

USING ARIMA MODELS FOR TURNOVER PREDICTION IN INVESTMENT PROJECT APPRAISAL

Zoran Petrović *

Tecon System d.o.o, Belgrade, Serbia

Dr Uglješa Bugarić

University of Belgrade, Faculty of Mechanical Engineering, Belgrade, Serbia

Dr Dušan Petrović

University of Belgrade, Faculty of Mechanical Engineering, Belgrade, Serbia

In the contemporary investment project analyses, most critical point is how to estimate daily turnover of production, or service, based system. In order to make prediction, for investment in certain type of equipment more accurate, daily turnover in the system for automated car wash was observed, along with weather conditions. According to observation, ARIMA model for daily turnover and weather condition is created, according to Box-Jenkins procedure. Conclusion was made that daily turnover can be analytically expressed through daily weather conditions. Validity of observation is checked on second system that is installed in different town in Serbia. According to compared results, conclusion was made that ARIMA model of system daily turnover, predicted by dependent variable, can be generally used as good predictor in investment analyses, or selective criteria for investment decisions.

Key words: Arima, Box-Jenkins, Investments, Predictions, Turnover

INTRODUCTION

Life cycle of project is determined by, at least four phases. In the first phase, also called initial phase, feasibility study is performed and decision for continuing or cancelation of the project, (if not feasible), is made. In the production systems if project is evaluated as feasible, other phases can be carried out (planning and construction, production and final operational phase), Figure 1.

From the Figure 1, it is hard to conclude that first phase of the project is most important one. Conclusions from the first phase will have vital influence to the project successful finishing [03].

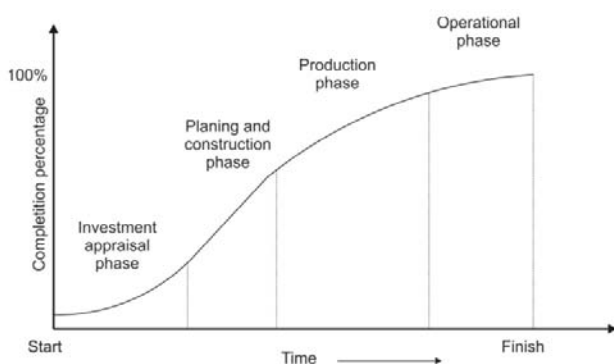


Figure 1: Project cycle phases [04]

Considering risk distribution across project life cycle, beginning of the project is bearing most of the risk, since in the beginning number of available information is relatively small.

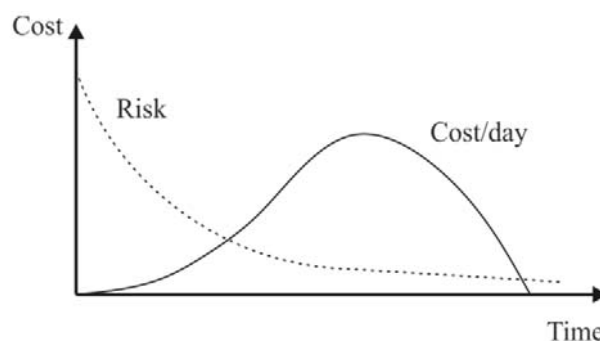


Figure 2: Diagram of project life cycle considering project expenditure and risk [03]

Usually, in investment analyses average daily turnover, measured on existing system is used to estimate turnover for system that is analyzed from the point of investment appraisal. Problem in such approximation is that fluctuation of average daily turnover of the system that is analyzed as potential investment, can differ significantly from any other, previously installed system.

In order to predict daily turnover, two variables were measured on existing system. First variable is Daily turnover and second one is Daily weather condition.

HYPOTHESIS

Hypothesis 1.: H0: Daily turnover is dependent from Daily weather condition.

Hypothesis 2.: H0: Measure of dependence can be expressed in analytical form.

Hypothesis 3.: H0: Daily weather condition can be used as predictor for estimation of daily turnover in investment analyses.

METHODOLOGY

Automated car wash was observed in such way that daily records of turnovers and weather conditions were registered in research protocol. Case study was created from daily records taken from 25.05.2009. until 21.5.2010. During mentioned time 362 data records were taken.

From research protocol two variables were created – Daily turnover and Daily weather condition. Daily turnovers are represented as values in RSD and Daily weather conditions are represented as one of four different, weather conditions. If weather during most of the day was sunny, number 1 was assigned for Daily weather condition. If weather during most of the day was cloudy, number 2 was assigned. To mostly cloudy with rainy period was assigned number 3 and to the rainy, or snowy, day was assigned number 4.

Variables defined in research were modeled as ARIMA time series according to Box-Jenkins modeling strategy [01]. Such variables are used for forecasting of turnover on the system [02]. All calculations were done in statistical software IBM SPSS Statistics 20.

Dependency between those variables was examined in order to check if daily turnover depends from daily weather condition. In the results, after confirmation of the Hypothesis 1, analytical model of dependency was created and Hypothesis 2 was confirmed.

