International Journal of Physical Sciences Vol. 7(38), pp. 5460-5467, 9 October, 2012
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS12.427
ISSN 1992 - 1950 ©2012 Academic Journals

Full Length Research Paper

Forecasting the output of Taiwan's integrated circuit (IC) industry using empirical mode decomposition and support vector machines

Kwo-Liang Chen[1], Ching-Chiang Yeh[2]* and Tz-ling Lu[3]

[1]Department of Industrial Engineering and Management, China University of Science and Technology, No.245, Sec. 3, Academia Rd., Nangang Dist., Taipei City, Taiwan 115, R.O.C.
[2]Department of Business Administration, National Taipei College of Business, No.321, Sec. 1, Ji-Nan Rd., Zhongzheng District, Taipei City, Taiwan 10051, R.O.C.
[3]Department of Business Administration, Soochow University, No.56. Kueiyang Street, Section 1, Taipei, Taiwan 100, R.O.C.

Accepted 10 August, 2012

As the production values in the integrated circuit (IC) industry are inherently nonlinear and non-stationary, it is regarded as one of the most challenging tasks for practitioners and academics. This study proposed a hybrid methodology by combining empirical mode decomposition (EMD) and support vector regression (SVR) in production values forecasting. The proposed approach first uses EMD, which can adaptively decompose the complicated raw data into a finite set of intrinsic mode functions (IMFs) and a residue. After identifying the IMF components, residue are then modeled and forecasted using SVR. The final forecasting value can be obtained by the sum of these prediction results. Experimental results show that the proposed approach outperforms the SVR model without EMD preprocessing.

Key words: Integrated circuit (IC) industry, production values forecasting, empirical mode decomposition, support vector regression.

INTRODUCTION

The IC industry in Taiwan is one of the main factors for Taiwan's economy growth. Since the strong demand from consumer electronics industry, the production value of Taiwan's IC industry came to new Taiwan dollar 1 trillion 768,600 million in 2010. With the change and globalization of the environments and rapid new technology development among the enterprises in the IC industry in Taiwan, the entire scenario has become ever more fierce. Forecasting of total production output in IC industry is useful for decision makers to prepare marketing strategies/production capacity planning and for financial institutions to make investment decisions.

Few forecasting models have been developed for the IC industry until recently. Three grey theory-based models have been presented to forecast the output of IC industry in Taiwan (Hsu, 2003; Hsu and Wang, 2007; Wang and Hsu, 2008). Generally, production values in the IC industry change over time. The changes thus can be treated as a time series process. However, there exist some specific characteristics in the IC industry, such as capital-intensive, short product life cycle, production technology rapidly changed over time and severe competition. Therefore, production values are inherently nonlinear and non-stationary in the IC industry. The non-stationary characteristic implies that the statistical properties of the data change over time. The main cause of this is the effect of various business and economic

*Corresponding author. E-mail: ychinc@webmail.ntcb.edu.tw. Tel: +886 2 23226161. Fax: +886 2 23226404.

cycles while the nonlinear characteristic makes linear parameter models information difficult to use (Hussain et al., 2008; Yaser and Atiya, 1996). Therefore, production values forecasting is regarded as one of the most challenging tasks for practitioners and academics.

Support vector machines (SVM), developed by Vapnik (1995), is a novel machine learning technique and is gaining popularity due to many attractive features and excellent generalization performance on a wide range of problems. It captures geometric characteristics of feature space without deriving weights of networks from the training data. Also, SVM embodies the structural risk minimization principle (SRM), which has been shown to be superior to traditional empirical risk minimization principle (ERM) employed by conventional neural networks. SRM minimizes an upper bound of generalization error as opposed to ERM that minimizes the error on training data. SVM guarantee global optima, where as neural networks face the risk of getting stuck in local optima and are not guaranteed to achieve global optima (Tay and Cao, 2001, 2003). In particular, with the introduction of Vapnik's ε- insensitivity loss function, the regression model of SVMs, called support vector regression (SVR), has also been receiving increasing attention to solve nonlinear regression estimation problems (Vapnik, 1995, 1999), which have been shown to exhibit excellent performance. Recently, SVR has been successfully adopted to solve forecasting problems in many fields such as forecast tourist arrivals, engine reliability prediction, wind speed prediction (Mohandes et al., 2004), financial time series (stocks index and exchange rate) forecasting (Cao, 2003; Cao and Tay, 2001; Ince and Trafalis, 2006; Kim, 2003; Tay and Cao, 2001, 2003; Thissen et al., 2003), electric load forecasting (Pai and Hong, 2005; Pai and Lin, 2005) and product demand forecasting (Guajardo et al., 2006). The SVR model has also been successfully applied to production value forecast of IC industry (Pai et al., 2009), it motivates our research work by using SVR for production value forecasting.

