INTERNATIONAL JOURNAL OF BUSINESS AND MANAGEMENT STUDIES
Vol 3, No 1, 2011 ISSN: 1309-8047 (Online)

MODELING ISE100 WITH CONTINUOUS AR(1) MODEL

Yavuz Yıldırım
Yeditepe University
İnönü Mah. Kayışdağı Cad. 26 Ağustos Yerleşimi, 34755, Ataşehir - İstanbul
yavuz.yildirim@std.yeditepe.edu.tr

Gazanfer Ünal
Yeditepe University
İnönü Mah. Kayışdağı Cad. 26 Ağustos Yerleşimi, 34755, Ataşehir - İstanbul
gunal@yeditepe.edu.tr

─ Abstract ─
Great majority of the studies on Istanbul Stock Exchange market (ISE100) have focused on various type of discrete modeling such as AR/MA, ARIMA, GARCH, Vector AR and extensions of GARCH modeling. The importance of finding a suitable model for a stock exchange market and having an efficient forecast results from the model is undisputable. In this study we will model ISE100 with simple AR(1) model and taking a step further in analysis to continuous modeling. Recent challenge in financial time series modeling is to find an appropriate continuous model for the data used. In our case continuous AR(1) (CAR(1)) model will be applied to ISE100 and the results of the financial modeling will be evaluated.

Key Words: ISE100, Continuous Modeling, CAR(1), Discrete Modeling, Ornstein-Uhlenbeck Process

JEL Classification: C60 - General

1. INTRODUCTION

Modelling Istanbul Stock Exchange Market (ISE100) is a long and non-ending effort. The reasons behind it are obvıous. Most researches have tried to model ISE100 using discrete modelling. Whereas this study focuses on building up a continuous model for ISE100. Recent studies have shown that modeling returns in discrete time are just not enough to have a close relationship with the real time. Brockwell and Klüpperg has shown that continuous time modelling gets us closer to the real world.

We take a simple AR(1) model for ISE100. The coefficients of the discrete model are converted to continuous time. For the conversion process the autocavariances of discrete and continuous processes are considered. Eventually, we carry out simulations in contiuous time.

2. DATA

The data that has been used in this study is between 03/01/1994 and 23/06/2010. The summary statsitics of the log return of ISE100 is as follows:

Table1: summary statistics of log return of ISE100

min	1Q	median	3Q	max
-0.1998	-0.01318	0.001559	0.01593	0.1777
mean	std	skewness	kurtosis	Observations
0.001351	0.02808	-0.07164	7.101	4106

3. METHODOLOGY

3.1 Discrete modelling

3.1.1 Unitroot and statıonary tests

Any trending in the data should be removed before carrying out a unit root test. This can be explored by looking at the time series plot of the data. From the graph, one might obtain that whether there is a trend, cycle or seasonal pattern in the data. If any of these appear they should be removed accordingly.
The most popular unit root test is called Augmented Dickey Fuller (ADF) test.

ADF Test

$$\Delta z_t = \theta z_{t-1} + \alpha_1 \Delta z_{t-1} + \cdots + \alpha_p \Delta z_{t-p} + a_t$$

ADF test's hypothesis is:

H0: there is a unit root (data needs to be differenced to make it stationary)
H1: there is not a unit root (data is stationary)

When ADF is used, the choice of the correct lag requires particular attention, falling to do so will result in errors biasing the test. According to Ng and Perron (1995); first an upper limit for lag should be set, and then; if the absolute value of the t-statistic, to test the significance of the last lagged difference is less than 1.6 then, it is required to reduce the lag length and then repeat the process. The starting point for the lag length is suggested by Schwert(1989) to be:

pmax = 12 *(T/100)1/4

KPSS Test

The null hypothesis of KPSS test is that an observable series is stationary around a deterministic trend. The series is expressed as the sum of deterministic trend, random walk, and stationary error,
The KPSS test statistic is given by

$$KPSS = N^{-2}\sum_{t=1}^{N} S_t^2/\sigma^2(p)$$

3.1.2 AR(1) model

The notation AR(p) refers to the autoregressive model of order p. The AR(p) model is written as:

$$X_t = \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + a_t$$

where $\alpha_1, \ldots\ldots, \alpha_p$ are the parameters of the model, c is a constant and ε_t is white noise. The constant term is omitted by many authors for simplicity.

An autoregressive model is essentially an all-pole infinite impulse response filter with some additional interpretation placed on it.

