Using a principal component analysis for multi-currencies-trading in the foreign exchange market

Article *in* Intelligent Data Analysis · June 2015

DOI: 10.3233/IDA-150738

5 authors, including:

Hyun Woo Byun
Yonsei University

Seungho Baek
City University of New York - Brooklyn College

Jae Joon Ahn
Yonsei University

Kyong Joo Oh
Yonsei University

Intelligent Data Analysis 19 (2015) 683–697
DOI 10.3233/IDA-150738
IOS Press

Using a principal component analysis for multi-currencies-trading in the foreign exchange market

Hyun Woo Byun[a], Seungho Baek[b], Jae Joon Ahn[c], Kyong Joo Oh[a,*] and Tae Yoon Kim[d]
[a]*Department of Information and Industrial Engineering, Yonsei University, Seoul, Korea*
[b]*Booth School of Business, University of Chicago, Chicago, IL, USA*
[c]*Department of Information and Statistics, Yonsei University, Wonju, Seoul, Korea*
[d]*Department of Statistics, Keimyung University, Daegu, Korea*

Abstract. This study proposes a multi-currencies trading algorithm that applies a stock-trading algorithm to the foreign exchange (FX) market. Our algorithm applies a principal component analysis and artificial neural networks to produce an induced classifier from the FX market. Our algorithm yields reasonable profits. In addition, we discuss a basic procedure that the currency-trading algorithm must follow.

Keywords: Multi-currencies-trading algorithm, principal component analysis, induced classifier, artificial neural networks

1. Introduction

The foreign exchange (FX) market is one of the most popular financial markets in which advanced analytical tools can be actively engaged [7,10,22]. Specific examples include a modified balance portfolio model [6], an FX-trading system based on adaptive reinforcement learning [22], a foreign exchange rates analysis based on rough set theory and acyclic graph support vector machines [23], exchange rate volatility prediction based on quantile regression [1] and trading strategy via logistic regression [15]. This trend has been particularly reinforced after the unprecedented volatility increase of FX market due to the recent financial crisis.

In FX markets, exogenous events such as news announcements, scheduled economic data releases, and other events usually drive traders to repeatedly compete for a limited resource (i.e., traders buy and sell currencies). Although the exact nature of the FX market is difficult to understand, the prices of currencies are expected to reflect the complex patterns of actions, feedback, and adaptation of traders. Thus, the time series of prices are the manifestation of the market's nature [5]. Under these assumptions, the FX-trading algorithm (FXTA) in this article was developed based on the correlations of multi-currency prices. Correlations in price movements are used for two reasons: risk management and trading signal detection. The likelihood of large losses can be significantly higher when the currencies held in

*Corresponding author: Kyong Joo Oh, Department of Information and Industrial Engineering, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, Korea. Tel.: +82 2 2123 5720; Fax: +82 2 364 7807; E-mail: johanoh@yonsei.ac.kr.

a multi-currency portfolio are positively correlated; therefore, an understanding of the correlation between different currencies can help in managing the risk associated with a portfolio. With regard to trading signal detection, correlations between currencies sharply increase when the key currency (e.g., the US dollar; USD) dominates the market. For instance, a sharp rise in correlations has been observed when the USD dominated the market during the 2008 collapse of Lehman Brothers. Thus, monitoring dynamic correlation changes can help in detecting trading signal.

In this study, we constructed an FXTA as an induced classifier that applies principal component analysis (PCA) and popular classifiers from stock-trading algorithms. This FXTA was designed for a market in which multi-currencies are traded in USD. FXTA has three phases. The first phase is to build a portfolio allocation among multi-currencies; the second phase is to construct an algorithm trading on the portfolio; and third phase is to install a moving window scheme for dynamic portfolio trading. The first and second phases are founded on PCA to use the correlations between multi-currencies. Note that PCA is used there because it explains variance-covariance structure by employing a linear combination of the original random variables. The second and third phases of the FXTA require a classifier for training data generated by the PCA and issuing trading signals. Our experiment indicates that the FXTA performs best when it engages the PCA via artificial neural networks (ANN) as its classifier. This paper consists of the following sections. Section 2 reviews the technical tools employed for FXTA. Section 3 describes the FXTA in detail. Section 4 tests the FXTA using empirical data, and concluding remarks are made in Section 5.

2. Review of technical tools

2.1. Principle component analysis (PCA)

PCA is a statistical method that explains variance-covariance structure by employing a linear combination of the original p random variables $X_1, \ldots, X_p$. Let the random vector $\underline{X} = (X_1, \ldots, X_p)$ have the covariance matrix $\sum$ with eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3 \ldots \geqslant \lambda_p$ and the corresponding eigenvectors $\mathbf{e}_1, \ldots, \mathbf{e}_p$, where

$$Var(\underline{X}) = \sum = \begin{bmatrix} \sigma_{11} \ldots \sigma_{1p} \\ \sigma_{21} \ldots \sigma_{2p} \\ \ldots \\ \sigma_{p1} \ldots \sigma_{pp} \end{bmatrix}.$$

Then, the ith principal component is given by

$$P_i = \mathbf{e}_i' \underline{X} = e_{i1}X_1 + \ldots + e_{ip}X_p \text{ for } i = 1, \ldots, p.$$

With these choices, we have

$$\begin{aligned} &Var(P_i) = \mathbf{e}_i' \sum \mathbf{e}_i = \lambda_i \text{ for } i = 1, \ldots, p \\ &Cov(P_i, P_k) = \mathbf{e}_i' \sum \mathbf{e}_k = 0, \text{ for } i \neq k \qquad (1) \\ &\mathbf{e}_i'\mathbf{e}_j = 1, \quad i = j \text{ and } \mathbf{e}_i'\mathbf{e}_j = 0, \quad i \neq j. \end{aligned}$$

Because

$$\sum_{i=1}^{p} Var(X_i) = tr\left(\sum\right) = tr(\mathbf{E}\Lambda\mathbf{E}') = tr(\Lambda\mathbf{E}'\mathbf{E}) = tr(\Lambda) = \sum_{i=1}^{p}\lambda_i \tag{2}$$

where $\mathbf{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_p)'$ and Λ is a diagonal matrix whose i-th diagonal elements is λ_i, the proportion of variance explained by the ith principal component $P_i = \mathbf{e}'_i\underline{X}$ is,

$$\frac{\lambda_i}{\sum_{k=1}^{p}\lambda_k}. \tag{3}$$

Together with $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3 \ldots \geqslant \lambda_p$, this equation implies that the variance is explained by the principal components in the order of $P_1, P_2, \ldots, P_p$. Later in Section 3.1 how PCA works for real problem will be illustrated.

