Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2013, 5(12):579-586

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Forecasting of core inflation

Linyun Zhang*, Jinchang Li and Bo Liu

Department of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou, China

---

ABSTRACT

Nonparametric support vector regression (SVR) approach has recently attracted much attention as a new technique for forecasting in economics. The chief advantage of this new approach is that such models are free from the large sample assumption that is often adopted to make the traditional models. In addition, SVR inherits the strong fitting ability of neural network and its minimum structural risk endows it stronger predictive ability than neural network. This paper uses SVR approach to predict core inflation series estimated with variance paring, compares forecast result with that obtained by maximum likelihood estimation (MLE) and back propagation (BP) neural network. It suggests that SVR presents the best predictive ability. On this basis, this paper uses SVR approach conducts trend extrapolation on core inflation and concludes that China will continue with inflation in the next year but such inflation will fall slightly.

Key words: Core inflation, validity, support vector regression, neural network

---

INTRODUCTION

Inflation is a core element that determines harmonious and healthy development of national economy. Accurate and reliable forecast of inflation can not only provide the basis to government decision-making departments and macro economic decision-making of the central bank but also provide reference to economic activities of enterprises and residents as well as anticipated inflation formation.

In macroeconomic analysis, people generally use CPI, GDP deflator, commodity retail price index, producer price index, export price index and import price index to measure inflation. However, in view of practical significance, data timeliness, reliability and availability, CPI is the mostly widely used inflation index. As CPI is often affected by abnormal price fluctuations of some commodities in a short term, it inevitably harbors "noise" and cannot faithfully reflect the relationship between aggregate supply and aggregate demand, which may be misleading in economic situation. For this reason, Eckstein [1] put forward core inflation and defined it as “trend rate of increase of the price of aggregate supply”, namely potential inflation regardless of the influence by temporary factors and reflects. In terms of prediction, core CPI, an index that eliminates volatility and reflects the general trend of price change, is more suitable to forecast inflation rate.

Inflation rate can be forecast through two methods, namely statistical survey analysis and measurement model. The first method is to average forecast on people. Its basic approach is to select particular samples and make survey on the samples to get forecast on inflation in the future and then average sample forecast data to get the predicted value of future inflation, such as Livingston survey and the Michigan survey. The Livingston survey is the oldest survey of economists' expectations. It summarizes the forecasts of economists from industry, government, banking, and academia, while the Michigan survey measures the consumers’ confidence. They are both devised in the 1940s. The Federal Reserve Bank of Philadelphia has taken the responsibility for conducting the Livingston survey since 1990, and Thomson Reuters with the University of Michigan publish the consumer confidence index monthly. There also many scholars use this method in China, such as [2] [3]. The second method aims to establish one or more models based on inflation historical data or other economic variable data to predict inflation. Commonly used models include auto-regression model, structure model and simultaneous model. Auto-regressive models use the inflation historical data to forecast future data. As the most basic models, auto-regressive models are often used in comparison with other methods. The key characteristic of structure models is the use of influencing variables. The structure models apply the influencing variables for the equation and then get the inflation prediction from the equation. Linear equations, grey models and nonparametric regression models are all included. Simultaneous models' basic thinking is viewing the macro economy as a whole. In this system, variables including inflation interact and inter-depend on each other. [4] [5] [6] are all the second type of methods. The merit of survey analysis lies in faithful and accurate reflection of market forecast while the demerit is the complete dependence on sample and quantity. In addition, it is hysteretic from data acquisition to forecast results. Measurement model can generate forecasts more conveniently. But for forecast requiring high accuracy, traditional parametric model is likely to create error. Xue [7] uses nonparametric neural network measurement model to predict inflation, the results show high precision. Theoretically, there is a new nonparametric support vector regression (SVR) more powerful than neural network prediction ability. SVR inherits the strong fitting ability of neural network and its minimum structural risk endows it stronger predictive ability than neural network [8].

Currently, very few scholars apply SVR to inflation. Therefore, in view of foresight, convenience and accuracy, SVR is suitable to forecast inflation. This paper makes innovative work from the following aspects. (1) A new variance paring approach is adopted to estimate the core CPI. (2) Valid text suggests the series estimated by variance paring is valid. (3) This paper compares the predictive ability between SVR, MLE and BP neural network. The result shows SVR outperformed MLE and BP neural network. (4) SVR is proposed to conduct the trend of core CPI.