For confirming Hypothesis 3, new system in different town was observed. Observation period was defined as 100 days, on which both variables (Daily turnovers and Daily weather conditions) were recorded and registered in research protocol. Based on the analytic expression of the dependency between variables in first case, daily turnover on second system was modeled, using daily weather conditions as predictor.

Data for both variables were tested against Hypothesis that they can fit to the Normal distribution. Goodness of fit for both variables was tested with Kolmogorov-Smirnov test, with significance level $\alpha = 0,05$.

Students T test was used in order to test both variables against Hypothesis 3, that there is no significant differences between distribution of two samples.

RESULTS

According to proposed methodology time series of daily turnover is given on the Figure 3:

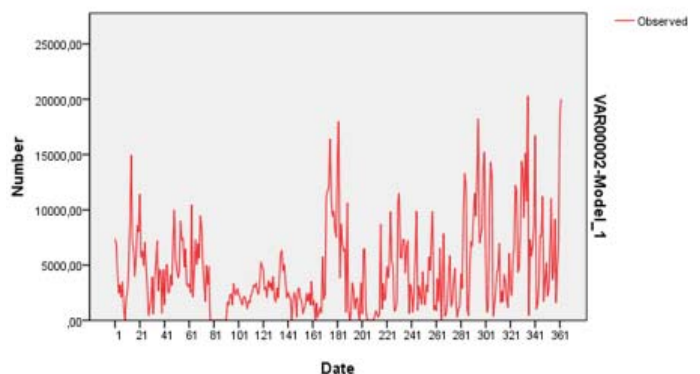


Figure 3: Observed daily turnover

Time series of measured data compared to time series of modeled data using ARIMA methodology is given on the Figure 4.

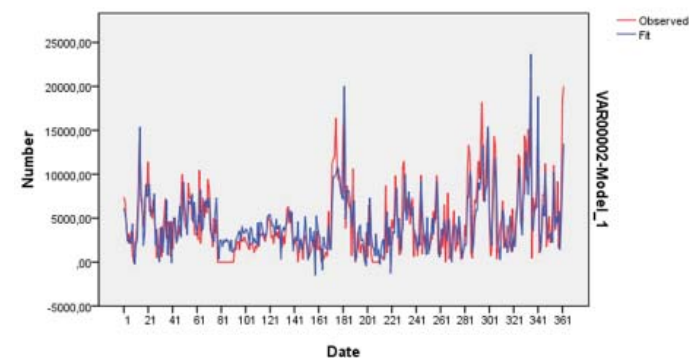


Figure 4: Observed and modeled daily turnover

According to the calculated data best fitted model that describes analytical relationship between Daily turnover and Daily weather condition was: ARIMA (1,0,14). Parameters of ARIMA (1,0,14) model are given in the Table 1.

Analytical expression for relationship between variables is given as:

$$Day_n = 8447,649 + 0,696 \cdot Day_{n-1} - 0,296 \cdot \sum_{k=n-13}^n \frac{Day_k}{n-1} - 1789,317 W_c$$

where:

$\frac{Day_k}{n-1}$ is average turnover for past 14 days.
 n is number of observation $n \geq 1$.

Wc is daily weather condition on observed day as categorical variable that can take value 1-4.

Table 1: ARIMA model paramters

ARIMA Model Parameters					Estimate	SE	t	Sig.
VAR02-Model_1	VAR02	No Transf.	Constant		8447,649	557,341	15,157	,000
			AR	Lag 1	,696	,041	17,070	,000
	VAR01	No Transf.	MA	Lag 14	-,265	,055	-5,431	,000
			Num.	Lag 0	-1789,317	115,705	-15,464	,000

According to Ljung-Box procedure, statistical model is good fitted to observed data. Significance of the test was 0,081, so difference be-

tween fitted and measured values are not statistically significant. Values of the Ljung –Box test for goodness of fit are given in the Table 2.

Table 2: Model goodnes of fit

Model	Number of Predictors	Model Fit statistics		Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared		Statistics	DF	Sig.	R-squared
VAR02-Model_1	1	,609		23,158	15	,081	6

Table 3: Extreme values (outliers)

			Estimate	SE	t	Sig.
VAR02-Model_1	14	Additive	7889,826	2094,530	3,767	,000
	182	Additive	10126,487	1996,093	5,073	,000
	300	Additive	10890,786	2045,844	5,323	,000
	335	Additive	13578,128	2011,825	6,749	,000
	341	Additive	10432,565	2030,952	5,137	,000

Number of extreme cases (outliers) is 6, which is less than 2%, of all observed cases. Extreme values (outliers) are given in the Table 3.