2.2. *Logistic regression (LR)*

In statistics, the logistic (or logit) regression (LR) model is used to predict the probability of the occurrence of an event by fitting the data to a logistic curve. This analysis uses several predictor variables that can be either numerical or categorical to estimate the probability of an event occurrence. LR is used extensively in a variety of areas such as predicting patient survival rate and general financial risk management as well as measuring the default probability of a financial product. The basic expression for LR is

$$P = \frac{e^{\alpha+\beta' X}}{1 + e^{\alpha+\beta' X}} = \frac{1}{1 + e^{-(\alpha+\beta' X)}}, \tag{4}$$

where P is the probability of success, and α and β are parameter vectors. The alternative expression is

$$\log\left(\frac{P}{1-P}\right) = \alpha + \beta' X \tag{5}$$

which enables simple linear regression. Note that LR is a parametric method based on a specific model.

2.3. *Decision tree (DT)*

DT learning is one of the most widely used methods of inductive inference. This method approximates discrete-valued functions that are robust regarding noisy data and capable of learning “disjunctive” logic. DT is generally acceptable for predictive modeling since it is easy to interpret and able to model complex input/output associations via the automatic handling of missing values [16,23]. Thus, DT is likely a parametric-type machine learning algorithm.

Various algorithms can be used for DT learning. Quinlan (1986) developed C4.5 and C5.0 by employing information and measure theories [12]; Breiman et al. (1984) based CART (classification and regression tree) on an inductive learning approach [16]. Our empirical study uses the standard CART algorithms designed for both real-valued and integer-valued responses. CART algorithms use the Gini diversity index to split the tree nodes and the cross-validation technique to prune the trees as default. Note that the Gini diversity index plays a key role in choosing the variables at each step and determining the next best variable for splitting the set of items.

2.4. Artificial neural networks (ANN)

ANN have recently gained popularity with regard to testing the movement of a variety of financial variables [11,14,25]. In particular, ANN are universal function approximator that maps any nonlinear function [13]; thus, they are less sensitive to error term assumptions and can tolerate noise, chaotic characteristics, and heavy tails better than most other statistical methods by properly designing their network. Our empirical study used a three-layer, fully connected backpropagation neural network (BPN) among various ANN algorithms. The most challenging task of the BPN training process is determining the correct independent variables and the number of neurons in the hidden layer as well as the learning rate, momentum rate, and transfer function. In practice, no robust theoretical foundation can be used to select the appropriate BPN architecture. Therefore, trial-and-error experiments were conducted to build an appropriate BPN architecture. In our empirical study of Section 4, ANN was set via 4 hidden nodes, and the logistic transfer function with learning rate, momentum and initial weight given by 0.1, 0.1 and 0.3, respectively. Note that ANN are the most nonparametric method due to its data-oriented tendency.

2.5. Charting

Charting, also known as "technical analysis charting", has been a financial analytical tool for many decades [2]. Charting was the original investment analysis and first used in the early 1800s [26]. Furthermore, most chartings have been made with regard to stock market analysis. Beginning in the early 1970s, charting extended its use to foreign currency trading [4]. This tool includes buy and sell signals based on simple trend-detecting techniques.

In our empirical example of Section 4, we have induced mapping by equation

$$f_t : \mathbf{Z}_{(t)} \to Y_{(t)}$$

where predictor $\mathbf{Z}_{(t)} = (\mathbf{z}_{\cdot 1(t)}, \mathbf{z}_{\cdot 2(t)}, \mathbf{z}_{\cdot 3(t)}, \mathbf{z}_{\cdot 4(t)})$ is 55 $\times$ 4 matrix and response $Y_{(t)} = (-1, \ldots, 1)'_t$ is 55 $\times$ 1 vector with -1 representing short position and $+1$ long position (see Eq. (13) below). Then f_t was estimated as $\hat{f}_t$ by employing various techniques (e.g., LR, DT, and ANN discussed above). Here we used four popular chartings for producing $\mathbf{Z}_{(t)}$ in FXTA [18,19]: relative position, relative strength index (RSI), momentum, and moving average convergence/divergence (MACD). These chartings are displayed in Table 1. See Remark 3 of Section 3.

3. Trading algorithm

This section describes the specific FXTA procedure. The FXTA consists of three phases. The first phase constructs the multi-currency portfolio; the second phase builds a trading rule on the portfolio; and the third phase builds a moving window scheme for the dynamic implementation of the FXTA. PCA plays a key role. At the end of each subsection below, empirical studies are referred to for illustrating implementation of each phase and they are the empirical analysis of Section 4.

3.1. Phase 1: Portfolio

Let $\mathbf{X}_{(t)}$ be a $n \times p$ matrix at time t whose element $x_{ij(t)}$ denotes the exchange rate of the jth currency in USD on the ith day when the exchange trading (buying or selling) of the multi-currencies is conducted

Table 1
Chartings used for FXTA (g_i denotes the closing value at time i)

Technical indicator	Formula $\delta(g_i, \ldots, g_{i-m+1}) \in R$
Relative position	$RP_i(m) = \dfrac{g_i - \pi_m}{\eta_m - \pi_m}$ where π_m and η_m are the minimum and maximum values of $(g_{i-m+1}, \ldots, g_i)$
Relative strength index (RSI)	$RSI_i(m) = 100 - \dfrac{100}{1 + RS_i(m)}$ where $RS_i(m) = \sum_{l=1}^{m} U_{i-m+l} \Big/ \sum_{l=1}^{m} D_{i-m+l}$ where $U_i = \max(g_i - g_{i-1}, 0)$ and $D_i = \max(g_{i-1} - g_i, 0)$
Momentum	$Momentum_i(m) = \dfrac{g_i}{g_{i-m}} \times 100$
Moving average convergence-divergence (MACD)	$MACD_i(m_1, m_2) = EMA_i(g, m_1) - EMA_i(g, m_2)$ where $EMA_i(g, m) = \dfrac{\sum_{l=1}^{m} a^{m-i} g_{i-m+l}}{\sum_{l=1}^{m} a^{m-i}}$ where a is an exponential factor ($0 < a < 1$)

Table 2
Currency distribution of FX market turnover

Rank	Currency	% Daily share
1	US dollar (USD)	43.15
2	Euro (EUR)	18.5
3	Yen (YEN)	8.25
4	Pound sterling (GBP)	7.5
5	Swiss franc (CHF)	3.4
6	Australian dollar (AUD)	3.35
7	Canada dollar (CAD)	2.1
8	Swedish Krona (SEK)	1.4
9	Hong Kong dollar (HKD)	1.4
10	Norwegian Krone (NOK)	1.1
11	New Zealand dollar (NZD)	0.95
12	Mexican Peso (MXN)	0.65
13	Singapore dollar (SGD)	0.6
14	On (KRW)	0.55
15	Others	7.1
	Total	100