Proposed Method

Support vector regression has support vector machine as its theoretical basis. Support vector machine is an algorithm as the extension of neural network with unique theoretical design - structural risk minimization. This algorithm was first proposed by mathematician Vapnik [9]. SVM was first put forward to solve problems of classification and identification, which is SVM for Classification (SVC). With strong generalization ability, it was then extended to regression, which is SVM for regression (SVR). The structure of support vector regression consists of input layer, hidden layer and output layer. When the training set finishes nonlinear transformation between the input layer and hidden layer, linear regression can be made in the output space. Input variables and output variables are in an unknown mapping function relationship g(x). When training with the training set, it is found f(x) is approximate to g(x). Then f(x) is used to make prediction. Formula (1) is the expression of decision function. Fig.1 is the structure diagram of SVR.

$$f(x)=\sum_{l=1}^{l} w_l \varphi_l(x)+b=w^T \phi(x)+b \tag{1}$$

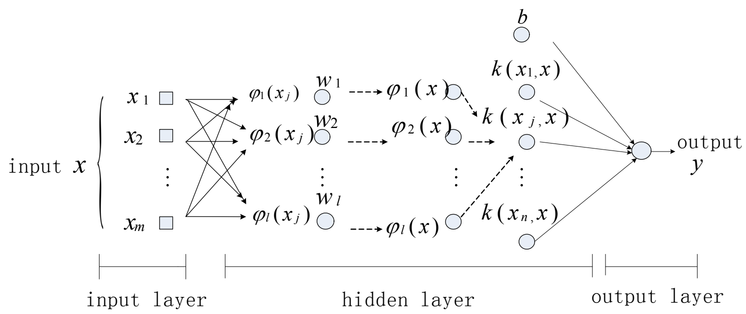

Figure1. The structure diagram of SVR

$\phi(x)=[\phi_1(x),...,\phi_l(x)]^T$ is the nonlinear transfer equation from input layer to hidden layer. $w=[w_1,...,w_l]^T$ stands for the linear weight from hidden layer to output layer. b means the threshold. Parameters w and b must be estimated based on the training set in order to get f(x). The variables w and b are calculated by minimizing the regularized risk function (2),

$$\text{Minimize } \frac{1}{2}\|w\|^2+\frac{C}{n}\sum_{i=1}^{n} L_\varepsilon(y_i, f(x_i)) \tag{2}$$

$$L_\varepsilon(y_i, f(x_i))=\begin{cases}|y_i-f(x_i)|-\varepsilon, & |y_i-f(x_i)|\ge\varepsilon \\ 0, & \text{otherwise}\end{cases} \tag{3}$$

$\|w\|^2/2$ is a measure of the function flatness, $(1/n)\sum_{i=1}^{n} L_{\varepsilon}(y_i, f(x_i))$ stands for empirical risk determined by ε insensitive equation. C is considered to specify the trade-off between the empirical risk and the model flatness. Minimization of the sum of the two means SVR can not only minimize empirical risk but also maximize promotion ability to get balanced optimal solution of the two. Function $L_{\varepsilon}(y_i, f(x_i))$ is called ε -insensitive loss function. ε is error parameter of linear ε - insensitive loss function. When forecast error is less than ε , the loss can be ignored. Such loss function is unique to SVR method with the aim to simplify sample information quantity. As a result, even high sample dimension won't cause great trouble to computation and storage.

After slack variables ξ_i and ξ_i' introduced to represent the distance from actual values to the corresponding boundary values of ε -tube, (2) can be transformed to the primal function (4) as,

Minimize $\frac{1}{2}\|\mathrm{w}\|^2 + \mathrm{C}\sum_{i=1}^{n}(\xi_i + \xi_i')$ (4)

Subject to

$$w^T\varphi(x) + b - y_i \le \varepsilon + \xi_i\ , \quad i=1,2,\ldots,n \tag{5}$$

$$\mathrm{y}_i - w^T\varphi(x) - b \le \varepsilon + \xi_i', \quad i=1,2,\ldots,n \tag{6}$$

$$\xi_{\mathrm{i}} \ge 0, \quad \xi_{\mathrm{i}}' \ge 0, \quad i=1,2,\ldots,n \tag{7}$$

