

## A comparison study between time series model and ARIMA model for sales forecasting of distributor in plastic industry

Patimaporn Udom<sup>1,a</sup>, Naragain Phumchusri<sup>1,b</sup>

<sup>2</sup> Department of Industrial Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand

**Abstract:** - The number of distributors in plastic industry is increasing every year. Therefore, they got a problem about competitive cost more than in the past. One of the ways to solve this problem is management by inventory. Before they improve inventory, the forecasting is important for this process. In the past, they were looking for conventional time series for forecasting sales volume before manages inventory. This study would like to compare the application of three forecasting methods on the amount of the sales volume for plastic distributor, the ARIMA time series method, Moving average method and Holt's and Winter exponential method. After applying five data sets of raw material from plastic distributor by four forecasting methods, the ARIMA model shows better results than other models when compare with other methods by using MAPE (Mean Absolute Percentage Error).

**Keywords:** - Forecasting, ARIMA, Inventory, Naive.

### I. INTRODUCTION

Thermoformed plastics are plastics used in the process of injection, extrusion and thermoforming. The regularly used polymers for the process include acrylonitrile butadiene styrene (ABS), high impact polystyrene (HIPS), polypropylene (PP) and polymethyl methacrylate (PMMA) among others. These plastics have applications in automotive part, electronics part, food packaging and consumer product. The main requirement for the plastic industry is a distributor to provided plastic to plastic factory.

The plastics industry shapes the world we live in today whether it is industrial, technological or commodities used on a regular basis. Quality market analysis, forecasts and trends determined from key market drivers help shape the future of the plastics industry. The current domestic market for distributor business is competitive rather high when compare with 5-6 years ago. Business and dealer happens quite a lot in the plastic industry. In the past, plastic distributor also have a small amount so they can add more margin to cover their cost but currently we have plastic distributor more than 40 company in Thailand market. And the way to keep their market has reduced prices and to analyze ways to reduce costs. Forecasting, it is the beginning of the administration in the other due to forecasting analysis will help them to understand the emerging technologies and applications by key market.

For forecasting process, it purposes to find the best model with historical data. Nevertheless, uncertainty of data collection will be occurring due to the error caused in time lag and effect between variables. However, from the traditional model, most of forecasting methods usually use a statistical model as a tool for forecasting future data. In industry segment, usually use normal forecasting model to forecast future data such as simple moving average, Exponential smoothing and Holt's and Winter exponential, etc. For example, Suwadee [2] find forecasting model for automotive part and they found Regression suitable for automotive part data.

In recent years, due to the innovation of forecasting methods and improvement of forecasting accuracy, the forecasting methodology becomes indispensable for further decision making process in both industry and government. For example, Zadeh [3] defines a fuzzy algorithm is an ordered set of fuzzy instructions which upon execution yield an approximate solution to a specified problem. The concept is applied to different fields of data to solve the problems. Currently, the concept of fuzzy logic is successfully to analysis forecasting methods. Volkan S. Edigera and Sertac Akar [4] used ARIMA method to forecast of primary energy demand by fuel in Turkey. The ARIMA forecasting of the total primary energy demand appears to be more reliable than the summation of the individual forecasts. Chi-Chen Wang [5] study forecasting model by comparison between fuzzy time series model and ARIMA model for forecasting Taiwan export. When the sample period is shorter with only a small set of data available, the fuzzy time series models can be utilized to predict export values accurately, outperforming the ARIMA model. James W. Taylor [6] found the results indicate strong potential for the use of seasonal ARIMA modeling and the extension of Holt-Winters for predicting up to about two to three days ahead and that, for longer lead times, a simplistic historical average is difficult to beat.

However, never saw the research pay attention to discuss the forecasting model for distributor in plastic application although plastic fields. Besides, previous researches only compute the forecasting data using a general forecasting model to forecast future data for plastic fields, and study new model of forecasting for

another fields except plastic or plastic distributor as well. This research attempts to use information of sales values from plastic pellets distributor as an example to test whether the ARIMA is indeed practical in its forecast of variables. By comparing the time series as Naïve method, Moving average and Winter methods by use MAPE value to decision. Furthermore, this paper intent to identify which model is suitable to provide information in decision making for plastic pellets distributor.