*Triennial Central Bank Survey, Bank for International Settlements (BIS), Dec 2007.

for $j = 1, \ldots, p$ and $i = t - n + 1, \ldots, t$ (i.e., p multi-currencies over the last n days from the current day t). To construct the multi-currency portfolio, we first apply PCA to $\mathbf{X}'_{(t)}\mathbf{X}_{(t)}$ (an estimate of $\sum_{(t)}$) and obtained its first principal component,

$$P_{1(t)} = \mathbf{e}'_{1(t)}\underline{X} = e_{11(t)}X_1 + \ldots + e_{1p(t)}X_p$$

where $\mathbf{e}_{1(t)} = (e_{11}(t), \ldots, e_{1p}(t))'$ is the eigenvector corresponding to λ_1 or the first PC vector of $\mathbf{X}'_{(t)}\mathbf{X}_{(t)}$. Using $\mathbf{e}'_{1(t)}\mathbf{e}_{1(t)} = 1$, the portfolio allocation among p multi-currencies was conducted according to weight

$$w_{j(t)} = e^2_{1j(t)} \text{ for } j = 1, \ldots, p \tag{6}$$

where $e_{1j(t)}$ is the jth element of the first PC vector $\mathbf{e}_{1(t)}$. In our empirical studies of Section 4 we select $p = 9$ multi-currencies over the last $n = 55$ days from the current day t and hence

$$\mathbf{X}_{(t)} = (x_{ij(t)}) = \begin{pmatrix} r_{1,t-54} & r_{2,t-54} & \cdots & r_{9,t-54} \\ r_{1,t-53} & r_{2,t-53} & \cdots & r_{9,t-53} \\ \vdots & \vdots & \vdots & \vdots \\ r_{1,t} & r_{2,t} & \cdots & r_{9,t} \end{pmatrix} \tag{7}$$

where r_{ij} is the exchange rate of the ith currency at jth day.

3.2. Phase 2: Trading algorithm

To generate training data conditional on $\mathbf{X}_{(t)}$, we used the following rule for response variable generation:

$$h(u_i) = 1,\ \text{if } u_i = \frac{g_{i(t)} - g_{i-k(t)}}{g_{i-k(t)}} > S \text{ for a fixed lag } k \geqslant 1 \tag{8}$$

$$h(u_i) = -1,\ \text{if } u_i = \frac{g_{i(t)} - g_{i-k(t)}}{g_{i-k(t)}} < -S \text{ for a fixed lag } k \geqslant 1 \tag{9}$$

where

$$g_{i(t)} = \mathbf{x}_{i\cdot(t)}\mathbf{e}_{1(t)},\ \ \mathbf{x}_{i\cdot(t)} = (x_{i1(t)}, \ldots, x_{ip(t)}) = (r_{1,t-n+i}, \ldots, r_{p,t-n+i}) \tag{10}$$

and S is a pre-given threshold, and 1 and -1 denote long and short positions, respectively. Here $g_{i(t)}$ is a projection score of $\mathbf{x}'_{i\cdot(t)}$ on the 1st PC vector $\mathbf{e}_{1(t)}$. Now, the response data are

$$Y_{(t)} = (h(u_1), \ldots, h(u_n))'_t \tag{11}$$

which is a record of buy or sell positions according to the rule Eqs (8) and (9) over the last n days from given time t. In our empirical studies, when $n = 55$, $k = 1$ and $S = 0.03$, $Y_{(t)} = (h(u_1), \ldots, h(u_{55}))'_t = (-1, \ldots, 1)'_t$ indicates that we sell on $t - 54$th day because

$$u_1 = \frac{g_{1(t)} - g_{0(t)}}{g_{0(t)}} < 0.03$$

where $g_{1(t)} = \mathbf{x}_{1\cdot(t)}\mathbf{e}_{1(t)}$, $g_{0(t)} = \mathbf{x}_{0\cdot(t)}\mathbf{e}_{1(t)}$, $\mathbf{x}_{1\cdot(t)} = (r_{1,t-54}, \ldots, r_{p,t-54})$ and $\mathbf{x}_{0\cdot(t)} = (r_{1,t-55}, \ldots, r_{p,t-55})$.

For predictor variables, we use relative position, RSI, momentum and MACD applied to $g_{i(t)} = \mathbf{x}_{i\cdot(t)}\mathbf{e}_{1(t)}$ for $i = 1, \ldots, n$. Indeed applying the formulae $\delta_v : R^m \to R$ from Table 1 to each $g_{i(t)}$ where δ_1, δ_2, δ_3 and δ_4 are the formulae for relative position, RSI, momentum and MACD, respectively, we obtain for $i = 1, \ldots, n$

$$\delta_{\nu(t)}(g_{i(t)}, \ldots, g_{i-m+1(t)}) = z_{i\nu(t)} \in R \text{ for } v = 1, 2, 3$$

$$\delta_{4(t)}(g_{i(t)}, \ldots, g_{i-m+1(t)}; m_1, m_2) = z_{i4(t)} \in R \text{ for } v = 4.$$

Note that there are two parameters m_1 and m_2 for δ_4. Now we might establish the $n \times 4$ matrix $\mathbf{Z}_{(t)}$ given by

$$\mathbf{Z}_{(t)} = (\mathbf{z}_{\cdot1(t)}, \mathbf{z}_{\cdot2(t)}, \mathbf{z}_{\cdot3(t)}, \mathbf{z}_{\cdot4(t)}). \tag{12}$$

Recall that $m(m = \max(m_1 \cdot m_2)$ for $v = 4)$ is a fixed number less than or equal to n. Thus, Eqs (11) and (12) yielded the training dataset $(\mathbf{Z}_{(t)}, Y_{(t)})$ that induced the mapping

$$f_t : \mathbf{Z}_{(t)} \to Y_{(t)}. \tag{13}$$

Next, f_t was estimated as $\hat{f}_t$ by employing various techniques (e.g., LR, DT, and ANN). In our empirical studies, when $v = 1$ and $m = 55$

$$z_{i1(t)} = \frac{g_{i(t)} - \pi_{55,i(t)}}{\eta_{55,i(t)} - \pi_{55,i(t)}}$$

where $\pi_{55,i(t)} = \min(g_{i-54(t)}, \ldots, g_{i(t)})$ and $\eta_{55,i(t)} = \max(g_{i-54(t)}, \ldots, g_{i(t)})$ for $g_{i(t)}$ given by Eq. (10).