Bringing all constraints into the objective function, we can construct the Lagrangean function (8) of the primal problem.

$$L(w,b,\xi_i,\xi_i',\alpha_i,\alpha_i') = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i') - \sum_{i=1}^{n}(\eta_i\xi_i + \eta_i'\xi_i')$$
$$-\sum_{i=1}^{n}\alpha_i(\varepsilon + \xi_i + y_i - w^T\varphi(x) - b) \tag{8}$$
$$-\sum_{i=1}^{n}\alpha_i'(\varepsilon + \xi_i' - y_i + w^T\varphi(x) + b$$
$$\eta_i \ge 0, \eta_i' \ge 0, \alpha_i \ge 0, \alpha_i' \ge 0,$$

According to the Karush-Kuhn-Tucker [10] optimality conditions, the partial derivatives of L with respect to the primal variables w, b, ξ_i, ξ_i' equal to zero:

$$\frac{\partial L}{\partial w} = 0 \rightarrow w - \sum_{i=1}^{n}(\alpha_i' - \alpha_i)\varphi(x) = 0$$

So $w = \sum_{i=1}^{n}(\alpha_i' - \alpha_i)\varphi(x)$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{n}(\alpha_i - \alpha_i') = 0 \tag{9}$$
$$\frac{\partial L}{\partial \xi_i} = 0 \rightarrow C - \eta_i - \alpha_i = 0$$
$$\frac{\partial L}{\partial \xi_i'} = 0 \rightarrow C - \eta_i' - \alpha_i' = 0$$

We obtain the dual problem by substituting (9) into (8). Specifically, the dual problem is as follows:

Max $$-\varepsilon\sum_{i=1}^{n}(\alpha_i' + \alpha_i) + \sum_{i=1}^{n} y_i(\alpha_i' - \alpha_i) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i' - \alpha_i)(\alpha_j' - \alpha_j)\varphi^{\mathrm{T}}(x_i)\varphi(x_j) \tag{10}$$

Subject to $\sum_{i=1}^{n}(\alpha_i^{'}-\alpha_i)=0 \quad \alpha_i^{'},\alpha_i \in (0,C)$

We can obtain the Lagrange multipliers $\alpha_i^{'}$ and α_i from this dual problem. Solving (9) for w, we have:

$$w^* = \sum_{i=1}^{n}(\alpha_i^{'}-\alpha_i)\varphi(x_i) \tag{11}$$

Using (11), we obtain the following expression for the regression function (1):

$$f(x) = w^{*\,\mathrm{T}}\cdot\varphi(x) + b^* = \sum_{i=1}^{n}(\alpha_i^{'}-\alpha_i)\varphi^T(x_i)\varphi(x_i)+b^* \tag{12}$$

We compute the optimal value of b from the complementary slackness conditions:

$$\begin{aligned}&\alpha_i(\varepsilon+\xi_i+y_i-w^T\varphi(x)-b)=0\\&\alpha_i^{'}(\varepsilon+\xi_i^{'}-y_i+w^T\varphi(x)+b)=0\\&\eta_i\xi_i=0\\&\eta_i^{'}\xi_i^{'}=0\end{aligned} \tag{13}$$

We obtain the b:

$$b^* = y_i - \sum_{i=1}^{n}(\alpha_i^{'}-\alpha_i)\varphi^T(x_i)\varphi(x_i)+\varepsilon \quad \text{for}\ \alpha_i \in (0,C) \tag{14}$$

$$b^* = y_i - \sum_{i=1}^{n}(\alpha_i^{'}-\alpha_i)\varphi^T(x_i)\varphi(x_i)-\varepsilon \quad \text{for}\ \alpha_i^{'} \in (0,C) \tag{15}$$

After exploiting the optimality constraints $w*$ and b^*, the decision function given by (1) has the following explicit form (16):

$$f(x) = \sum_{i=1}^{n}(\alpha_i^{'*}-\alpha_i^*)K(x_i,x)+b^* \tag{16}$$

In (16), $\alpha_i^{'*}$ and α_i^* have been figured out, $K(x_i,x)$ is defined as the kernel function. Any function that satisfies Mercer's condition can be used as the kernel function [11]. Commonly used kernel functions include linear kernel function, polynomial kernel function, radial basis function (RBF) and sigmoid kernel function. The difficulty for support vector regression method lies in the selection of kernel function and parameters. There hasn't been a universally recognized method for selection of the two. Among various kernel functions, RBF kernel is not only easy to perform but also effectively maps the training set to infinite space layer on a nonlinear basis, which is very suitable to deal with non-linear relation [12]. For this reason, this paper adopts software default RBF kernel. As for optimal selection of parameters, this paper uses sensitivity method [13]. Parameters need include loss function parameter ε, penalty factor C and kernel function parameter (RBF parameters shall select γ).