Some constraints are necessary on values of the parameters of AR(p) model in order that the model remains stationary. For example, processes in the AR(1) model with $|\varphi_1| \geq 1$ will not be stationary.

$$X_t = \phi_0 + \phi_1 X_{t-1} + a_t$$

Taking the expectation

$$E[X_t] = \phi_0 + \phi_1 E[X_{t-1}] + E[a_t]$$

$E[a_t] = 0$ from the properties of Wiener process and under stationarity condition

$$E[X_t] = E[X_{t-1}] = \mu$$

$$\mu = \phi_0 + \phi_1\mu \text{ , } \mu = \frac{\phi_0}{1-\phi_1} \text{ , } \phi_0 = \mu(1-\phi_1)$$

X_t can be rewritten as

$$X_t = \mu(1-\phi_1) + \phi_1 X_{t-1} + a_t$$

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + a_t \qquad (1)$$

To find the variance, take the square of (1) and then take the expectation

$(X_t - \mu)^2 = \phi_1^2(X_{t-1} - \mu)^2 + 2\phi_1(X_{t-1} - \mu)a_t + a_t^2$
$E[(X_t - \mu)^2] = \phi_1^2 E[(X_{t-1} - \mu)^2] + 2\phi_1 E[(X_{t-1} - \mu)a_t] + E[a_t^2]$
Under stationarity condition $E[(X_t - \mu)^2] = E[(X_{t-1} - \mu)^2]$ and $cov(X_{t-1}, a_t) = 0$
$Var(X_t) = \phi_1^2 Var(X_t) + 0 + Var(a_t)$

$$Var(X_t) = \frac{\sigma_a^2}{1 - \phi_1^2} = \gamma_0$$

The sufficient and necessary conditions for AR(1) process to be weakly stationary are $\phi_1^2 < 0$, $-1 < \phi_1 < 1$

To get the autocovariance funciton of (1), multiply (1) by $(X_{t-h} - \mu)$ and then take the expectation
$(X_{t-h} - \mu)(X_t - \mu) = \phi_1(X_{t-1} - \mu)(X_{t-h} - \mu) + a_t(X_{t-h} - \mu)$
$E[(X_{t-h} - \mu)(X_t - \mu)] = \phi_1 E[(X_{t-1} - \mu)(X_{t-h} - \mu)] + E[a_t(X_{t-h} - \mu)]$
As $E[a_t(X_{t-h} - \mu)] = 0$
$cov(X_t, X_{t-h}) = \phi_1 cov(X_t, X_{t-h})$
$\gamma_h = \phi_1 \gamma_{h-1}$
For h=1, autocorrelation for AR(1) process

$$\gamma_1 = \phi_1 \gamma_0 = \frac{\phi_1 \sigma_a^2}{1 - \phi_1^2}$$

3.2 Continuous Modelling

The study of continuous time models has recieved great attention from the very successful use of continuous time models in theoritical finance, particularly with the work of Black-Scholes and Medon on the pricing of options. The luxury of using the formula and the derivative pricing tools become available when modelling is done in countinuous time. The analysis of timeseries data observed at irregularly spaced times can be handled very convieniently via continuous time models as pointed out by Jones (1981, 1985).

The discrete ARMA model is explained in detail by Box and Jenkins (1970). Drawback of DARMA compared to CARMA is the issue of temporal aggregation. The aggregation of a DARMA process results in a model that depends on the observation frequency, Weiss(1984). When the data is modelled in continuous time, the problem of temporal aggragation is not an issue anymore.

The CARMA (Continuous Autıregrresive Moving Average) model is a continuous time model for the instantaneous short rate with p autoregressive terms and q moving average terms with p>q, CARMA(p,q). The Vasicek model is a CARMA(1,0).

The Ornstein-Uhlenbeck process is;

$dX_t = -aX_t dt + b dW_t$

$Y_t = k(t)X_t + l(t)$

$dY_t = \left(\frac{\partial Y}{\partial t} + f\frac{\partial Y}{\partial X} + \frac{1}{2}g^2\frac{\partial^2 Y}{\partial X^2}\right)dt + g\frac{\partial Y}{\partial X}dW_t$

$\frac{\partial Y_t}{\partial t} = \frac{dk}{dt}X_t + \frac{dl}{dt} = \dot{k}X_t + \dot{l}$