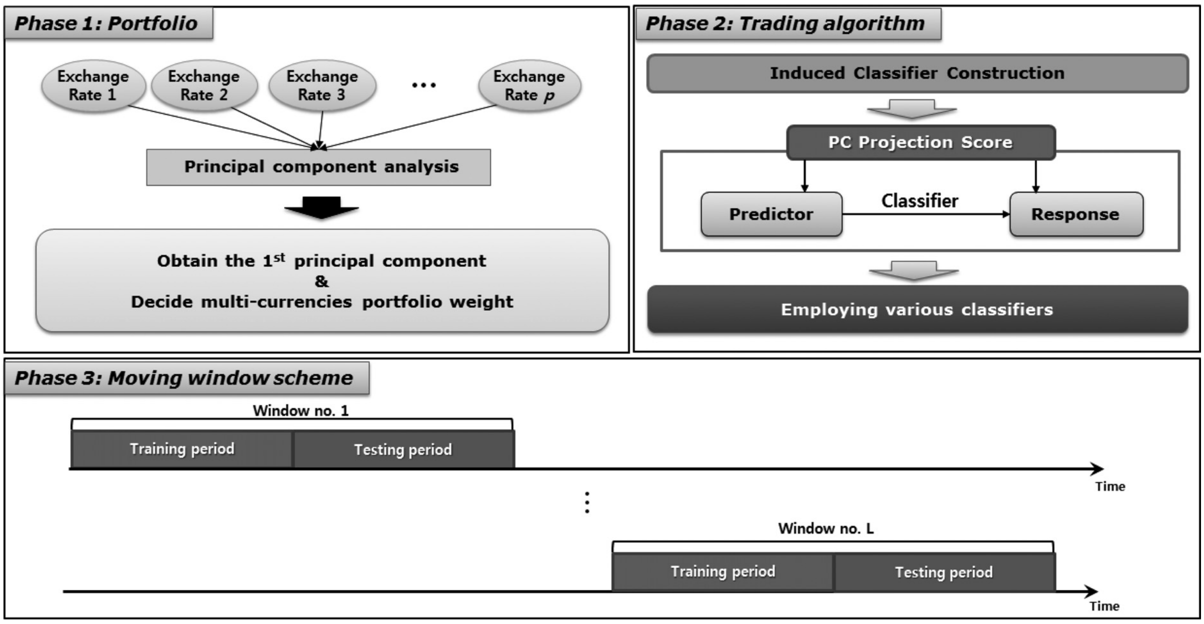

Fig. 1. Architecture of FXTA. (Colours are visible in the online version of the article; http://dx.doi.org/10.3233/IDA-150738)

3.3. Phase 3: Moving window scheme

A moving window training scheme with a fixed window width was used to build a dynamic FXTA (refer to [8,9] or Fig. 4 below). Using a moving window at $t = \tau_l$ for $l = 1, 2, 3 \ldots, L$, we established the training dataset $(\mathbf{Z}_{(\tau_l)}, Y_{(\tau_l)})$ and its induced mapping

$$f_{\tau_l} : \mathbf{Z}_{(\tau_l)} \to Y_{(\tau_l)} \ \text{ for } l = 1, \ldots, L \tag{14}$$

where $(\mathbf{Z}_{(\tau_l)}, Y_{(\tau_l)})$ was obtained using Eqs (11) and (12) from the $n_l \times p$ matrix $\mathbf{X}_{(\tau_l)}$ where n_l was the number of daily exchange rates inside the left window width of the moving window at $t = \tau_1$ (or $\mathbf{x}_{i\cdot(\tau_l)} = (x_{i1(\tau_l)}, \ldots, x_{ip(\tau_l)}) \in \mathbf{X}_{(\tau_l)}$ for $i = \tau_l - n_l + 1, \ldots, \tau_l - 1, \tau_l$, equivalently). Next, f_{τ_l} was estimated as $\hat{f}_{\tau_l}$, which is used as a trading rule for the testing period of the lth window (or $\mathbf{x}_{i\cdot(\tau_l+n_l)} = (x_{i1(\tau_l)}, \ldots, x_{ip(\tau_l)}) \in \mathbf{X}_{(\tau_l+n_l)}$ for $i = \tau_l + 1, \ldots, \tau_l + n_l$, equivalently). Recall that the moving portfolio weights for the lth window are given by

$$\left(w_{1(\tau_l)} = e^2_{11(\tau_l)}, \ldots, w_{p(\tau_l)} = e^2_{1p(\tau_l)}\right)$$

via Eq. (6) where $e_{1j(\tau_l)}$ is the jth element of the first PC vector $\mathbf{e}_{1(\tau_l)}$ from $\mathbf{X}_{(\tau_l)}$. In our empirical studies, with the window width $n_l = 55$, 55×9 matrix $\mathbf{X}_{(\tau_l)}$ is training data and another 55×9 matrix $\mathbf{X}_{(\tau_l+55)}$ is the test data. The overall structure of the multi-currency trading system FTXA is shown in Fig. 1.

Remark 1. As Laloux et al. (1999) discussed [17], an increase in the variance accounted for by a PC might be due to increases in the correlations among a limited number of currencies (which then have large PC coefficients) or a market-wide effect in which many currencies begin to significantly contribute to the component. This distinction is important with regard to the optimal portfolio selection of Phase

Fig. 2. Projection score $g_{iC} = \mathbf{x}_{i\cdot C}\mathbf{e}_{1C}$ for $i = 1, \ldots, T_C$ (Prior checking period: 1/1/2000–12/31/2004).

Fig. 3. Projection score $g_{iE} = \mathbf{x}_{i\cdot E}\mathbf{e}_{1E}$ for $i = 1, \ldots, T_E$ (Experimental period: 01/01/2005–12/31/2009).

1 because these changes have different financial implications. If the correlations between all currencies increase, then it becomes more difficult to reduce risk by diversifying across different currency classes. In contrast, increases in the correlations within a currency class (e.g., "Euro class" refers to the set of currencies that behave similar to the Euro) that are not accompanied by increases in correlations between currency classes affect diversification or portfolio selection less significantly. Our FXTA handles these risks related to Phase 1 "dynamically" by finding optimal portfolio selection among multi-currencies via the moving window scheme.

Remark 2. Trading rules Eqs (8) and (9) of Phase 2 are based on the following strategy: "Buy the multi-currency portfolio when the key currency weakens, and sell the multi-currency portfolio when the key currency strengthens." This strategy is based on the fact that the projection score of the multi-currencies from the first PC (or $g_{i(t)} = \mathbf{x}_{i\cdot(t)}e_{1(t)}$ for given ith day) tends to rise (fall) when the key currency is weak (strong). Refer to Figs 2 and 3 and their explanations in Section 4.2 below. This strategy may be also viewed as one based on FX volatility: "Buy the multi-currency portfolio when FX volatility increases, and sell the multi-currency portfolio when FX volatility decreases". This is true because FX volatility tends to increase when the key currency weakens and vice versa. Recall that PCA factors FX volatility on principal components as discussed in Section 2.1. Thus our algorithm FXTA is expected to work reasonably well in a highly volatile environment as long as the key currency exists.