RESULTS AND DISCUSSION

This paper uses variance trimming method to estimate core inflation and support vector regression to predict period core inflation, and compares the result with that based on maximum likelihood estimation (MLE) and BP neural network. Finally, support vector regression extrapolation is used to predict CPI in 12 periods and get the inflation trend in China.

3.1. Core Inflation Estimate and Validity Test

This paper uses variance trimming [14] to estimate core inflation with eight group indexes of CPI between January 2001 and June 2013: household appliances and services index, transportation and communication index, housing index, entertainment and educational products and services index, medical care and personal products index, clothing index, alcohol and tobacco products index, food products index. Data source: Phoenix Net Data Center. Fig.2 shows core CPI between January 2001 and June 2013.

With core inflation series between January 2001 and June 2013, validity test is carried out on three items. Marquesa et al. [15] believed that a valid measure of core inflation should satisfy three conditions:
(1) $x_t = \pi_t - \pi_t^*$ is a stationary series, where π_t is CPI series and π_t^* for core CPI series.

(2) There exists an error correction mechanism in $\Delta\pi_t$, that is, π_t^* is the attractor of π_t, if $\pi_t > \pi_t^*$, then the future π_t will fall, if $\pi_t < \pi_t^*$, then the future π_t will rise, and this requires the error correction coefficient γ to be less than zero in (17):

$$\Delta\pi_t = \sum_{i=1}^{m} \alpha_i \Delta\pi_{t-i} + \sum_{i=1}^{n} \beta_i \Delta\pi_{t-i}^* + \gamma(\pi_{t-1} - \pi_{t-1}^*) + \varepsilon_t \tag{17}$$

(3) π_t^* is strongly exogenous, that is, the lagged difference terms of π_t have no effect on π_t^* under the premise that inflation π_t is not the attractor of core inflation π_t^*, and this requires the coefficient of error correction terms γ and the coefficients of π_t 's lagged difference terms a_i (i=1, 2, ……, r) to be equal to zero in (18):

$$\Delta\pi_t^* = \sum_{i=1}^{r} a_i \Delta\pi_{t-i} + \sum_{i=1}^{s} b_i \Delta_{t-i}^* + \lambda(\pi_{t-1} - \pi_{t-1}^*) + \eta_t \tag{18}$$

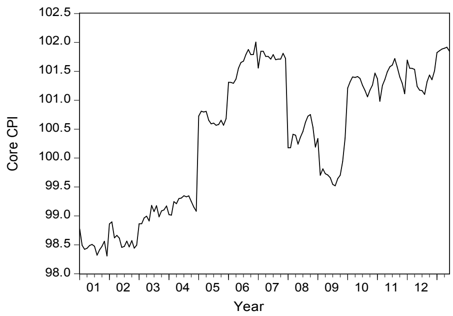

Figure 2. The core CPI between January 2001 and June 2013

This paper verifies core inflation series estimated by variance trimming in three items. x_t is a stationary series (p<0.1), γ is less than zero (p<0.1), λ and a_i equal zero (P_λ >0.1, P_{ai} >0.1) (see Table1 for the results). Lag phase is verified according to AIC criterion. $\Delta\pi_t$ harbors error correction mechanism with three periods lagged behind. π_t^* is strongly exogenous model with one period lagged behind. The results show that variance trimming calculated core inflation is valid core inflation measurement.

Table 1: Validity test of core inflation series

	x_t has a unit root	γ=0	$\lambda = a_1 = 0$
core CPI series	T=-3.287 P=0.017	γ=-0.041 P=0.028	P_λ=0.800 P_{a1}=0.298

3.2. MLE, BP Neural Network and SVR Core Inflation Period Forecast
MLE-based linear regression model is

$$y_t = \alpha + \beta_1 y_{t-1} + \varepsilon_t \text{, t=1,2,3…,T} \tag{19}$$

Nonlinear regression model estimated by SVR and BP neural network is

$$y_t = g(y_{t-1}) + \varepsilon_t \text{, t=1,2,3…,T} \tag{20}$$

Where, y_t stands for core CPI of period t, lag period is made one according to AIC criterion.