## II. METHODOLOGY

### 2.1 Time series

Time series is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals. Examples of time series is the annual flow volume of the rain-fall and show frequently plotted via line charts. Time series analysis comprises methods for analyzing time series data in order to extract parameter statistics and other characteristics of the data to predict future values based on previously observed values. While regression analysis is often employed in such a way as to test theories that the current values of one or more independent time series affect the current value of another time series, this type of analysis of time series is not called "time series analysis", which focuses on comparing values of time series at different points in time. A stochastic model for a time series will generally reflect the fact that observations close together in time will be more closely related than observations further apart. In addition, time series models will often make use of the natural one-way ordering of time so that values for a given period will be expressed as deriving in some way from past values, rather than from future values. [7]

(1) Moving average model

$$F_{t+1} = \frac{(A_t + A_{t-1} + A_{t-2} + \dots + A_{t-n+1})}{n}$$

$A_t$  = Data at time t

$F_{t+1}$  = Forecast value at time t+1

n = Number of data

(2) Holt's and Winter exponential model

$$S_t = \alpha \frac{A_t}{I_{t-L}} + [(1 - \alpha)(S_{t-1} + T_{t-1})]$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1}$$

$$I_t = \gamma \frac{A_t}{S_t} + (1 - \gamma)I_{t-L}$$

$$F_{t+p} = (S_t + T_t p)I_{t-L+p}$$

$S_t$  = Smoothing value

$\alpha$  = Constant of smoothing value ( $0 < \alpha < 1$ )

$A_t$  = Actual value at time t

$\beta$  = Constant of trend value ( $0 < \beta < 1$ )

$T_t$  = Trend parameter at time t

$\gamma$  = Constant of seasonal value ( $0 < \gamma < 1$ )

$I_t$  = Seasonal parameter at time t

p = Number of forecasting

L = Lead time

$F_{t+p}$  = Forecasting value at time p

### 2.2 Auto Regression Investigate Moving Average (ARIMA) model

The model was propose by Box–Jenkins in 1970, which examines each variable by using auto-regression, AR (P) and Moving Average, MA (Q) to investigate the historical data and economic fluctuations. The algorithm presented as follows: [8]

- (1) The first step in developing ARIMA model: determine if the series is stationary and if there is any significant seasonality that needs to be modeled. The autocorrelation functions (ACF) or partial autocorrelation functions (PACF) are used to define the distribution of sample data. If data set are not stationary, differencing to achieve stationary.
- (2) Identification: identifying the phase of the series by using autocorrelation functions (ACF) and partial autocorrelation function (PACF).
- (3) Estimation: the good conditional and exact likelihood are used to estimate the parameters.
- (4) Diagnostic check: the process of diagnostic check involves testing the assumptions of the model to identify any areas where the model is not enough. The statistical identification process includes parameters statistical significance and the residual. If the model is found to be out of control, it is necessary to calculate and repeat Step (4) until a better model is identified..