Remark 3. We used relative position, RSI, momentum and MACD applied to the projection score for predictor variables. These indicators are related to monitoring volatility or the trend of the projection score. In fact, relative position and momentum are related to trend, whereas RSI and MACD are related to volatility. From modeling aspect, one might notice that these predictors selected enable FXTA to work like error correction model [18] which dynamically monitors trend as well as volatility for checking market stability of multi-currencies (or cointegration variables). Notice that the four indicators were selected among technical indicators of kind mainly because simulated experiments favor them over others.

Remark 4. As Kühl (2007) discussed [19], the correlation between the exchange rates expressed in the same currency must automatically be stationary when their cointegration holds. In a monetary system with target zones, the exchange rates expressed in a currency to which it does not belong shall be cointegrated; therefore, the correlation should move within the defined ranges and exhibit mean reverting behavior. Although the multi-currencies under investigation are not under a tightly controlled monetary system, they are partially controlled by USD. Thus, the correlation might display some stable character though it behaves in a non-stationary manner depending on t. The moving window scheme is desirable since it is able to capture such a non-stationary mapping f_t of Eq. (13) via a non-homogeneous Markovian manner. Indeed the moving window scheme is based on the assumption that the f_t having Markovian character changes its transition mechanism over time (non-homogeneous Markov transition). Consequently, the moving window scheme here might apply to various complex FX products as long as they are to a degree controlled by some key currencies. For related discussions, refer to Tables 3–5 with their discussions in Section 4. In addition, note that the charting tools that transform numerical values within a defined range usually stabilize f_t.

Remark 5. To build a better FXTA mapping, three classifiers (LR, DT, and ANN) were considered. Note that these classifiers are parametric, close to nonparametric and purely nonparametric, respectively. Thus, the classifier match to the FXTA depends on the type of mapping induced by FXTA via Eq. (14). Our experiments discussed in the next section indicate that ANN best matches FXTA, which implies that FXTA tends to induce a pure nonparametric or nonlinear mapping. Recalling that our moving window is based on non-homogeneous Markovian process, it is not surprising that ANN designed for nonlinear fitting matches the FXTA adequately. In the meantime, it is worth mentioning that FX market expert's opinion is quite important for successful FXTA implementation. For instance, there are a number of parameters to be selected by the FX market experts, e.g. the window width of the moving window and p, the number of multi-currencies. In our empirical studies of Section 4, we select $p = 10$ and the window width 55 by following the expert's opinion. Regarding selection of p, our simulation observe that PCA does not change much when p exceeds 10 indicating $p = 10$ to be optimal. Regarding selection of window width, market experts recommended 55 corresponding to the number of exchange rates during one quarter period to reflect real economic situation efficiently (e.g., quarterly scheduled economic data releases from local governments), which is also checked and supported by simulations there. Finally notice that FXTA is here designed to trade multi-currencies whose daily trading volume exceeds at least 1 percent of the entire FX market trading volume and that its trading is based on the training of the past exchange rates of the multi currencies. Thus FXTA is not recommended for trading pegged or renamed currencies since they are scarcely traded in FX market. In addition, when some currencies happen to merge to a new currency, as in the case with the EUR, FXTA may use the exchange ratios between the merged currencies at the time of the merge for its trading.

This section closes with presentation of the pseudo code of the FTXA algorithm.

Pseudo code of the algorithm

Phase1

1) From $\mathbf{X}'_{(t)}\mathbf{X}_{(t)}$, obtaining the first principal component with its corresponding eigenvector, $\mathbf{e}_{1(t)}$ by using PCA.
2) Decide multi-currencies portfolio weight $(e^2_{1j(t)})_{j=1,\dots,p}$.

Phase2

1) Produce response variable for trading algorithm for $i = 1, \dots, n$.
   {

Table 3
Eigenvalues and the proportion of variance explained by PCs (Prior checking period: 01/01/2000–12/31/2004)

Contents	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Eigenvalue	7.873	0.831	0.151	0.065	0.037	0.017	0.015	0.009	0.003
Var. explained	87.5%	9.2%	1.7%	0.7%	0.4%	0.2%	0.2%	0.1%	0%

*PC: principal component.

a) if (return rate of the 1st PC score $(g_{i(t)}) > S$ (pre-given threshold))
{generate buy signal ($h(u_i) = 1$)}
b) else if (return rate of the 1^{st} PC score $(g_{i(t)}) < -S$(pre-given threshold))
{generate sell signal ($h(u_i) = -1$)}

2) Produce technical indicators for predictor variables
(relative position ($\mathbf{z}_{\cdot 1(t)}$), RSI ($\mathbf{z}_{\cdot 2(t)}$), momentum ($\mathbf{z}_{\cdot 3(t)}$) and MACD ($\mathbf{z}_{\cdot 4(t)}$))
3) Employ various classifiers (LR, DT and ANN) for generating trading rule
}

Phase3

1) Apply trading rule based on moving window training scheme
For ($l = 1$ to L;)
{
// trading rule generation from training period
do Phase 1 and Phase 2 at τ_l
// trading rule application to testing period
do trading with the generated rule at τ_l
}

4. Empirical studies

4.1. Data

The data in this study were collected from the economic statistics system (ECOS) of the Bank of Korea (BOK). To develop and test the FXTA, the top nine foreign currencies ($p = 9$) with regard to daily trading volume were selected, and their daily prices were investigated. Daily currency prices were taken as the exchange rates of each currency in USD. Table 2 describes the daily trading proportions of the various currencies offered via BIS of the Triennial Central Bank Survey [3]. Note that Hong Kong dollar (HKD) currency was intentionally excluded from our selection because China has governed Hong Kong since 1997 when Hong Kong's sovereignty was officially transferred from the United Kingdom to the China.

4.2. FXTA experimental results

To determine validity with regard to using FXTA, the movements of the exchange rates of 9 currencies with regard to USD during experimental period (01/01/2005–12/31/2009, say $\mathbf{X_E}$) and prior checking period (01/01/2000 to 12/31/2004, say $\mathbf{X_C}$) were analyzed separately via principal component analysis.

Table 4
Eigenvalues and proportion of variance explained by PCs (Experimental period: 01/01/2005–12/31/2009)

Contents	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Eigenvalue	5.87	2.36	0.35	0.25	0.06	0.04	0.03	0.02	0.01
Var. explained	66%	26%	4%	3%	1%	0%	0%	%	0%

*PC: principal component.

