Sugarcane transportation process modeling by time series approach

F. Afsharnia¹*, A. Marzban¹, A. Abdeshahi²

¹Department of Agricultural Machinery and Mechanization Engineering, Agricultural Sciences and Natural Resources University of Khuzestan, Ahvaz, Iran, I. R. Iran
²Department of Agricultural Economic, Agricultural Sciences and Natural Resources University of Khuzestan, Ahvaz, Iran, I. R. Iran

* Corresponding Author: afsharniaf@yahoo.com

ARTICLE INFO

ABSTRACT- Sugarcane is one of the severely perishable crops that is used as raw material for white sugar production. Sucrose content of the sugarcane which is of high commercial value decreases in quality due to pre-harvest burning, high ambient temperature, kill-to-mill delays as well as microbial contaminations. Delays in sugarcane transportation are the most important risks which can affect the quality and quantity of the product. Delay in milling of the harvested sugarcane is caused by various reasons in agro-industry units including factory downtime, breakdowns of tractors in the waiting line at factory, tractor accident in factory yard and staff shift changes creating long queues. In order to reduce delays, the present study attempted to forecast arrival and service level of tractor drawn carts which transfer burned or cut canes from farm to mill. The univariate ARMA models were applied to forecast arrival and service level. The RMSE and MAPE were also used to evaluate precision of our forecast. The results of models demonstrated that ARMA(4,3) and ARMA(4,2) models are suitable for arrival and service level of tractor drawn carts, respectively. The predicted values trend of arrival, and service level truly reflected the actual values of arrival and service level as well as queue system tendency. The values of MSE, RMSE and MAPE that indicate accuracy of the forecasted carts arrival and service level were relatively low. The estimated models can be used to forecast values of arrival and service levels of tractor drawn carts for subsequent hours during harvest season.

INTRODUCTION

Sugarcane, as a commercial and versatile crop, plays an important part role in Iran agriculture and industry economy. The market demand for sugarcane will continue to rise due to sucrose production and value-added products such as blackstrap molasses, bagasse, ethanol, chipboard as well as animal feed, since as one of the strategic goods is always considered and supported by governments (Le Gal et al., 2008; Chunhawong et al., 2018).

Commercial sugarcane crops (cp69-1062) are grown by furrow irrigation in Sugarcane and Byproducts Development Company, Khuzestan province, Iran. Sugarcane industry is Iran's first-largest agro-based industry providing employment to about 50,000 people and contributing to the growth of the vital rural economy. The Sugarcane and Byproducts Development Company of Khuzestan, as the largest sugar producer in Iran, was able to produce 700 thousand tons of sugar which was half of the country's required sugar. Over the last 45 years, sugarcane cultivation area remained relatively constant in Iran and covered 768,300 hectares in 2016-2017 growing season, while total sugarcane production gradually increased, from around 594,493 tons in 1971-1972 growing season to 7.52 million tons in 2016-2017 seasons due to increase in extracting sugarcane juice from the sugarcane sticks from 72.58% in 1971-1972 seasons to 83.48% in 2016-2017 seasons. Furthermore, the sugarcane yield has shown an increase by 102.83% from 24 tons per hectare to 53 tons per hectare in the years 1971-2016 (ISCRTI, 2017).

The harvesting and transportation operations for sugarcane mill are extremely dynamic. As the cane is harvested in the field, the harvester also cuts cane into uniformly sized billets (0.304-0.457 meters) and continuously feeds billets into cart that pulled by a medium powered tractor, usually Massy Freguson 399 (MF-399) tractor. This infield transporter and cart combination runs alongside the harvester during the harvest operations, and when the cart is filled, the tractor runs to mill (sugar factory).