In the modeling of output value forecasting using SVR, as the existing approaches would either involve cost lots of calculating resources or sensitive to parameter selection (Lu et al., 2009). Moreover, this variability makes it difficult for any single artificial technique to capture the non-stationary property of the data (Mitani et al., 2003). To avoid the limitations of the existing approach and reduce the influence of non-stationary property, a hybrid methodology by combining empirical mode decomposition (EMD) and support vector regression is proposed in the study for production values forecasting.

Recently, a new signal analysis method, namely EMD proposed by Huang et al. (1998), is based on the local characteristic time scale of the signal and can decompose the complicated signal into a number of intrinsic mode functions (IMFs). By analyzing each resulting IMF component which involves the local characteristic of the signal, the characteristic information of the original signal can be extracted more accurately and effectively. In addition, the frequency components involved in each IMF not only relate to the sampling frequency but also change with the signal itself. Therefore, EMD is a self-adaptive signal processing method that can be applied to nonlinear and non-stationary time series analysis (Huang et al., 1998; Huang et al., 1999). Due to the simplicity of its algorithm, the EMD method has been successfully applied in several fields such as ocean waves (Hwang et al., 2003), biomedical engineering (Balocchi et al., 2004; Jiang and Yan, 2008; Liang et al., 2005; Su et al., 2008), mechanical fault diagnosis (Cheng et al., 2004; Vincent et al., 1999; Yu et al., 2005), signal processing (Blanco-Velasco et al., 2008; Guo et al., 2008; Li and Meng, 2006; Rai and Mohanty, 2007; Tao et al., 2005; Xie and Wang, 2006), wind engineering (Li and Wu, 2007), and earthquake engineering (Dong et al., 2008). However, most applications are primarily limited to the studies of nature science and engineering (Guo et al., 2008). Moreover, in existing literatures there are still few applications using EMD to forecast the output value.

In this paper, we introduce EMD and SVR to forecast the output value. The EMD is used to adaptively decompose the original time series data into a finite and small number of oscillatory modes based on the local characteristic time scale to improve the performance of SVR. A hybrid methodology by combining EMD and SVR is proposed in this study for output value forecasting. Firstly, this approach employs EMD to decompose the original output value data into several IMFs and a residue. Secondly, the tendencies of these IMFs and the residue are then modeled and forecasted using SVR. Finally, these results prediction will be integrated to get a final forecasting value. The proposed approach was compared with the existing SVR, thus demonstrating that the proposed model can result in an enhancement of prediction accuracy and reduction of the influence of non-stationary property.

The remainder of this paper is organized as follows: brief methodology on overviews of EMD and SVR; discussion on hybrid models; results of the experiment; the conclusion that describes the contribution of this paper is summarized and several future researches are also listed.

METHODOLOGY

This paper constructs a support vector regression (SVR) predicting model to mitigate the problem of production values forecast. It is aided by the utilization of empirical mode decomposition (EMD) (called EMD- SVR model). By using EMD, any complicated data can be decomposed into a finite and often small number of intrinsic mode functions (IMFs) (Huang et al., 1998). Then, a SVR, trained by IMFs and the corresponding demand data is used to predict these IMFs of the future production values. Finally, the production values are forecasted by summing the predicted IMFs. The detailed introduction and literature review of each method can be seemed in the following areas.