Table 5
The first principal component vector

	PC1 ($\mathbf{e}_{1C}$*)	PC1 ($\mathbf{e}_{1E}$**)
AUD	0.348009	0.399167
GBP	0.348437	0.263088
CAD	0.334693	0.365748
EUR	0.351889	0.361662
JPY	0.223918	0.007371
NOK	0.327696	0.403572
SEK	0.352934	0.392985
CHF	0.339704	0.278987
NZD	0.352351	0.334192

*$\mathbf{e}_{1C}$: Prior checking period (01/01/2000–12/31/2004);
**$\mathbf{e}_{1E}$: Experimental period (01/01/2005–12/31/2009).

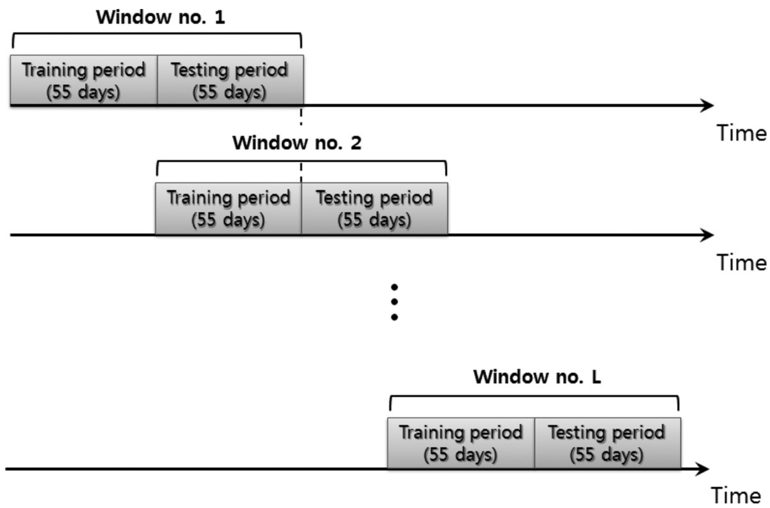

Fig. 4. Moving window scheme with L windows for FXTA. (Colours are visible in the online version of the article; http://dx.doi.org/10.3233/IDA-150738)

As shown in Table 3 for the prior checking period, the first PC explained 87.5% of the movement, and the second PC explained 9.2%, for a total of 96.7%. Note that Table 3 calculates this by finding the Eigen values $\lambda_1, \ldots, \lambda_9$ of $\mathbf{X_C}'\mathbf{X_C}$ and then figuring out $\lambda_i / \sum_{k=1}^{9} \lambda_k$ for $i = 1, \ldots, 9$ (refer to Eq. (2)). Table 4 shows that the first PC of $\mathbf{X_E}'\mathbf{X_E}$ explained 66% of the movement during the experimental period, and the second PC explained 26%, for a total of 96.7%. Together, these results indicate that the first PC might well serve the FXTA, although its pivotal role varied throughout the period. The amount of variance in multi-currency movements explained by the first PC is not striking and demonstrates that a large amount of "common" variation is found in FX markets. This finding highlights the expected close links between the multi-currencies. The decrease in the variance explained by the first PC during the experimental period also implies that the multi-currencies became less related as USD depreciation continued over time. Keep in mind that the significant rise in the variance explained by the first PC implies that the correlation among the currencies increased or were dominated by a strong USD.

To understand how the first PC relates to economic reality, the first PC $\mathbf{e}_{1C}$ ($\mathbf{e}_{1E}$) and its projection score $g_{iC} = \mathbf{x}_{i \cdot C}\mathbf{e}_{1C}$ ($g_{iE} = \mathbf{x}_{i \cdot E}\mathbf{e}_{1E}$) for $i = 1, \ldots, T_C$ ($i = 1, \ldots, T_E$) during the prior checking period (the experimental period) are shown in Table 5 and Figs 2 and 3. Recall that $\mathbf{x}_{i \cdot}$ denotes the values of the 9 currencies in USD on ith trading day. Table 5 shows a significant down-weight of JPY (from 0.2239 to 0.0073) during the experimental period, which appears to be mainly due to severe JPY depreciation during that period. Because the rise (fall) of the projection score is closely related to value depreciation (appreciation) of the USD, the PC projection score reflects these economic episodes fairly well. In fact, Figs 2 and 3 indicate that the USD continued to appreciate until February 2002 and then began to depreciate as the IT bubble (having captivated the US over decades) finally began to burst. The

Table 6
First PC and portfolio weights for the first window

USD against foreign currencies	AUD	GBP	CAD	EUR	JPY	NOK	SEK	CHF	NZD
Component 1	−0.0328	0.2612	0.3309	0.4426	0.2059	0.4340	0.4471	0.4409	0.0148
Weight for PC1	0.11%	6.83%	10.95%	19.59%	4.24%	18.84%	19.99%	19.44%	0.02%

continuous rally of the USD until 2002 was possible because global investment money flowed into US thanks to its IT business boom, despite its severe current-account deficit. After 2002, the central banks of other countries continuously cut their shares of USD (particularly in 2007; see Fig. 3). This depreciation has continued since that time, although it significantly rallied when the global financial crisis occurred in 2008. Figure 3 displays the 2008 global financial crisis vividly. When the 2008 crisis occurred, most investors in need of USD sold their investment assets, which engendered a severe demand increase of USD in FX markets. This demand increase from the 2008 financial crisis is highlighted by the Lehman Brothers bankruptcy.

The moving window scheme employed in this study is illustrated in Fig. 4, where window width for training and testing was set at 55 days. Using the initial window at $t = \tau_1$ (or t is at 03/01/2005), Table 6 calculates the first principal component vector $\mathbf{e}_{1(\tau_1)}$ from the given $n_1 \times$ 9 matrix $\mathbf{X}_{(\tau_1)}$ with $n_1 = 55$ taken as the number of daily exchange rates of the 9 currencies during the last 55 days from $t = \tau_1$. The portfolio weights are found by

$$(w_{1(\tau_1)} = e^2_{11(\tau_1)}, w_{2(\tau_1)} = e^2_{12(\tau_1)}, \ldots, w_{9(\tau_1)} = e^2_{19(\tau_1)})$$

via Eq. (6), where $e_{1j(\tau_1)}$ is the jth element of $\mathbf{e}_{1(\tau_1)}$. This calculation determines the portfolio of the 9 currencies for the initial window. Note that $e_{11(\tau_1)} = -\,0.03$ for AUD denotes that the opposite position should be taken for AUD whenever portfolio trading is executed (see Table 6). When generating training data conditional on $\mathbf{X}_{(\tau_1)}$, we apply Eqs (8) and (9) with $k = 1$ and $S = 0.03$ (i.e.,

$$h(u_i) = 1, \text{ if } u_i = \frac{g_{i(\tau_1)} - g_{i-1(\tau_1)}}{g_{i-1(\tau_1)}} > 0.03,$$

$$h(u_i) = -1, \text{ if } u_i = \frac{g_{i(\tau_1)} - g_{i-1(\tau_1)}}{g_{i-1(\tau_1)}} < -0.03,$$