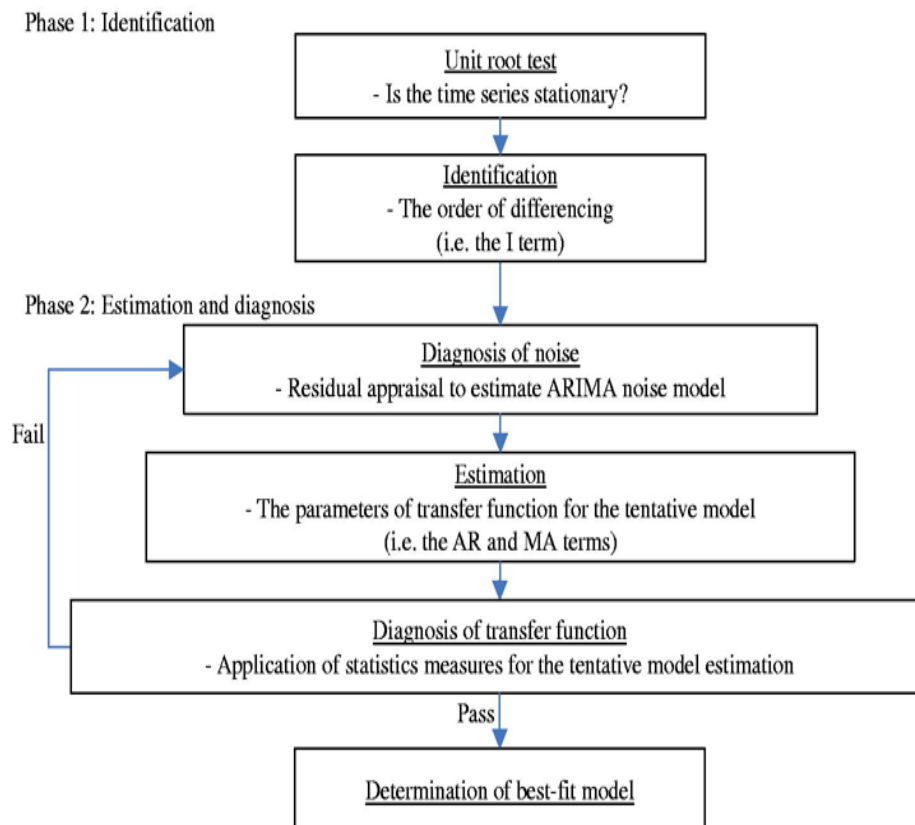


Fig. 1 ARIMA flow diagram [8]

The accuracy of the fitted model is measure by the mean average percentage error (MAPE), which is a measure of accuracy of the fitted model. The mean average percentage error (MAPE) is not very informative by itself, but it can be used to compare fits of different models that we use in this paper. For all measures, smaller values generally indicate a better fitting model for quantity of sale for plastic distrutor. [5]

### III. RESULT AND DISCUSSION

#### 3.1 Data preparing

The selected data were the amount of plastic sale from plastic distributor in Thailand during the time period between January 2004 and December 2012. A total of data is 108 months and 90 months for training and 18 months for testing. Plastics distributor have many items in their stock therefore we should select data by ABC analysis. It's showed in the table 1. We chose data by use summary value of each product items and we select product items have summary value more than 75% from all of product items. We got five items to find forecasting model for suitable with each data.

Table 1 ABC analysis for plastic distributor company

Products Code	Annual Sales (kg)	Average Unit cost per kg (25/2/13)	Usage THB per year	
1101	34,657,528.19	52.3	1,812,588,724.09	50%
2301	3,123,480.51	112.25	350,610,686.87	59%
2101	1,053,726.58	178.86	188,469,536.10	64%
1202	4,150,138.67	36.45	151,272,554.40	68%
2302	1,374,427.77	85.25	117,169,967.68	72%
1199	1,201,626.84	89.141	107,114,218.14	75%
1299	1,021,221.07	98.65	100,743,458.23	77%
1108	793,589.45	121.15	96,143,362.43	80%
2204	550,545.74	167.65	92,298,993.09	82%
2104	1,024,882.99	79.88	81,867,653.24	85%
2105	433,083.28	188.54	81,653,521.61	87%
2205	347,303.73	209.89	72,895,580.59	89%
2207	269,843.77	241.1	65,059,333.75	91%
2304	197,743.07	302.8	59,876,602.20	92%
2399	223,825.33	245.45	54,937,928.07	94%
2303	251,411.67	174.5	43,871,335.83	95%
2202	458,716.97	85.56	39,247,824.24	96%
2201	387,399.35	67.24	26,048,732.34	97%
1105	298,627.13	84.5	25,233,992.77	97%
1102	378,945.79	49.86	18,894,236.92	98%
2305	153,786.20	98.24	15,107,956.29	98%
1201	436,850.00	34.26	14,966,481.00	99%
2102	77,951.73	145.6	11,349,772.37	99%
1203	186,436.47	45.12	8,412,013.38	99%
2103	41,159.53	201.45	8,291,587.45	100%
1107	36,254.93	102.54	3,717,580.86	100%
1103	66,455.67	51.52	3,423,795.95	100%
2206	15,453.52	158.65	2,451,700.95	100%
3102	7,039.46	345.5	2,432,134.81	100%
1104	64,280.00	36.2	2,326,936.00	100%
2203	9,208.03	90.15	830,103.60	100%
			3,659,308,305.26	

**3.2 Unit root test**

To determine ARIMA (p,d,q)(P,D,Q)<sub>x</sub> model, we have to test review and trend stationary by use ACF and PACF as showed in Fig 3. We can analyze by calculate data at time t and t+1, if the data a bit different from 1 so we have to re-calculate. Moreover, ACF and PACF provide a statistical summary at a lag. The maximum number of lags is determined by dividing number of actual data by 4, for a series with less than 240 based on Box and Jenkins method. The result of ACF and PACF of product item 1101 were not stationary so we have to differentiated data 1 time can see in Fig 4. If the data not stationary we can not forecast by ARIMA because forecast value have an error from data. From Fig.2 standard deviation of product item 1101 is 595,792.62 and mean is 2,888,127.3. Example for unit root test for product item 1101 showed as below.