Empirical mode decomposition

The empirical mode decomposition (EMD) technique, proposed by Huang et al. (1998), is a form of adaptive time series decomposition technique using the Hilbert–Huang transform (HHT) for nonlinear and non-stationary signals. The basic principle of EMD is to decompose a time series into a sum of oscillatory functions, namely, intrinsic mode functions (IMFs). In the EMD, the IMFs must satisfy two conditions: (1) the number of extrema (sum of maxima and minima) and the number of zero crossing differs only by one, and (2) the local average is zero. The condition that the local average is zero implies that envelope mean of the upper envelope and lower envelope is equal to zero. The first condition is similar to the traditional narrow band requirements for a stationary Gaussian process. The second condition modifies classical global requirement to a local one; it is necessary so that the instantaneous frequency will not have the unwanted fluctuations induced by asymmetric wave forms. The detail algorithm for EMD is shown as follows (Flandrin et al., 2004; Huang, 2001; Huang et al., 1998; Wu and Hu, 2006):

Step 1: Identify the entire local extrema (including local maxima and minima), and then connect all the local maxima by a cubic spline line as the upper envelope.

Step 2: Repeat the procedure for the local minima to produce the lower envelope. The upper and lower envelopes should cover all the data between them.

Step 3: The mean of the upper and lower envelope value is designated as $m_1(t)$, and the difference between the time series $x(t)$ and $m_1(t)$ is the first component, $h_1(t)$, that is,

$$h_1(t) = x(t) - m_1(t). \tag{1}$$

Ideally, if $h_1(t)$ is an IMF, then $h_1(t)$ is the first component of $x(t)$.

Step 4: If $h_1(t)$ is not an IMF, $h_1(t)$ is treated as the original time series and repeats Step 1 to 3; then

$$h_{11}(t) = h_1(t) - m_{11}(t) \tag{2}$$

in which, $m_{11}(t)$ is the mean of the upper and lower envelope value of $h_1(t)$. After repeated sifting, i.e. up to k times, $h_{1k}(t)$ becomes an IMF, that is

$$h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t); \tag{3}$$

then, it is designated as

$$c_1(t) = h_{1k}(t), \tag{4}$$

Here, to guarantee that the IMF components retain enough physical sense of both amplitude and frequency modulations, a criterion for stopping the sifting process is used. This is accomplished by limiting the size of the standard deviation, denoted as *SD*, which is calculated from two consecutive sifting results as

$$SD = \sum_{t=0}^{N} \left[\frac{\left| h_{1(k-1)}(t) - h_{1k}(t) \right|^2}{h_{1(k-1)}^2(t)} \right]. \tag{5}$$

A typical value for *SD* can be set between 0.2 and 0.3.

Step 5: Separate $c_1(t)$ from $x(t)$, we could get:

$$r_1(t) = x(t) - c_1(t), \tag{6}$$

$r_1(t)$ is treated as the original data and the above processes are repeated, the second IMF component $c_2(t)$ of $x(t)$ can be achieved. Let us repeat the process as described above for n times, then n-IMFs of time series $x(t)$ can be got. Then,

$$\begin{aligned} r_2(t) &= r_1(t) - c_2(t) \\ r_3(t) &= r_2(t) - c_3(t) \\ &\vdots \\ r_n(t) &= r_{n-1}(t) - c_n(t). \end{aligned} \tag{7}$$

The decomposition process can be stopped when $r_n(t)$ becomes a monotonic function from which no more IMF can be extracted. By summing up Equations 6 and 7, we finally obtain

$$x(t) = \sum_{j=1}^{n} c_j(t) + r_n(t) \tag{8}$$

Thus, residue $r_n(t)$ is the mean trend of $x(t)$. The IMFs $c_1(t)$, $c_2(t)$, . . . , $c_n(t)$ include different frequency bands ranging from high to low. The frequency components contained in each frequency band are different and they change with the variation of time series $x(t)$, while $r_n(t)$ represents the central tendency of time series $x(t)$.

Support vector regression (SVR)

SVR first non-linearly transforms the original input space x into a higher dimensional feature space. That is, in order to learn non-linear relations with a linear machine, it is required to select a set of non-linear features and to express the data in the new representation. This is equivalent to applying a fixed non-linear mapping of the data to a feature space which the linear machine can be used in. This transformation can be achieved by using various types of non-linear mapping. Non-linear regression problems in an input space can become linear regression problems in a feature space. The regression function can be formulated as follows:

$$f(x) = w \cdot \phi(x) + b \tag{7}$$

where w is a weight vector, b is a constant, $\phi(x)$ denotes the feature of the inputs, and ($w \cdot \phi(x)$) describes the dot production in the feature space F. In SVR, the problem of nonlinear regression in the lower dimension input space(x) is transformed into a linear regression problem in a high dimension feature space F. That is, the original optimization problem involving a nonlinear regression is recast as a search for the flattest function in the feature space, not in the input space.