In order to verify the advancement of SVR prediction, this paper makes forecast for three times. In the first time, SVR, BP neural network and linear regression are used to model for core CPI data from January 2001 to June 2013, and then core CPI from January 2001 to June 2013 is fitted based on the model. In the second time, data from January 2001 to December 2012 is used for modeling, and then core CPI from January to June 2013 is forecast based on the model. In the third time, data from January 2001 to June 2012 is used for modeling, and then core CPI from July 2012 to June 2013 is forecast based on the model. SVR method is achieved by Matlab2010a which includes five steps.

(1) Data preprocessing. Raw data shall under go dimensionality reduction or normalization. This step is carried out in line with data quality as appropriate.

(2) Determination of training set. This paper has carried out three forecasts. The first training set contains core CPI data from January 2001 to June 2013. The second training set contains core CPI data from January 2001 to December 2012. The third training set contains core CPI data from January 2001 to June 2012.

(3) Determination of kernel function and optimal parameters. Three forecasts all use RBF kernel with loss function parameter ε of 0.1. Parameters are all obtained based on sensitivity analysis optimization with C as 620, 400 and 340, γ as 720, 8 and 16 respectively.

(4) Fitting or prediction results.

(5) De-normalization for final forecast result. This step corresponds to the first step. If normalization is not carried out, de-normalization is not necessary.

In this paper, neural network method adopted is counter-propagation neural network, with three-neuron nonlinear hidden layer and one-neuron linear output layer. Learning rate parameter is made 0.1.

Fig. 3 shows that SVR, MLE and BP neural network forecast the core CPI from January 2001 to June 2013. Fig. 4 shows that SVR, MLE and BP neural network forecast the core CPI from January 2013 to June 2013. Fig. 5 shows that SVR, MLE and BP neural network forecast the core CPI from July 2012 to June 2013.Table 3 shows the comparison of the predicted value and real value with three methods. As for forecast results, two quantitative comparison standards are used root mean square error (RMSE) and mean absolute error (MAE) [16]. Smaller RMSE and MAE indicate more closeness between the predicted values and the real value, thus a better predictive effect.

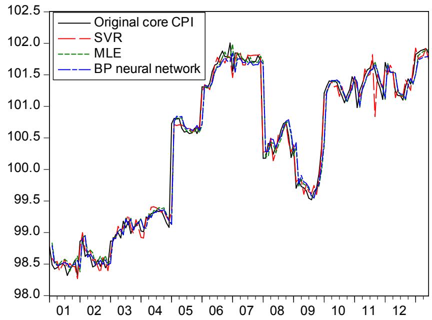

Figure 3. The comparison of forecasting series (from January 2001 to June 2013)

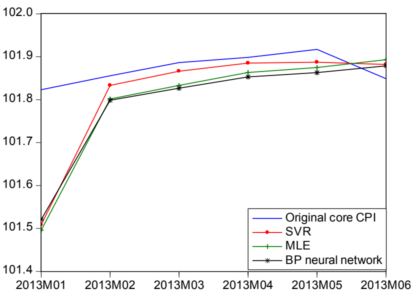

Figure 4. The comparison of forecasting series (from January 2013 to June 2013)

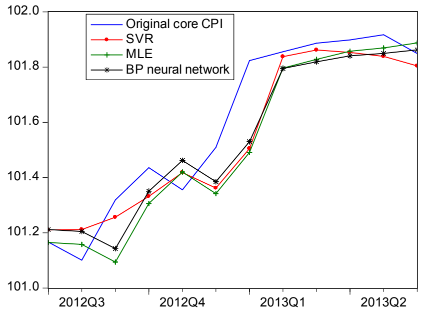

Figure 5. The comparison of forecasting series (from July 2012 to June 2013)

Table2: Comparison of three forecasts

	Fitting		12 periods		6 periods	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
MLE	0.0658	0.1430	0.0015	0.1020	0.0008	0.0927
BP	0.0629	0.1403	0.0012	0.1002	0.0007	0.0914
SVR	0.0258	0.1004	0.0011	0.0885	0.0007	0.0722