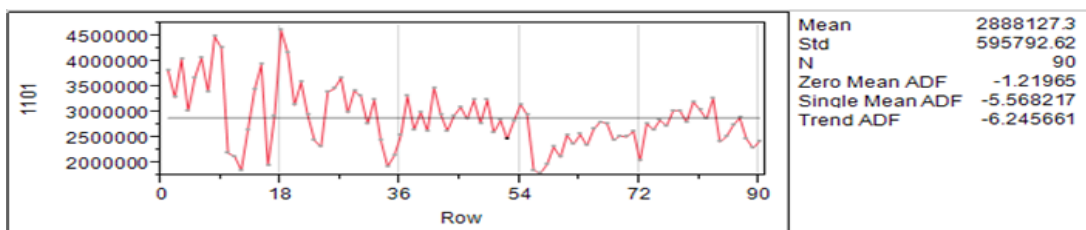


Fig. 2 Time series of item 1101

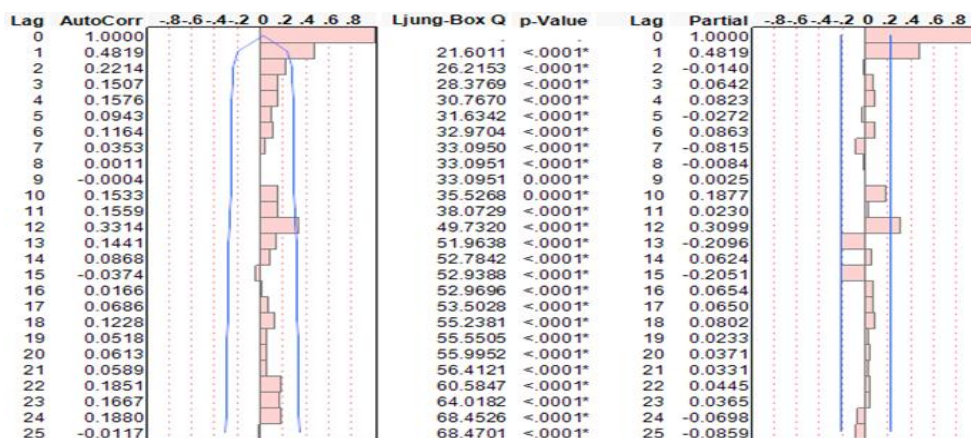


Fig. 3 ACF and PACF graph of item 1101

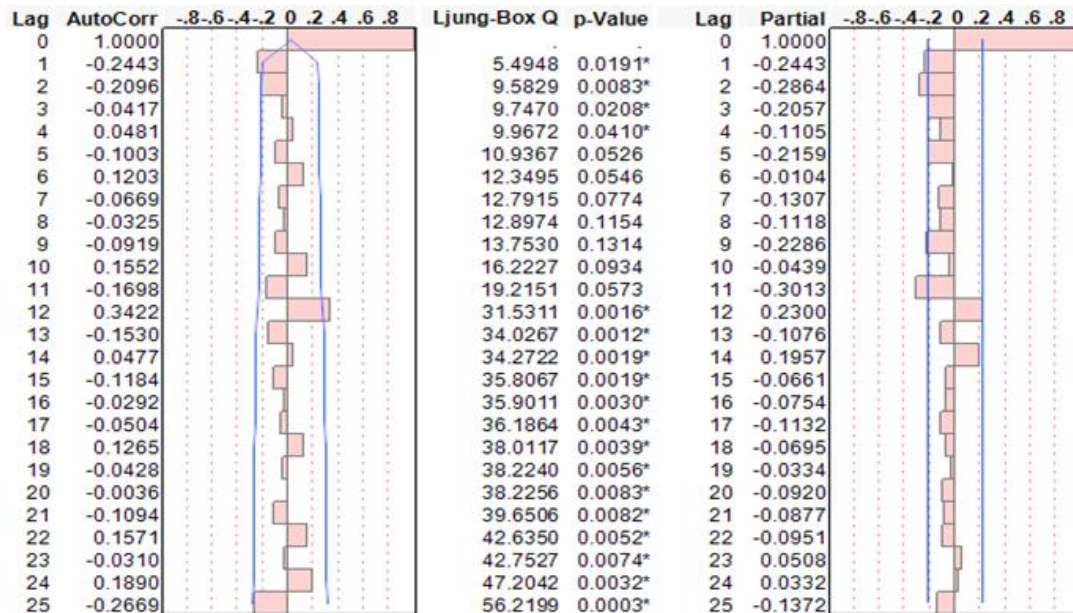


Fig. 4 ACF and PACF graph of item 1101 (difference 1 time)

### 3.3 Estimation Parameter

When we get the stationary data set from unit root test we will find the parameter for ARIMA model. From Fig 4. We get the stationary data for product item 1101 then identify model to get the parameter and we get model ARIMA is ARIMA (1,0,1)(1,0,1)<sub>12</sub>. Parameter estimate from this model can show as Fig 5. The ARIMA model has been originated from the autoregressive model (AR), the moving average model (MA) and the combination of the AR and MA, the ARMA models. From Fig 5 show estimate parameter for product item 1101 and have seasonal for ARIMA as AR1(1) = 0.147, AR2(12) = 0.854, MA1(1) = -0.524, MA2(12) = 0.455 and B = 2,937,546 and can show the formular as below.