There are many problems causing delays in milling on time. Delays in the billets milling has many reasons in agro-industry companies including sugar mill downtime, tractors failures in the waiting line, accident due to crowd of tractors in yard of the factory and shift changes, all these created quite long queues. The actual occurrence of a long queue in the queuing system is more due to sugar mill downtime. Furthermore, tractors repair takes more than a day due to the problems of acquiring replacement parts and the work complicated
in repair (Afsharnia et al., 2013; Afsharnia and Marzban, 2017). So, the amount of time when the tractor operators spend queuing is still very high. The majority of tractor operators in these agro-industries have contract with Sugarcane and Byproducts Development Company to attribute rights and responsibilities between both sides based on the amount of tonnage delivered to the mill. So, they have to wait in a queue prior to unloading their sugarcane at the mill. The tractors unload base on first-come-first-serve. The waiting time for each tractor may take up to 20 hours to complete the handling process, which cause significant increases in costs. These costs could not only include sucrose loss of sugarcane as a result of delay in milling, but also include opportunity cost, accidents costs and the cost of operators servings. Furthermore, under normal conditions, time delay in harvesting sugarcane slightly reduces the amount of commercial sugar mostly due to increase of cane dextran during harvesting to milling which in turn reduces the quality of produced sugar. Clarke (1991) observed that 0.04% of sucrose will decrease at 0.1% of dextran level. Sayed (1972) and later Legendre (1985) concluded that if the burned or cut cane is not crushed within several hours after harvesting, the available sucrose in cane scions will be inverted in the fractured parts of scions by breathing process. Furthermore, cane quality and amount of recoverable sugar could be reduced by loss of injured cells sap and attack of invertase microorganisms to scars.

Queuing system management is a critical component in any sector of business which can reduce the waiting time for operators that are in queues and increase the quality of service of the intended system. The appropriate management decreases the costs and maintains the quality of sugarcane billets. The objective of this study was to combine practical means for forecasting arrival and service level of the sugarcane carts to obtain the best model for the accurate time forecasting when the tractors operators spend in queuing of sugarcane delivering process to mill.

MATERIALS AND METHODS

Data Collection

Sugarcane has been widely cultivated agro-industry companies in Khuzestan province, Iran during the last decades. The average of cane harvested was 9000 hectares per year. These companies harvested cane between November and March when rainfall was less frequent and the plant's sugar content was at its highest level. There were 24 sugarcane harvester machine (Austoft 7000), 90 MF399 tractors and 90 carts in these companies. Sugar mill was located near the sugarcane farms at this companies site (Fig. 1). The average distance between the farthest sugarcane farms to the sugar mill was 13 km. Based on preliminary observations of these companies, the factory mill operated and crushed sugarcane billets at 24 hours per day where there was a feed elevator to convey billets for processing plant. The plant was experiencing heavy queues and therefore each cart has to wait in a considerable time for a service to begin. The queuing system in Sugarcane and Byproducts Development Company was a M/M/1 queue in which there was one server and one channel (Arifin et al., 2015). The required data related to arrival and service times of tractors was collected from October 2017 to May 2018 and comprised 56000 observations. Furthermore, factory breakdowns data including conveyors, mill, steam furnace and production hall failures were collected and derived from company records valid in the study region. Data analysis was performed according to the queuing models in the agro-industry unit (Arifin et al., 2015).

Model Description

ARMA (Autoregressive Moving Average) Process

The Autoregressive and Moving Average models (ARMA) often provide actual time series data (Fan and Fan, 2015). The ARMA model integrated a combination of AR and MA into the same equation which is the most general category of models. An ARMA model is merely a stationary series in level of the variable. If the series are not stationary, the Autoregressive Integrated Moving Average model (ARIMA) can be used.

The most general ARMA model simply combines AR and MA Eqs. where p is the order of the autoregressive part and q is the order of the moving average part (Chatfield, 2000) as follows:

\[
x_t = \delta + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}
\]

In this formula, \(\delta, \phi_1 \ldots \phi_p, \theta_1 \ldots \theta_q\) are fixed parameters, \(\phi_p x_{t-p}\) is the autoregressive (AR) part, \(\theta_q e_{t-q}\) is the moving average (MA) part and \(e_t \sim N(0, \gamma)\) is white noise with \(\gamma\) mutually independent for all \(t\) (Thiesson et al., 2004).
Box-Jenkins Methodology

This methodology consists of four phases (Gujarati and Damodar, 2003). For better clarification, the original equation and symbols have been remained unchanged which are as follows;

