cost is c dollars and unit holding cost is h dollars. Each unit brings a price of w dollars when it is sold, where $w > c$. Average fixed ordering cost per period is given by A/y_t; while revenue earned per period is $(w-c)y_t$, the average inventory per period is $h(y_t - x_t)/2$. In a periodic (y_t, L) replenishment policy, the aggregated total cost (TC) is given by $TC = (h/2)(y_t - x_t) + A/y_t$. The inventory cost and variable cost per unit per period (holding cost, setup cost, shortage costs etc.) are listed in Table 5. Holding cost rate is 30% per year. Therefore, holding cost h_t per month is, h_t = ($25)(0.30/year)/(12 months/year) = $0.624/month.

Table 5: Unit costs applied to the inventory model

	Parameters	Jul	Aug	Sep	Oct	Nov	Dec
D_t	Actual Demand (in million, $)	2.87	4.33	5.30	5.46	3.41	2.49
A_t	Fixed cost (in thousand, $)	15.0	14.0	16.0	16.0	17.0	19.0
c_t	variable cost ($)	25.0	25.0	24.0	25.0	26.0	30.0
π_t	shortage cost ($)	5.0	5.0	5.0	5.0	5.0	5.0
h_t	inventory cost ($)	0.624	0.624	0.624	0.624	0.624	0.624

The steps to compare the forecasting models using the inventory costs are described in the following,

Step 1: Find customer service level by specifying the probability (P_1) of no stock-out using periodic review policy.

Step 2: Select the safety factor z to satisfy $P(Z)$ = (1- P_1). The value of unit normal variable, $P(Z)$ can be obtained from Z-table or from inverse function of normal distribution.

Step 3: Determine the ordering quantity, Q, for each forecast using lead time $L(\hat{y}_L)$ and the safety stock (SS) where SS = $Z\sigma_L$ (σ_L is standard deviation of the forecast error) Therefore, the ordering quantity is

$$Q = \hat{y}_L + Z\sigma_L ,$$

(Ordering quantity may be increased to the next higher integer).

Step 4: Compute the optimal inventory cost of actual demand and demand forecasts using dynamic programming (DP) algorithm.

Step 5: Calculate the relative cost of each demand forecast with respect to the cost derived from actual demand and choose the best forecasting model that gives the minimum costs.

The inventory cost associated with actual demand is the least inventory cost, which is considered the base cost reference to the demand forecasts. The relative percent of inventory cost (RPIC) for each forecast is determined. These percentages for each forecast are compared and the minimum percentage value is considered the best forecast for the selected time series. The standard errors of forecast, MAPE (mean absolute percentage error), inventory costs and the percent above the least inventory cost for each demand forecast for the given data set are presented in Table 6.

Table 6: Inventory cost for each forecasting model and actual demand

Forecast Models	Std. Error (million)	MAPE	Inventory Cost (million)	Relative Percentages
S-ARIMA	0.55	13.18%	$666.75	27.05%
BS-ARIMA	0.31	7.43%	$561.46	6.99%
Exponential Smoothing (M-ES)	0.48	10.54%	$695.26	32.49%

6. Conclusion

In this paper an ARIMA approach is used to forecast the demand of a seasonal product. Based on the demand pattern, the *S-ARIMA* (0,1,1) $(1,1,0)_{12}$ model has been identified to be the best fit model for the time series. For a non stationary stochastic time series (such as winter apparel), the forecasting model often becomes complicated. In *S-ARIMA* model, forecast errors are incorporated to refine the predicted value, so the model gradually improves toward the end of the time series and provides satisfactory forecasting accuracy.

There are major advantages of using Bayesian methodology to forecast non-stationary demands. As classical ARIMA requires significantly long data series, a Bayesian-sampling based ARIMA (*BS-ARIMA*) model has been proposed to forecast from incomplete data with missing values. In *BS-ARIMA* model, it is assumed that data points at stage-2 (July to December) in 2004 were unavailable. A number of non-informative priors were used for the model parameters (α, β, τ). The posterior values of the parameters were computed numerically using the Markov Chain Monte Carlo (MCMC) simulation and BUGS/WinBUGS software. A multiplicative approach of *exponential smoothing* (*M-ES*) technique is considered as the base reference to forecast seasonal demand

A periodic review model has been used to evaluate the inventory costs for each forecast and acknowledged the cost savings due to improved forecast. The dynamic programming algorithm is used to derive the lowest inventory cost for each demand forecast and the actual data. Comparing the cost percentages of each demand forecast above the inventory cost of actual demand data, standard error, and mean absolute percent error, the analysis suggests that the *BS-ARIMA* model is well-performed forecasting model for time series and advantageous over many of the traditional forecasting models.

Reference

Carlin, B. and Gelfand, A.E., “A Sample Reuse Method for Accurate Parametric Empirical Bayes Confidence Intervals,” *Journal Royal Stat. Soc., Ser.* B 53, 189-200, 1990.

Congdon, P., “Applied Bayesian Modeling,” *England: John Willey & Sons*, 2003.

Gardner, E.S, and Diaz-Saiz, J., “Seasonal adjustment of inventory demand series: a case study,” *International Journal of Forecasting*, 18(1): 117–123, 2002.

Gelman, A., Carlin J.B., Stern H.S. and Rubin, D.B., “Bayesian Data Analysis,” *Chapman and Hall/CRC*, 2004.

Hyndman, R.J., “The interaction between trend and seasonality,” International Journal of Forecasting, 20(4), 561-563, 2004.

Kurawarwala, A.A, Matsuo H., “Forecasting and inventory management of short life-cycle products,” *Operation Research*, 44 (1) 131-150, 1996.

Makridakis, S. Wheelwright and Hyndman, “Forecasting: Methods and Applications”, 3rd ed. N.Y., John Wiley & Sons, 1998.

Miller, D.M. and Williams, D., “Damping seasonal factors: Shrinkage estimators for the X-12-ARIMA program,” *International Journal of Forecasting*, 20(4), 529-549, 2004.