$$(0.147)(B)(0.854)(B^{12})(1 - B)^0(1 - B^{12})^1 Y_t = 366,079.238 + (-0.524)(B)(0.455)(B^{12})a_t$$

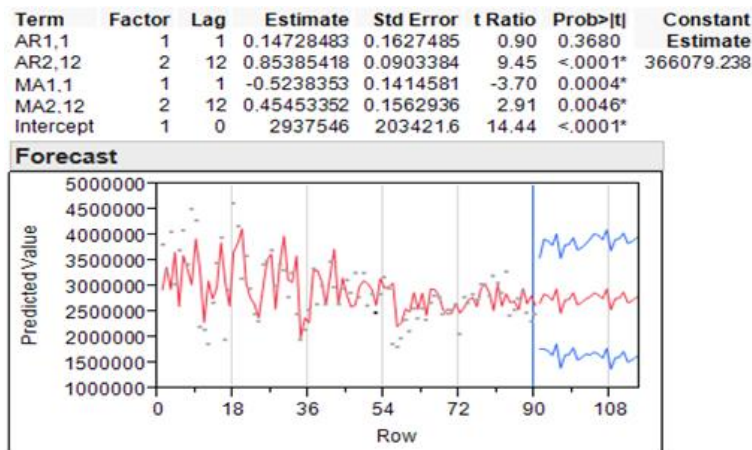


Fig. 5 Parameter estimate and forecasting of item 1101

### 3.4 Diagnostic check

In this case where have seasonal components are included in the ARIMA model, the model is called as the Seasonal ARIMA or SARIMA. The ARIMA forecasting gives the results in different options which are upper limits, lower limits, and forecasted values. Upper and lower limits provide a confidence interval of 95%, in other words any realization within the confidence limits will be acceptable. The process of diagnostic check for identification process includes parameters statistical significance can analyze by the residual in Fig 6 and 7. If the residual is random or upper and lower limits have a confidence interval more than 95%, its mean parameter estimate from model can accept.