In order to validate results, newly installed system was observed in different town. Average turnover form first model was used as starting value for modeling. Rest of time series was modeled from analytical expression for relationship between variables, based on weather variable that was recorded on the system, as predictor. Data sets from observed real system were tested against Hypothesis that they can be described with Normal distribution. Data sets from model, were tested against Hypothesis that they can be described with Log Normal distribution. In the first case results are distributed according to Normal distribution $N(2671,4912)$ and in the second, results were distributed according to Log Normal distribution $LogN(0.77,8.2)$. Histograms for both variables are presented on Figures 5 and 6.

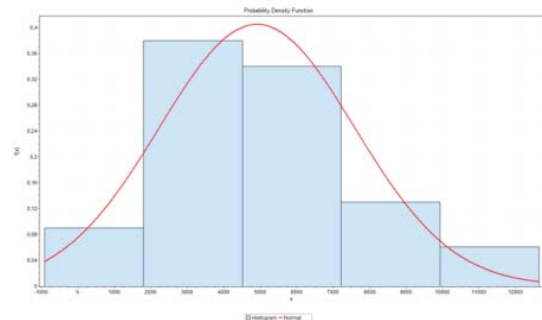


Figure 5: Histogram of data observed from real system

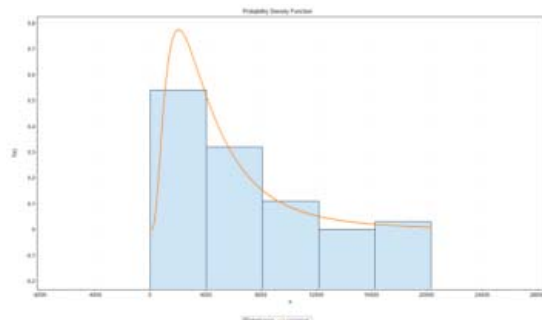


Figure 6: Histogram of data predicted with ARIMA model

Goodness of fit was tested with Kolmogorov-Smirnov test, for same significance level $\alpha=0,05$.

In the first case, where data on real system were tested against Hypothesis of Normal distribution, critical value for Kolmogorov –Smirnov test was 0,13403 for 100 recorder data sets and value of the test was 0,086.

In the second case, where modeled data were tested against Hypothesis of Log Normal distribution, critical value for Kolmogorov –Smirnov test was 0,13403 for 100 recorder data sets and value of the test was 0,078.

Student's T test was accomplished in order to test Hypothesis 3, that there is no significant difference between means on two independent samples.

According to calculated values of Student's T test, for significance level of $\alpha = 0,05$ and degrees of freedom $df=198$, value of the test was $t=0,553$. There is no significant difference between means on two variables.

DISCUSSION

Hypothesis 1 was tested and results confirmed stated Hypothesis that variable Daily turnover can be predicted by variable Daily weather condition. From the model fit it can be seen that goodness of the fit is 60,9%, which is considered as good model fit [06].

Number of outliers in fitted model was significantly small (below 2%), which also indicates goodness of the fit.

Hypothesis 2 was also confirmed. Analytical formulation of dependency was made, which also established mathematical model for next Hypothesis.

Hypothesis 3 was also confirmed. Based on proposed methodology, record sets of measurements from real system and record sets from modeled values were analyzed. Measured values were fitted to Normal distribution and modeled values were fitted to Log Normal distribution. Reason for this is in the fact that modeled values are dependent from previous record sets, so sudden changes from one weather condition to completely opposite one (from sunny weather to snow for example) can't be described completely by model.

Never the less model is in the end, for mentioned number of record sets giving good predictive results, with no statistically significant differences between two samples.

CONCLUSION

As described in the paper Daily turnover of automated car wash system depends from observed variable Daily weather condition. Measure of this dependence is calculated through ARIMA time series model. Results from modeling were compared to observed values got from another system. According to these results there is no statistically significant difference between two data sets, which implies that proposed ARIMA method can be used for prediction of daily turnover of car wash facilities.

Similar model can be used for estimation of daily turnover in other industry fields.

Future analyses will be in the direction of finding one or more variables that can be used for prediction of daily turnovers in other technical systems and comparing results with one published in this paper.

REFERENCES

- 1) Box, G.E.P., Jenkins, G.M., 1987. Time Series Analyses, Forecasting and Control. 2nd ur. San Francisco: Holden-Day.
- 2) Ho, S.L., 1998. The use of ARIMA models for reliability forecasting and analyses. Computers and industrial engineering, 35 (1-2), 213-216
- 3) Jiang, B., Heise, D.R., 2004. The Eye Diagram: New Perspective on the Project Life Cycle. Journal of education for Business, pp. 10-16
- 4) Morris, P. W., 1998. Managing Project interfaces: Key Points for Project Success. Englewood Cliffs, New Jersey: Prentice-Hall.
- 5) Newell, M.W., Grashina, M.N., 2004. The Project Management Question and Answer Book. New York: AMACOM.
- 6) Tabachnik, B.G., Fidell, L.S., 2007. Using multivariate statistics (5th edn.). Boston: Pearson Education

Paper sent to revision: 10.09.2012.

Paper ready for publication: 04.12.2012.