Model Identification

In modeling phase, it is of prime importance to check whether the series are stationary or not. This needs to be known since the estimation procedure is merely based on stationary series. Stationary in variance and means are two classes of stationary, stationary could be discerned by attending to the graph of the data, which are correlated either as partial correlation coefficient or autocorrelation structure. In cases in which the model is seemed to be non-stationary, differencing the series would be the solution. Some transformation modes such as Log transformation are able to get stationary in variance. In identification phase, obtaining the initial values of the order of non-seasonal and seasonal parameters (p, q, and P, Q) is the next step. One must look for significant autocorrelation and partial autocorrelation coefficients, to find numerical values of abovementioned parameters. It is essential that the sample coefficients of autocorrelation to be good estimates of the population autocorrelation coefficients. However, these estimations could be utilized as an initial guess of iterative steps. Another use of autocorrelation function is to find seasonal component. This happens when the coefficients of autocorrelation at lags between ‘t’ and ‘t-12’ are significant, otherwise, the coefficients will not have a significant difference with zero.

Model Estimation

At the previous phase, a few models were temporarily picked. Those models must provide satisfactory representation of the available data. Then, the precise estimates of the model parameters can be obtained by the least squares method which was described by Box and Jenkins (Box and Jenkins, 1976). In our study, data analysis were performed by Eviews 8 and XLSTAT 2016 to find the estimates of related parameters using repetitive procedures.

Diagnostic Checking

Diagnostic checking for model adequacy was the next step when we estimate the parameters by ARMA model. For adequacy model checking, the back forecasting method was used that described as the estimation of εᵢ for ε ∈ {-M,-M+1,−M+2,...,−2,−1} when the following conditions are jointly satisfied, and Autocorrelation Function (ACF) and Partial ACF (PACF) of residuals could be examined. If the residuals were random, the tentatively identified ARMA model would be adequate. When the residuals of ACF and PACF are random, all their ACF should be within the following range (Jadhav et al., 2017):

\[ \frac{1}{\sqrt{n-12}} \leq 1.96 \]

Ljung and Box ‘Q’ statistic could be used to test the autocorrelation of residuals. The autocorrelations of residuals should have a significant different with zero to get significant correlation in the residuals series. So, there is information left in the residuals which should be used in computing forecasts. This can be calculated as follows:

\[ Q = n(n + 2) = \sum_{k=1}^{h} (n - k)^{-1} r_k^2 \] (2)

In the above formula h is the considered as the largest lag, n is the observation number and rₖ is the ACF for lag k. The Q approximately follows a Chi-square distribution with (h-m) degree of freedom. The term ‘m’ is the number of parameters (p+q+P+Q). Under \( H_0 \) the statistic Q asymptotically follows a \( \chi^2_m \). For significance level α, the critical region for rejection of the hypothesis of randomness is:

\[ Q > \chi^2_{1-\alpha,m} \]

Selection of suitable forecast models was done by Akaike’s Information Criteria (AIC) and Schwarz Basic Criteria (SBC). In analyzing of time series AIC and SBC are being used as standard tools for model quality assessment. A variety of models were estimated and the one with lowest AIC and SBC was chosen as best model. The AIC enables us to determine the differencing order (d, D) which is required to obtain stationary and the proper number of AR and MA parameters. The AIC and SBC were calculated as Eqs. (3) and (4):

\[ AIC = n \left( \ln \left( \frac{\text{RSS}}{n} \right) + 1 \right) + 2m \] (3)

\[ SBC = \log \sigma^2 + \frac{m \log n}{n} \] (4)

In this formula, RSS is the estimated residual sum of squares, \( \sigma^2 \) is the estimated MSE, ‘n’ is the total number of observations and ‘m’ is the sum of estimated parameters (p+q+P+Q). Sometimes SBC is used instead of AIC (Jadhav et al., 2017).