$\frac{\partial Y_t}{\partial X} = k(t), \frac{\partial^2 Y}{\partial X^2} = 0$

$dY_t = \left(\dot{k}X_t + (-aX_t)k + \dot{l}\right)dt + bk dW_t$

Let $\dot{k} - ak = 0$, $\dot{l} = 0$

$k(t) = e^{at}$, $l = 0$

Hence $Y_t = e^{at}X_t \rightarrow dY_t = e^{at}dW_t$

Integration from “0” to “t” gives

$Y_t - Y_0 = \int_0^t e^{au}dW_u \rightarrow Y_t = Y_0 + \int_0^t e^{au}dW_u$

$Y_0 = e^{a(0)}X_0 = X_0$, $X_t = e^{-at}Y_t$

$X_t = e^{-at}\left(X_0 + b\int_0^t e^{au}dW_u\right) \rightarrow X_t = e^{-at}X_0 + b\int_0^t e^{a(u-t)}dW_u)$

$E[X_t] = e^{-at}E[X_0] + bE\left|\int_0^t e^{a(u-t)}dW_u\right|)$

Remark 1: $cov(X_t, X_s) = E[(X_t - E[X_t])(X_s - E[X_s])]$

Let $E[X_0] = X_0$

$X_t - E[X_t] = e^{-at}X_0 - e^{-at}X_0 + b\int_0^t e^{a(u-t)}dW_u$

$X_s - E[X_t] = b\int_0^s e^{a(u-s)}dW_u$

$$cov(X_t, X_s) = b^2 E\left[\left(\int_0^t e^{a(u-t)} dW_u\right)\left(\int_0^t e^{a(u-s)} dW_u\right)\right]$$

$$= b^2 e^{-a(t+s)} E\left[\left(\int_0^s e^{au} dW_u\right)^2\right]$$

Remark 2: $E\left[\left(\int_0^s e^{au} dW_u\right)^2\right] = \int_0^s E[e^{2au} du] = \int_0^s e^{2au} du = \frac{1}{2} e^{2as}$

Autocovariance of CAR(1) is

$$cov(X_t, X_s) = b^2 \frac{e^{-a(t+s)}}{2a} e^{as} = \frac{b^2}{2a} e^{-at-as+2as} = \frac{b^2}{2a} e^{-a(t-s)}$$

Let s=t-h

$$cov(X_t, X_{t+h}) = \frac{b^2}{2a} e^{-a(t-t+h)} = \frac{b^2}{2a} e^{-ah} \ , \ h > 0$$

The autocorrelation for CAR(1)

$$\gamma_Y(h) = \frac{b^2}{2a} e^{-a|h|}$$

The autocorrelation for AR(1)

$$\gamma_X(h) = \frac{\sigma^2 \phi^{|h|}}{1-\phi^2}$$

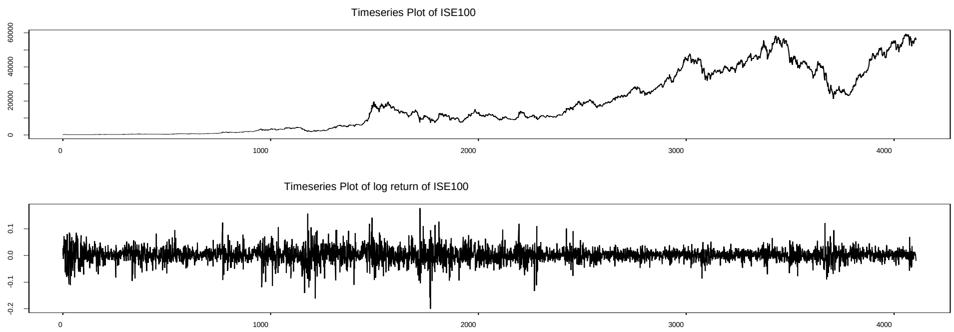
Setting autocorrelation of CAR(1) equal to AR(1)'s autocorrelation, we end up with

$$a = -\ln\phi \ , \ b^2 = \frac{-2\sigma^2}{1-\phi^2} \ln\phi$$

4. RESULTS

From figure 1, it is straight forward to see that the log return of ISE100 is a stationary process. This can also be tested via untiroot and stationarity tests.

Figure 1: The timeseries plot of ISE100 and the log return plot of ISE100



Source: own study

The results below indicate that log return of ISE100 has no unit root with the ADF's p-value of zero. KPSS test supports the stationarity of the return data by the test statsitcs 0.2691 which is greater than 0.05.