3.3. SVR–Based Core Inflation Forecast

In respect of fitting and forecast as shown in Table2, RBF kernel SVR outshines BP neural network and linear regression in both fitting and prediction ability with RMSE and MAE less than that of linear regression and neural network. In view of implementation process, SVR inherits neural network in no requirement for normal assumption and large sample, overcomes unstable forecast results of neural network in extrapolation, strong trend and unsatisfactory precision of linear regression. Therefore, SVR performs excellently in forecast. Based on empirical SVR forecasting ability, this paper uses SVR method to forecast core CPI of 12 periods from July 2013 to June 2014. The model used is SVR nonparametric model trained with core CPI data from January 2001 to December 2012. The forecast method is recursive prediction where the training model is used to forecast core CPI in July 2013 which is then used as independent variable to forecast core CPI in August, 2013. This process is repeated 12 times until the core CPI in June 2014, core CPI estimation series and CPI comparison are forecast and worked out. Fig.6 is the comparison of core CPI (including forecast of 12-period) series and CPI series.

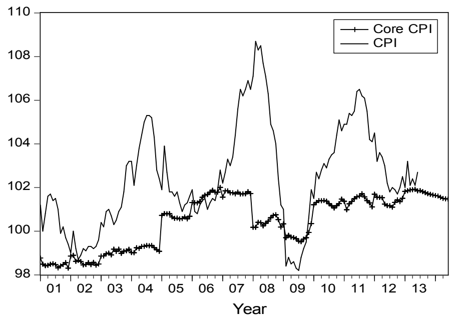

Figure 6. The comparison of core CPI (including forecast of 12-period) series and CPI series

CONCLUSION

It can be seen from Fig. 6 that core CPI and CPI overall trend are close to each other. In years of high CPI, core CPI indicates inflation. In years of low CPI, core CPI indicates deflation. From the perspective of core CPI, domestic price level had experienced four stages, namely deflation from 2001 to 2004, inflation from 2005 to 2008, deflation in 2009, and inflation from 2010 to 2013. It can be seen directly that core CPI volatility is less that of CPI and core CPI series are more stable than CPI series, which conforms to the definition of core inflation as inflation regardless of temporary fluctuations. Based on core CPI forecasting data, China will continue with inflation next year, remaining almost the same with the first half of 2013 but inflation rate will fall down slightly.

REFERENCES

[1] O Eckstein. *Core inflation*, 1st Edition, Englewood Cliffs, Prentice Hall, **1981**, 7-12.
[2] ZY Xiao; YB Chen. *Financial Research*, **2004**, (11), 1-18.

[3] ZY Xiao; YM Yao; SL Tang. *Statistical Research*, **2011**, 28(3), 8-14.
[4] E Nakamura. *Economics Letters*, **2005**, 86(3), 373-378.
[5] P McAdam; P McNelis. *Economic Modelling*, **2005**, 22(5), 848-867.
[6] C Broto. *Economic Modelling*, **2011**, 28(3), 1424-1434.
[7] YG Xue. *Commercial Times*, **2010**, (18), 71-72.
[8] KR Muller; AJ Smola; G Ratsch; B Scholkopf; J Kohlmorgen; V Vapnik. *Lecture Notes in Computer Science*, **1997**, (1327), 999-1004.
[9] V Vapnik. *The Nature of Statistical Learning Theory*, 2nd Edition, New York: Springer, **2000**, 267-287.
[10] HW Kuhn; AW Tucker. *Nonlinear Programming, Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability*, 1st Edition, Berkeley: University of California, **1951**, 481-492.
[11] LJ Cao; FEH Tay. *Neural Computing and Applications*, **2001**, (10), 184–192.
[12] WC Hong. *Energy Conversion and Management*, **2009**, 50(1), 105-117.
[13] SY Chen. *The Journal of Quantitative and Technical Economics*, **2007**,(5), 142-150.
[14] M Pedersen. *German Economic Review*, **2009**, 10(2), 139-164.
[15] CR Marquesa; PD Neves; LM Sarmentoa. *Economic Modelling*, **2003**, (20), 765-775.
[16] S Moshiri; N Cameron. *Journal of Forecasting*, **2000**, (19), 201-217.