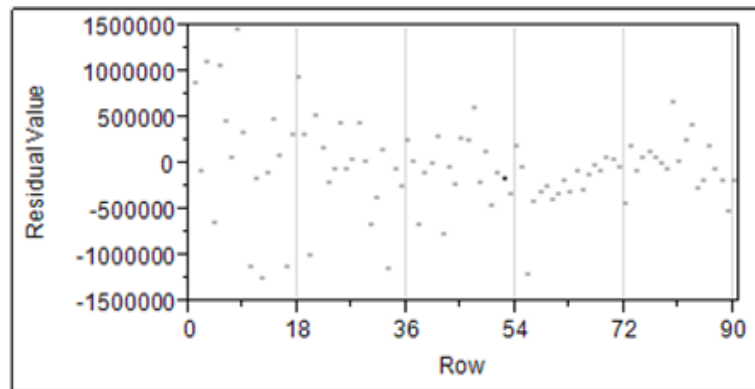


Fig. 6 Residual from ARIMA (1,0,1)(1,0,1)12 of item 1101

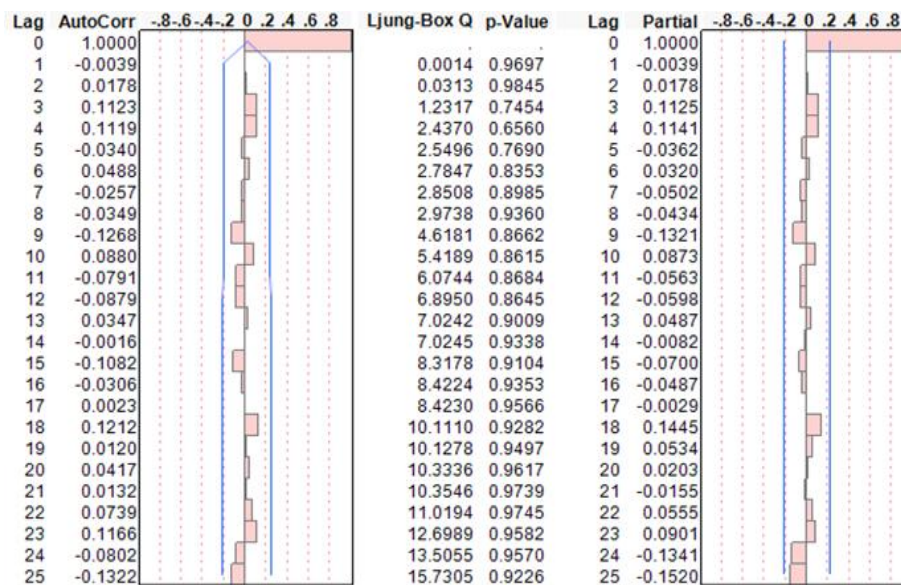


Fig. 7 ACF and PACF from residual of ARIMA (1,0,1)(1,0,1)12 of item 1101

After test all of product items that we chose 5 items then will define mean average percentage error (MAPE) for each model and each item. The result can show in table 2. We find the best model from four model for forecasting sale volume of plastic distributor in Thailand is ARIMA.

Table 2 MAPE from four forecasting model of the result

Items	MAPE Value (%)			
	Naive	Moving Average	Winter Exponential	ARIMA
1101	16.37	15.15	15.17	12.32
2101	16.76	15.46	16.46	14.89
2301	40.42	37.97	43.24	33.44
1202	22.73	22.98	22.66	21.54
2302	43.23	39.46	48.53	37.64

#### IV. CONCLUSION

The researcher tries to find a way to forecast more accurately than the old method. The five data sets were collected from the sales volume from one of distributor for plastic industry. Then, the four models were used to apply into five of the raw data sets in order to find suitable forecasting model which get better error more than currently model. The results showed that the ARIMA model get the best MAPE (Mean Absolute Percentage Error) which is the measurement of this study.

### REFERENCES

- [1] S. Samuthananon, "Inventory management system development in spare-part business", Engineering management, Chulalongkorn University, 2007.
- [2] M.Z. Babai, M.M. Ali and J.E. Boylan, "Forecasting and inventory performance in a two-stage supply chain with ARIMA (0,1,1) demand: Theory and empirical analysis," *International Journal of Production Economics*, 2011.
- [3] J. W. Taylor, "A comparison of Univariate Time Series Methods for Forecasting Intraday Arrivals at a Call Center," *Management Science*, vol. 54, pp. 253–265, 2008.
- [4] L. Bianchi, J. Jarrett and R. Choudary Hanumara, "Improving forecasting for telemarketing centers by ARIMA modeling with intervention," *International Journal of Forecasting* 14, pp. 497–504, 1998.
- [5] W. A. NG, "A simple classifier for multiple criteria ABC analysis," *European Journal of Operational Research*, vol177, pp. 344–353, 2007.
- [6] Chi-Chen Wang. A comparison study between fuzzy time series model and ARIMA model. Department of Financial Management National Defense University. *Expert Systems with Applications* 38,pp. 9296-9304, 2011.
- [7] Lisa Bianchi, Jefft Jarrett and R. Choudary Hanumara. Improving forecasting for telemarketing centers by ARIMA modeling with intervention. *International Journal of Forecasting* 14, pp. 497-504, 1998.