where $g_{i(\tau_1)} = \mathbf{x}_{i\cdot(\tau_1)}\mathbf{e}_{1(\tau_1)}$ and $\mathbf{x}_{i\cdot(\tau_1)} = (x_{i1(\tau_1)}, \ldots, x_{i9(\tau_1)})$ for $i = 1, \ldots, 55$). Thus, the response is

$$Y_{(\tau_1)} = (h(u_1), \ldots, h(u_n))_{\tau_1}.$$

$k = 1$ and $S = 0.03$ are selected using previous simulated results. For predictor, we used relative position ($\mathbf{z}_{\cdot 1(\tau_1)}$), RSI ($\mathbf{z}_{\cdot 2(\tau_1)}$), momentum ($\mathbf{z}_{\cdot 3(\tau_1)}$) and MACD ($\mathbf{z}_{\cdot 4(\tau_1)}$) applied to $g_{i(\tau_1)} = \mathbf{x}_{i\cdot(\tau_1)}\mathbf{e}_{1(\tau_1)}$. In fact, the 55×4 matrix $\mathbf{Z}_{(\tau_1)}$ is given by

$$\mathbf{Z}_{(\tau_1)} = (\mathbf{z}_{\cdot 1(\tau_1)}, \mathbf{z}_{\cdot 2(\tau_1)}, \mathbf{z}_{\cdot 3(\tau_1)}, \mathbf{z}_{\cdot 4(\tau_1)}),$$

where $\mathbf{z}_{\cdot\nu(\tau_1)} = (z_{1\nu(\tau_1)}, \ldots, z_{n_1\nu(\tau_1)})$ and $z_{i\nu(\tau_1)} = \delta_{\nu(\tau_1)}(g_{i(\tau_1)}, \ldots, g_{i-m+1(\tau_1)}) \in R$ for $i = 1, \ldots,$ 55, $\nu = 1, 2, 3$ and $m = 55$, and $z_{i4(t)} = \delta_{4(t)}(g_{i(t)}, \ldots, g_{i-m+1(t)}; m_1, m_2) \in R$ for $m_1 = 12$ and $m_2 = 26$. Thus, we obtained the training dataset $(\mathbf{Z}_{(\tau_1)}, Y_{(\tau_1)})$, which induced the mapping

$$\hat{f}_{\tau_1} : \mathbf{Z}_{(\tau_1)} \to Y_{(\tau_1)},$$

Table 7
Summary statistics for FXTA trading experiment

Classifier	Number of trading	Winning ratio	Max	Min	Average holding period
LR	84	69.0%	4.8%	−4.9%	12.1 days
DT	88	71.6%	5.0%	−4.4%	9.6 days
ANN	81	77.8%	4.9%	−4.4%	11.6 days

Table 8
Annual return rates of FXTA over 5 years and their averages

Year	Earning rate (%)		
	LR	DT	ANN
2005	4.08	5.99	7.31
2006	7.41	8.44	15.07
2007	6.29	7.44	10.86
2008	4.15	3.96	8.18
2009	8.97	12.90	18.75
Average	6.18	7.75	12.03

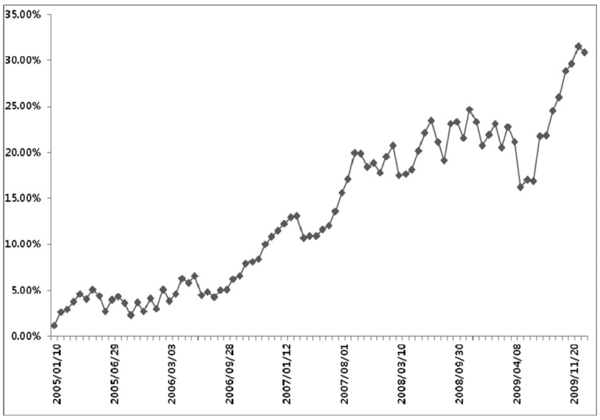

Fig. 5. Cumulative return rate for FXTA with LR.

and $\hat{f}_{\tau_1}$ was used for the testing period of the initial window. Because the window moves based on the moving window scheme (see Fig. 4), we generated $\hat{f}_{\tau_l}$ for $l = 2, 3 \ldots, L$. The last L was 12/31/2009.

To choose an appropriate classifier for our FXTA, we implemented LR, DT, and ANN for induced mappings and calculated their cumulative return rates. Cumulative return rate is given by

$$cr_t = \sum_{i=1}^{t} \sum_{j=1}^{9} w_{ij} \log \frac{x_{ij}}{x_{i-1,j}}, \tag{15}$$

where w_{ij} (x_{ij}) is the weight (exchange rate) for the jth currency at time i. Note that w_{ij} are constants depending on j for a given moving window. Figures 5–7 illustrate the calculated cumulative return rates when LR, ANN, and DT are employed. All methods showed increasing cumulative return rates in the following order: ANN, DT, and LR. Note that the cumulative return rates shown in Figs 5–7 were calculated only under conditions of portfolio trading. Table 7 provides the summary statistics of the trading results for each classifier. In fact, winning ratio (i.e., the number of profitable trades out of total trading), maximum and minimum earning rates as well as average holding period are provided. Earning rates were calculated by

$$er_\tau = \sum_{j=1}^{9} w_{\tau j} \frac{x_{\tau j} - x_{b(\tau),j}}{x_{b(\tau),j}}, \tag{16}$$

where τ is the sell time by FXTA, and $b(\tau)$ is the previous corresponding buying time by FXTA (i.e.,

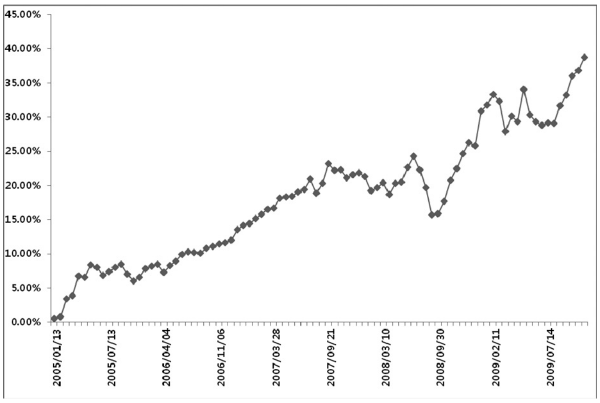

Fig. 6. Cumulative return rate for FXTA with DT.