Table2: Unitroot and Stationarity Tests Results

Augmented DF Test; Null Hypothesis: there is a unit root	Test Statistic: -11.25 P-value: 5.203e-23
KPSS Test;Null Hypothesis: stationary around a constant	Test Statistics: 0.2691

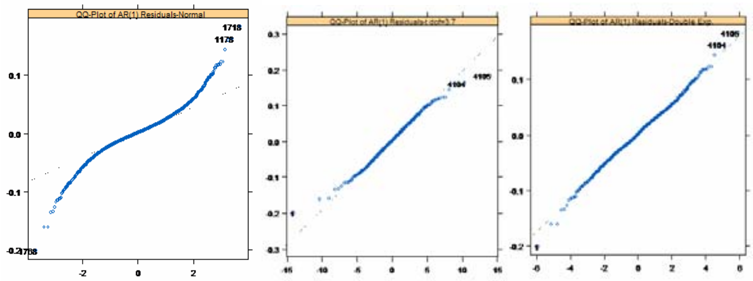
Next step is to consider whether AR(1) model is a satsifactory model for ISE100. With the t-values of that are greater than 1.65 states that the coeffcients are statistically significant at 95% confidence level. Hence AR(1) can be accepted as a good candidate model.

Table3: Coefficients of AR(1) model

	Estimate	Std. Error	t value	Pr(>|t|)
ar1	0.0621878	0.0155770	3.992	6.54e-05 ***
intercept	0.0012632	0.0004379	2.885	0.00392 **
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

The following figures show the QQ-plot of the residuals of AR(1) model. The usual expectation for stock exchange index values is that they have heavy tailed distributions. The summary statsitics for ISE100 sugeested a heavy tail distribution with kurtosis 7.101. The best fitted distribution function for the residuals is double exponential distribution.

Figure2: QQ-Plot of AR(1) Residuals



Source: own study

After specifiying a candidate discrete model and distirbutional properties of residuals, continuous modelling can be applied taking the coefficients of AR(1) model into account.

The general CAR(1) stochastic model, X_t

$$dX_t = -aXdt + b\,dW_t$$

The numerical solution for X_t is

$$X_t = e^{-a}X_{t-1} + b\sqrt{\frac{1-e^{-2a}}{2a}}W_t$$

Let's get CAR(1) parameters from AR(1) parameters;

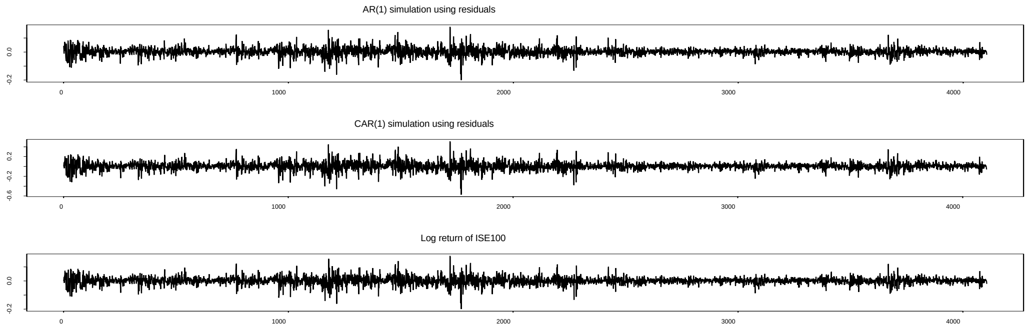
$\phi = 0.0621878$, $a = -\ln\phi$, $a = -\ln 0.062$, $a = 2.778$, $b = \frac{2\sigma^2}{1-\phi^2}a$

$$b = \frac{2}{1-0.062^2}1.206 = 2.42$$

$$dY_t = -2.778Y_t dt + 2.42dW_t$$

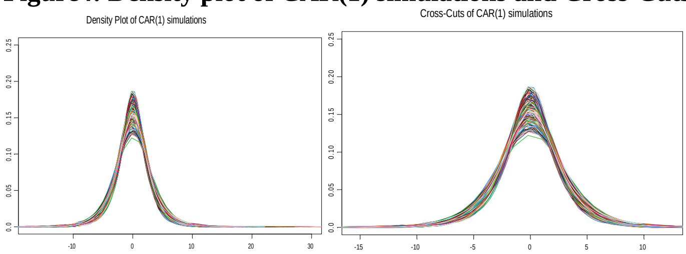
Figure 3 shows that continuous simulation gives the same outcomes as the log return data when the residuals from AR(1) model is used as Wt process in CAR(1) process. From figure 4 and 5 it can be concluded the simulations of CAR(1) model is quite close to the real data. Figure 6 and 7 proves that the distribution of the simulations stays consistent, which is expected from a good and well fitted model.