Strength of a model to predict Ex-ante and Ex-post was verified by Mean Square Error (MSE) as Eq. (5) (Makridakis and Hibbon, 1979; Samarasinghe, 2007; Safa et al., 2015):

\[ MSE = \frac{\sum_{i=1}^{n} (\hat{Y}_i - y_i)^2}{n} \] (5)

In this formula \( Y_i \) is the actual output for the \( i^{th} \) value and \( \hat{Y}_i \) is the predicted output for the \( i^{th} \) value, and N is total number of observation.

Root Mean Square Error (RMSE) calculated as another error estimation index using Eq. (6) (Suresh and Priya, 2011):

\[ RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{Y}_i - y_i)^2}{n}} \] (6)

Mean Absolute Percentage Error (MAPE) makes use of all observations and is easy to grasp notation, since it is used most widely for reporting (Farjam et al., 2014; Jadhav et al., 2017). It is defined by the formula (7):

\[ MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100 \] (7)
where, $Y_t$ and $\hat{Y}_t$ are actual and predicted values, respectively.

In our study we used the above three test to verify our models. Finally, 70% of data was used for in-sample modeling and 30% of out-of-sample was considered for validation test.

RESULTS AND DISCUSSION

For arrival level changes, as observed on Fig. 2, the behavior of the in-process arrival prior to a sugar factory was provided by observing the queue at the end of each of the three shift work system. In a steady state condition, the trend of queuing system could vary about a mean level of 16.35 carts and standard deviation of 5.73 carts. Furthermore, according to Fig.2, the mean levels of the first, second and third shifts were 14.7, 16.72 and 19.33 carts, respectively. The arrival level of carts in third shift was higher as compared to other shifts. Out of 24 sugarcane harvester in this company, four or five sugarcane harvesters run at the night, since the arrival level of carts was minimum at the first shift. During every day, the arrival level was gradually increased by increasing the number of active harvesters and consequently increased the amount of harvested crops.

Fig. 3 shows the service level for sugar factory obtained by sampling at the end of each shift with a mean level of 16.82 carts and standard deviation of 7.39 carts. Moreover, according to Fig. 3 the mean service levels of the first, second and third shifts were 13.75, 17.37 and 18.26 carts, respectively.

Therefore, it is reasonable to state that the service level was almost agrees with arrival level. Nevertheless, up to 81.6% of carts must wait for more than 30 minutes to discharge billets on feed conveyor. Fluctuation shewing in Fig. 4 indicates that waiting time is not related to the number of carts in queue, but it is usually
due to breakdowns that occurred in sugar factory equipment. Poor maintenance of the systems can result in the equipment suddenly going dead. Maximum failure of the factory components were related to mill (73%) and feed conveyor (20%), respectively. Only the 25% of total failures were predictable by maintenance experts and the factory scheduled maintenance was only performed for these failures. So, there was not a specific schedule for 75% of other breakdowns that were so unpredictable. This has an impact on the costs and time of agro-industrial units (Afsharnia and Marzban, 2019). The waiting time of carts ranged from 5 to 763 minutes (Fig. 5). The highest mean waiting time of 206.1 minutes per day was observed in the sugar factory yard. The number of carts arriving ranged between 378 and 496 per day, resulted the total unload amounted to 6500 tons day\(^{-1}\). On an average, each cart should wait for 105.6 minutes in discharging queue. Delay in the milling process causes sucrose content degradation as well as dextran, ethanol and oligo saccharides formation in burned or cut canes (Solomon, 2000). Sucrose content of canes lose by 0.0135% per one hour delay (Noroozi, et al., 2015). Accordingly, the total of staled sugarcane would be equal to 1.5 tons per day.

Fig. 4. Waiting time change trend based on number of carts in the queuing system

Fig. 5. Waiting time for carts in the queuing system

Fig. 6. The queue length of tractor drawn carts in the system

A glance of the ACF and PACF reveals no evidence of seasonality in the data. The seasonal part of an AR or MA model will be seen in the seasonal lags of the PACF and ACF. Spikes in the ACF (at low lags) indicated non-seasonal MA terms. Spikes in the PACF (at low lags) indicated possible non-seasonal AR terms.

