Fuzzy Sets and Systems 118 (2001) 9–19

www.elsevier.com/locate/fss

# Fuzzy ARIMA model for forecasting the foreign exchange market

Fang-Mei Tseng[a], Gwo-Hshiung Tzeng[b,*], Hsiao-Cheng Yu[a], Benjamin J.C. Yuan[a]

[a] *Institute of Management of Technology, College of Management, National Chiao Tung University, No. 1001 Ta-hsieh Rd, Hsih-chu, Taiwan, ROC*

[b] *Energy and Environmental Research Group and Institute of Traffic and Transportation, College of Management, National Chiao Tung University, No. 1001 Ta-hsieh Rd, Hsih-chu, Taiwan, ROC*

Received June 1997; received in revised form May 1998

**Abstract**

Considering the time-series ARIMA($p, d, q$) model and fuzzy regression model, this paper develops a fuzzy ARIMA (FARIMA) model and applies it to forecasting the exchange rate of NT dollars to US dollars. This model includes interval models with interval parameters and the possibility distribution of future values is provided by FARIMA. This model makes it possible for decision makers to forecast the best- and worst-possible situations based on fewer observations than the ARIMA model. © 2001 Elsevier Science B.V. All rights reserved.

*Keywords:* ARIMA; Foreign exchange market; Fuzzy regression; Fuzzy ARIMA; Time series

## 1. Introduction

Since it has been suggested by Box–Jenkins [1] that the time-series ARIMA model has enjoyed fruitful applications in forecasting social, economic, engineering, foreign exchange, and stock problems. It assumes that the future values of a time series have a clear and definite functional relationship with current, past values and white noise. This model has the advantage of accurate forecasting in a short time period; it also has the limitation that at least 50 and preferably 100 observations or more should be used. In addition, this model uses the concept of measurement error to deal with the differences between estimators and observations, but these data are precise values that do not include measurement errors.

Tanaka et al. [9–11] have suggested fuzzy regression to solve the fuzzy environment and to avoid a modeling error. This model is basically an interval prediction model with the disadvantage that the prediction interval can be very wide if some extreme values are present.

Song and Chissom [6–8] presented the definition of fuzzy time series and outlined its modeling by means of fuzzy relational equations and approximate reasoning. Chen [2] presented a fuzzy time-series method based on the concept of Song and Chissom. An application of fuzzy regression to fuzzy time-series analysis was

* Corresponding author.


PII: S0165-0114(98)00286-3

found by Watada [12], but this model did not include the concept of the Box–Jenkins model. In this paper, based upon the works of time-series ARIMA$(p,d,q)$ model and fuzzy regression model, we combine the advantages of two methods to develop the fuzzy ARIMA model.

In order to show the applicability and effectiveness of our proposed method in practical application, we conduct an illustration for forecasting the foreign exchange market. In the results, we found that the proposed method makes good forecasts in several situations for which FARIMA appears to be the most appropriate tool. The situations are listed as follows:

(i) To provide the decision makers the best- and worst-possible situations.
(ii) The required number of observations is less than the ARIMA model requires, which is at least 50 and preferably more than 100 observations.

The structure of this paper is organized as follows: Concepts of time-series ARIMA and fuzzy regression are reviewed in Section 2. In Section 3, the FARIMA model is formulated and proposed. The FARIMA model is applied to forecasting the foreign exchange rate of NT dollars to US dollars in Section 4 and finally the conclusions are discussed.

## 2. ARIMA model and fuzzy regression model review

A time-series $\{Z_t\}$ is generated by an ARIMA$(p,d,q)$ process with mean $\mu$ of the Box–Jenkins model [1]

$$\text{if } \varphi(B)(1-B)^d(Z_t-\mu)=\theta(B)a_t, \tag{1}$$

where $\varphi(B)=1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_p B^p$, $\theta(B)=1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q$ are polynomials in $B$ of degree $p$ and $q$, $B$ is the backward shift operator, $p,d,q$ are integers, $Z_t$ denotes the observed value of time-series data, $t=1,2,\ldots,k$, and time-series data are the observations.

The ARIMA model formulation includes four steps:

1. Identification of the ARIMA$(p,d,q)$ structure. Use autocorrelation function (ACF) and partial autocorrelation function (PACF) to develop the rough function.
2. Estimation of the unknown model parameter.
3. Diagnostic checks are applied with the object of uncovering possible lack of fit and diagnosing the cause.
4. Forecasting from the selection model.

It is assumed that $a_t$ are independent and identically distributed as normal random variables with mean 0 and variance $\sigma^2$, and the roots of $\varphi(Z)=0$ and $\theta(Z)=0$ all lie outside the unit circle. If possible, at least 50 and preferably 100 observations or more should be used. In the real world, however, the environment is uncertain and changes rapidly, we usually must forecast future situations using little data in a short span of time, and it is hard to verify that the data is a normal distribution. So this assumption has limitations. This model uses the concept of measurement error to deal with the difference between estimators and observations, but these data are precise values and do not include measurement errors. It is the same as the basic concept of the fuzzy regression model as suggested by Tanaka et al. [11].

The basic concept of the fuzzy theory of fuzzy regression is that the residuals between estimators and observations are not produced by measurement errors, but rather by the parameter uncertainty in the model, and the possibility distribution is used to deal with real observations.

The following is a generalized model of fuzzy linear regression:

$$Y=\beta_0+\beta_1 x_1+\cdots+\beta_n x_n=\sum_{i=1}^{n}\beta_i x_i=\boldsymbol{x}'\boldsymbol{\beta}, \tag{2}$$

where $\boldsymbol{x}$ is the vector of independent variables, superscript $'$ denotes the transposition operation, $n$ is the number of variables and $\beta_i$ represents fuzzy sets representing the $i$th parameter of the model.