Figure3: AR(1) simulation, CAR(1) simulation, plot of log return of ISE100 using AR(1) residuals



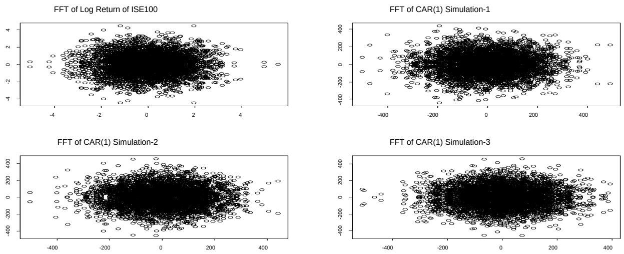
Source: own study

Figure4: Density plot of CAR(1) simulations and Cross-Cuts of CAR(1) simulations



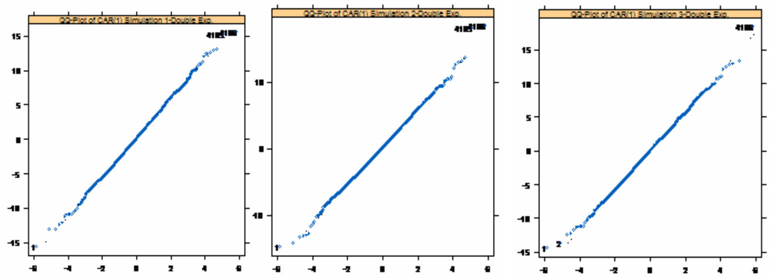
Source: own study

Figure5: Fast Fourier Transform (FFT) of log return data and CAR(1) simulations



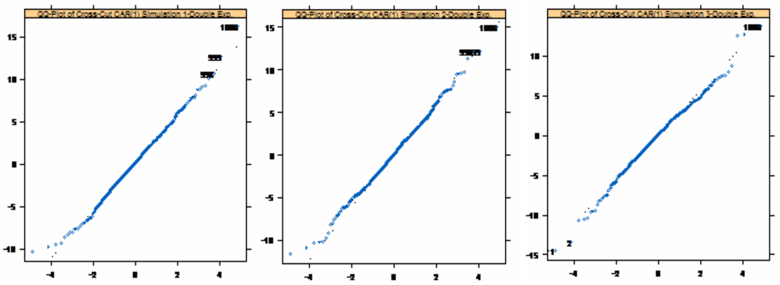
Source: own study

Figure6: QQ-Plot of CAR(1) Simulations



Source: own study

Figure7: QQ-Plot of Cross-Cut CAR(1) Simulations



Source: own study

CONCLUSION

This study focused on estimation of parameters for Istanbul Stock Exchange Market in continuous time. This is achieved by considering the properties of discrete and continuous modelling. Once the parameter conversion is done successfully, then the simulations in continuous time was carried out. The results and diagnostics checks have showed that using simple Ornstein-Uhlenbeck process as Continuous AR(1) model gives satisfactory outcomes. The simulations in continuous time is very close to the real log return data considering distributional properties and frequency domain.

BIBLIOGRAPHY

Box, G.E.P. and G.M. Jenkins (1970), Time series analysis: Forecasting and control, San Francisco: Holden-Day.

Brockwell, P. J. (1995), ‘A note on the embedding of discrete-time ARMA processes’, Journal of Time Series Analysis 16(5), 451–460.

Brockwell, P. J. (2000), Heavy-tailed and non-linear continuous-time ARMA models for financial time series, University of Hong Kong: Centre of Financial Time Series, Imperial College Press, London, pp. 3–23.

Jones, R. H. (1981), Fitting a continuous time autoregression to discrete data. In Applied Time Series Analysis //(Ed.. D. F. Findley), pp. 651-682. Academic Press, New York.

Jones, R. H. (1985), Time series analysis with unequally spaced data. In Time Series in the Time Domain, Handbook of Statistics 5 (Eds., E. J. Hannan, P. R. Krishnaiah and M. M. Rao), pp. 157— 178, North Holland, Amsterdam.

Phillips A.W. (1959), The estimation of parameters in systems of stochastic differential equations, Biometrika 46,67-76.