The results of the ADF unit root test is summarized in Table 1. Without differencing, the time series do not have trend or seasonal effects. The hourly arrival and service level dataset were stationary. According to ACF and PACF, the models were identified for the arrival and service level series. Table 2 gives the results of models identified for arrival and service level forecast of tractor drawn carts with their respective Akaike's Information Criteria (AIC) and Q statistics. It was found that the models ARMA (4,3) for arrival level and ARMA (4,2) for service level were good fit with the lowest AIC and SBC values to the rest models. After a repetitious process, the parameters related to the best models were estimated (Table 3).
The significance of fitted ARMA models indicated a good fit. Therefore, these coefficients in Table 3 can be used for predicting arrival and service levels. The back forecasting method was used to obtain residuals for checking the model adequacy of the best selected models. The ACF and PACF plots of the residuals for the selected models revealed that there was no autocorrelation between residuals. The Ljung-Box \( Q \) statistic was used for model adequacy checking. Non-significant \( Q \)-statistic for testing the null hypothesis in which autocorrelations up to lag \( k \) equal zero, indicated white noise of series. Therefore, these tests demonstrate that ARMA(4,3) and ARIMA (4,2) models were adequate for forecasting levels of arrival and service, respectively.

The behavior of Ex-ante and Ex-post forecasts related to arrival and service level of tractor drawn carts are illustrated in Figure 8. The relatively lower values of MSE, RMSE and MAPE compared to other models denoted validity of ex-ante that predicts corresponding to arrival and service level of the tractor drawn carts. There is no study having similar variables and statistical models to our study. The results of analyzing, actual and predicted arrival levels (MSE=4.6, RMSE= 2.14, MAPE= 3.86) revealed that the amplitude of graph fluctuation is generally very high at the first hours of harvesting and then reaches the mean interval level (Fig. 8). At the initial hours of sugarcane harvesting, operation was often irregular due to the low speed at beginning of the work due to calibrating the sugarcane harvesters which did not run on the fields at the last night. The trend of service level graph is illustrated in the Fig. 9 (MSE= 6.6, RMSE= 2.57, MAPE= 7.21).

### Table 1. The stationary survey of carts arrival and service level

<table>
<thead>
<tr>
<th>Test type</th>
<th>variable</th>
<th>Stationary degree</th>
<th>Critical value</th>
<th>Calculated value</th>
<th>Test result</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF</td>
<td>Arrival level</td>
<td>I(0)</td>
<td>-2.87</td>
<td>-6.78</td>
<td>Stationary</td>
</tr>
<tr>
<td>ADF</td>
<td>Service level</td>
<td>I(0)</td>
<td>-2.87</td>
<td>-8.40</td>
<td>Stationary</td>
</tr>
</tbody>
</table>

### Table 2. Models identified for arrival and service level forecast of tractor drawn carts

<table>
<thead>
<tr>
<th>Models</th>
<th>Arrival level</th>
<th>Service level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q-Stat</td>
<td>df</td>
</tr>
</tbody>
</table>

Fig. 7. Autocorrelation and partial autocorrelation function curve of carts arrival and service level
Table 3. Summary of the statistical parameters of the best fitted ARIMA models on waiting time

<table>
<thead>
<tr>
<th>Arrival level</th>
<th>Service level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AR1</td>
<td>1.48</td>
<td>0.06</td>
</tr>
<tr>
<td>AR2</td>
<td>-0.76</td>
<td>0.12</td>
</tr>
<tr>
<td>AR3</td>
<td>-0.16</td>
<td>0.12</td>
</tr>
<tr>
<td>AR4</td>
<td>0.16</td>
<td>0.07</td>
</tr>
<tr>
<td>MA1</td>
<td>-1.03</td>
<td>0.09</td>
</tr>
<tr>
<td>MA2</td>
<td>0.32</td>
<td>0.14</td>
</tr>
<tr>
<td>MA3</td>
<td>0.45</td>
<td>0.09</td>
</tr>
<tr>
<td>Constant</td>
<td>17.13</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Fig. 8. Actual and predicted values of arrival level
The curve in Fig. 9 illustrated that instability was retained and oscillations occurred within service interval time. The unloading of tractor drawn carts terminates in sugarcane factory equipment failure time. In other words, when the factory equipment brake down, the delivery of sugarcane billets to sugar factory was stopped. The service level of tractor drawn carts attained higher level after 20 hours at 34 cart per hour and reached the lowest service level of 0 cart per hour after 16 hours. It is clear that the graph oscillations of service level were higher compared to the arrival level. According to failure data, the most problems refers to sugar factory breakdowns. So, the failures of harvesters as well as tractors which used to draw the carts can be ignored.