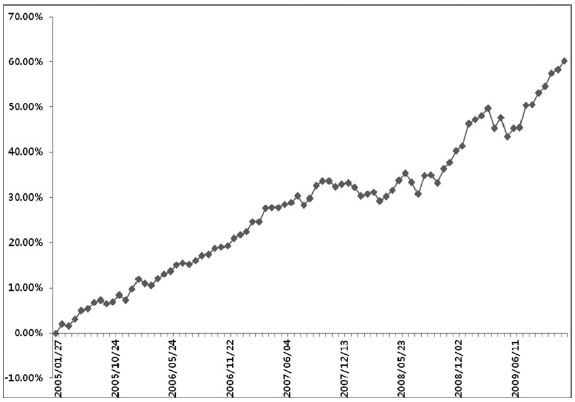

Fig. 7. Cumulative return rate for FXTA with ANN.

$b(\tau) < \tau$). Finally, Table 8 provides annual earning rates over 5 years and the overall average earning rate. Tables 7 and 8 as well as Figs 5–7 clearly show that ANN have an edge over DT and LR. Figures 5–7 show that the cumulative return rates of ANN surpass DT and LR. Table 7 shows that ANN has a higher winning ratio (77.8%) than DT and LR (71.6% and 69.0%, respectively). The annual earning rate of Table 8 also confirms that ANN easily surpasses the other methods (12.03% vs. 7.75% and 6.18%, respectively). These results suggest that FXTA performs better with ANN, and this method is better suited for training nonlinear movements of data.

5. Concluding remarks

This study provides an FXTA for obtaining profits via multi-currency trading. Because FX markets are usually dominated by USD, it makes sense that the FXTA uses the correlations among multi-currencies via PCA. In other words, PCA determines the first PC that explains much of the movements of the multi-currencies except USD, and then the projection of the multi-currencies on the first PC is used as an imaginary key currency. We applied a standard method from stock market trading (i.e., the induced classifier construction method) to the imaginary key currency. The cointegration movement of the multi-currencies and USD is essential for this approach, which was verified in our unreported work. We believe that our approach for developing an FXTA provides a basic procedure that currency-trading algorithms seeking to apply a stock market trading tool must follow.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A2A1A11052670). T.Y. Kim's work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2011-314-C00046) and by the 2012 Bisa research grant of Keimyung University.

References

[1] A.Y. Huang, S.-P. Peng, F. Li and C.-J. Ke, Volatility forecasting of exchange rate by quantile regression, *International*

Review of Economics and Finance **20**(4) (2011), 591–606.

[2] A.W. Lo, H. Mamaysky and J. Wang, Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation, *The Journal of Finance* **55**(4) (2000), 1705–1770.

[3] Bank for International Settlements (BIS), *Triennial Central Bank Survey: Foreign Exchange and Derivatives Market Activity in 2007*, Basel, Switzerland: Ban for International Settlements Press and Communications, 2007.

[4] C.J. Neely, P.A. Weller and R. Dittmar, Is technical analysis in the foreign exchange market profitable? A genetic programming approach, *Journal of Financial and Quantitative Analysis* **32**(4) (1997), 405–426.

[5] D. Fenn, Network communities and the foreign exchange market, Ph D thesis, University of Oxford, 2010.

[6] D.O. Cushman, A portfolio balance approach to the canadian – U.S. exchange rate, *Review of Financial Economics* **16**(3) (2007), 305–320.

[7] F.R. Fernando, S.R. Simon and A.F. Julian, Exchange-rate forecasts with simultaneous nearest-neighbor methods: Evidence from the EMS, *International Journal of Forecasting* **15** (1999), 383–392.

[8] G.S. Jang, F. Lai, B.W. Jiang, C.C. Pan and L.H. Chien, Intelligent stock trading system with price trend prediction and reversal recognition using dualmodule neural networks, *Applied Intelligence* **3**(3) (1993), 225–248.

[9] H.B. Hwarng, Insights into neural-network forecasting of time series corresponding to ARMA(p, q) structures, *Omega* **29**(3) (2001), 273–289.

[10] J. Beran and D. Ocker, SEMIFAR forecasts, with applications to foreign exchange rates, *Journal of Statistical Planning and Inference* **80**(1–2) (1999), 137–153.

[11] J.J. Ahn, S.J. Lee, K.J. Oh and T.Y. Kim, Intelligent forecasting for financial time series subject to structural changes, *Intelligent Data Analysis* **13**(1) (2009), 151–163.

[12] J.R. Quinlan, Induction of decision tree, *Machine Learning* **1**(1) (1986), 81–106.

[13] K. Hornik, M. Stinchcombe and H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* **2**(5) (1989), 359–366.

[14] K.J. Oh, T.Y. Kim and C. Kim, An early warning system for detection of financial crisis using financial market volatility, *Expert Systems* **23**(2) (2006), 83–98.

[15] K.S. Shim, J.J. Ahn and K.J. Oh, Multi-currencies portfolio strategy using principal component analysis and logistic regression, *Journal of the Korean Data and Information Science Society* **23**(1) (2012), 151–159. (in Korean).

[16] L. Breiman, J.H. Friedman, R. Olshen and C.J. Stone, *Classification and Regression Trees*, Belmont, CA: Wadsworth, 1984.

[17] L. Laloux, P. Cizeau, J.-P. Bouchaud and M. Potters, Noise dressing of financial correlation matrices, *Physical Review Letters* **83** (1999) 1467–1470.

[18] F. Hayashi, *Econometrics*, Princeton University Press, New Jersey, 2000

[19] M. Kühl, Cointegration in the foreign exchange market and market efficiency since the introduction of the Euro: Evidence based on bivariate cointegration analyses, CEGE discussion paper 68.

[20] M.A.H. Dempster and C.M. Jones, The profitability of intra-day FX trading using technical indicators, Research papers in management studies, 1–64, 2000.

[21] M.A.H. Dempster, T.W. Payne, Y. Romahi and G.W.P. Thompson, Computational learning techniques for intraday FX trading using popular technical indicators, *IEEE Transaction on Neural Networks* **12**(4) (2001), 744–754.

[22] M.A.H. Dempster and V. Leemans, An automated FX trading system using adaptive reinforcement learning, *Expert Systems with Applications* **30**(3) (2006), 543–552.

[23] P.-F. Pai, S.-Y. Chen, C.-W. Huang and Y.-H. Chang, Analyzing foreign exchange rates by rough set theory and directed acyclic graph support vector machines, *Expert Systems with Applications* **37**(8) (2010), 5993–5998.

[24] R. Gencay, Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules, *Journal of International Economics* **47**(1) (1999), 91–107.

[25] V. Cherkassky and F. Mulier, *Learning from Data: Concepts, Theory and Methods*, New York: Wiley, 1998.

[26] W. Brock, J. Lakonishok and B. LeBaron, Simple technical trading rules and the stochastic properties of stock returns, *The Journal of Finance* **47**(5) (1992), 1731–1764.

[27] Y. Zhang and L. Wu, Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network, *Expert Systems with Applications* **36**(5) (2009), 8849–8854.