CONCLUSIONS

The predicted arrival and service level of tractor drawn carts were nearly identical to actual values with very good validation as demonstrated by the relatively low values of MSE, RMSE and MAPE. Therefore, the ARMA model serves as a reliable method to predict the magnitude of arrival and service level. The results of models fit demonstrated that ARMA (4,3) and ARMA (4,2) models were suitable for estimating arrival level and service level of tractor drawn carts, respectively. The trends of predicted values of arrival and service level truly reflected the actual values as well as queue system tendency. The identified models which can be used to forecast values for arrival and service level of tractor drawn carts are developed for subsequent hours at harvest season.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Iran's National Elites Foundation for financial support related to Agricultural and Biological System Analysis research core of Agricultural Sciences and Natural Resources University of Khuzestan.

REFERENCES


فاطمه افشازنا، افسین مرزبان، عباس عبد شاهی

گروه مهندسی مالی‌های کشاورزی و مکانیاسیون، دانشگاه علوم کشاورزی و منابع طبیعی خوزستان، اهواز، ج. ایران.

گروه اقتصاد کشاورزی، دانشگاه علوم کشاورزی و منابع طبیعی خوزستان، اهواز، ج. ایران.

چکیده- نیشکر یکی از محصولات به شدت فاسدپذیر است که به صورت مواج خام برای تولید شکر سفید کار می‌رود. این منبع تجاری ساکرز بدلیل سواراندن قبل از بردانش، دمای بالای محيط، تأخیر در اسباب و برنامه‌ریزی الوداگی‌هاي میکروبی و سرعت دچارات کمیته‌ی میوه، یکی از مهم‌ترین ریسک‌های حمل و نقل نیشکر تأثیرگذار است. که در این فاصله می‌توانند کمیته و کمیته محصول را تحت تأثیر قرار دهد. تأخیر در اسباب گردن نیشکر برداشت شده به دلایل مختلفی از جمله خرابی کارخانه، خرابی تراکتورها در صفحه قدرت تراکتورها در محیط کارخانه و نگهداری شیفت در کشت و صمت‌ها بوجود می‌آید که سبب اجبار صرف طولانی می‌گردد. از این رو، در این پژوهش تلاش گردید به پیشنهادی سیستم فحش تولید محصول نیشکر به کارخانه تولید شکر با استفاده از سری‌های زمانی برداخته شود تا زمینه‌های بهبادی آن را فراموش کند. مدل ARMA جهت پیش‌بینی نرخ ورود و نرخ سروبی‌خ تراکتورهای حمل نیشکر کار گرفته شد و شاخص‌های MAPE و RMSE جهت ارزیابی دقت پیش‌بینی استفاده شدند. نتایج برازش مدل‌ها نشان داد که بهترین مدل‌های ARMA(4,2) برای نرخ ورود و نرخ سروبی‌خ تراکتورهای حمل بی‌سبک بودند. روند مقادیر پیش‌بینی نیشکر به کار گرفته شد. مدل‌های ARMA(4,2) برای نرخ ورود و نرخ سروبی‌خ به خوبی بر مقادیر واقعی منطبق بود. یافته‌ها، مقادیر پیش‌بینی شده را می‌توان برای قیمت چاپ در دست نیشکر کار بر داد و از نتایج‌ها بوجود آمده که موجب ضایع شدن مقادیر زیادی از محصول می‌شود کاست.

اطلاعات مقاله

تاریخچه مقاله:
تاریخ دریافت: 1397/7/28
تاریخ پذیرش: 1398/2/9
تاریخ دسترسی: 1398/3/5

واژه‌های کلیدی:
ARMA
تأخیر در اسباب گردن
نیشکر
حمل و نقل