A number of cost functions such as the Laplacian, Huber's Gaussian, and ε-insensitive can be used in the SVR formulation. Among these, the robust ε-insensitive loss function (L_ε) which is the most commonly adopted is presented as follows:

$$L_\varepsilon(f(x), q) = \begin{cases} |f(x) - q| - \varepsilon & for \ |f(x) - q| \geq \varepsilon \\ 0 & otherwise \end{cases} \tag{8}$$

Where ε is a precision parameter representing the radius of the tube located around the regression function *f*(x). Under this loss function, errors below ε are not penalized; we can ignore the error and say the predicted *f (x)* has no loss.

The weight vector (W) and constant (*b*) in Equation 7 can be estimated by minimizing the following regularized risk function:

$$R(C) = C\frac{1}{n}\sum_{i=1}^{n} L_e(f(x_i), q_i) + \frac{1}{2}\|\omega\|^2 \qquad (9)$$

and $L_e(f(x), q)$ is the ε-insensitive loss function in Equation 8. The *f (x_i)* and q_i values are the actual and forecasting values, respectively. The second term, $\|\omega\|^2/2$ is the regularization term which controls the trade-off between the complexity and approximation accuracy of the regression model to ensure that the model possesses an improved generalized performance; and *C* is the regularization constant used to specify the trade-off between the empirical risk and regularization term. Both *C* and ε are user-determined parameters. Two positive slack variables, ξ_i and ξ_i^*, which represent the distance between the actual values and the corresponding boundary values of ε-tube, are introduced. Thus, Equation 9 is transformed into the following constrained form.

$$\text{Minimize}: \; R_{reg}(f) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \qquad (10)$$

with the constraints:

$$\begin{cases} q_i - w\cdot\phi(x_i) - b \le \varepsilon + \xi_i^* \\ w\cdot\phi(x_i) + b - q_i \le \varepsilon + \xi_i \\ \xi_i, \xi_i^* \;\ge\; 0, \qquad for\; i = 1,\ldots,n \end{cases}$$

By applying Lagrangian multipliers and Karush-Kuhn-Tucker conditions to equation (4), it thus yields the dual Lagrangian form as Equation 11.

$$\begin{aligned} \text{Minimize}: L_d(\alpha, \alpha^*) = -\varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) + \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)q_i \\ - \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, x_j) \end{aligned} \qquad (11)$$

subject to
$$\begin{cases} \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) = 0 \\ 0 \le \alpha_i \le C \quad, \quad i = 1,\ldots,n \\ 0 \le \alpha_i^* \le C \quad, \quad i = 1,\ldots,n \end{cases}$$

In Equation 11, α_i and α_i^*, are called Lagrangian multipliers. They satisfy the equality $\alpha_i \cdot \alpha_i^* = 0$. After the Lagrangian multipliers α_i and α_i^* have been calculated, the minimizing function can be written in the following form.

$$f(x, w) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x, x_i) + b \qquad (12)$$

Herein, $K(x, x_i)$ is called the kernel function whose value equals the inner product of two vectors, x_i and x_j, in the feature space, $\phi(x_i)$ and $\phi(x_j)$, meaning that $K(x_i, x_j) = \phi(x_i)\cdot\phi(x_j)$.

There are several types of kernel function. However, any function that satisfies Mercer's condition by Vapnik (1995) can act as the kernel function. The most widely used kernel function is the Gaussian radial basis function (RBF) defined as $K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2})$, where σ denotes the width of the RBF. Moreover, the RBF kernel is not only easier to implement, but also capable to non-linearly map the training data into an infinite dimensional space, thus, it is suitable to deal with non-linear relationship problems. Thus, the RBF is used in this work as the kernel function.

The combined EMD with SVR model

For the proposed hybrid forecasting method, EMD is the first applied to decompose the original production value data into a finite set of IMFs (Here the residual r_n (*t*) also be considered as an IMF). Second, each IMF is modeled and forecasted using the SVR model, as is the residue. The step involved in SVR modeling are (1) selecting a suitable kernel function and kernel parameter (kernel width *σ*), (2) specifying the "*ε*" insensitive parameter, and (3) specifying the capacity parameter cost, "*C*".

As mentioned earlier, the RBF kernel function is adapted in this

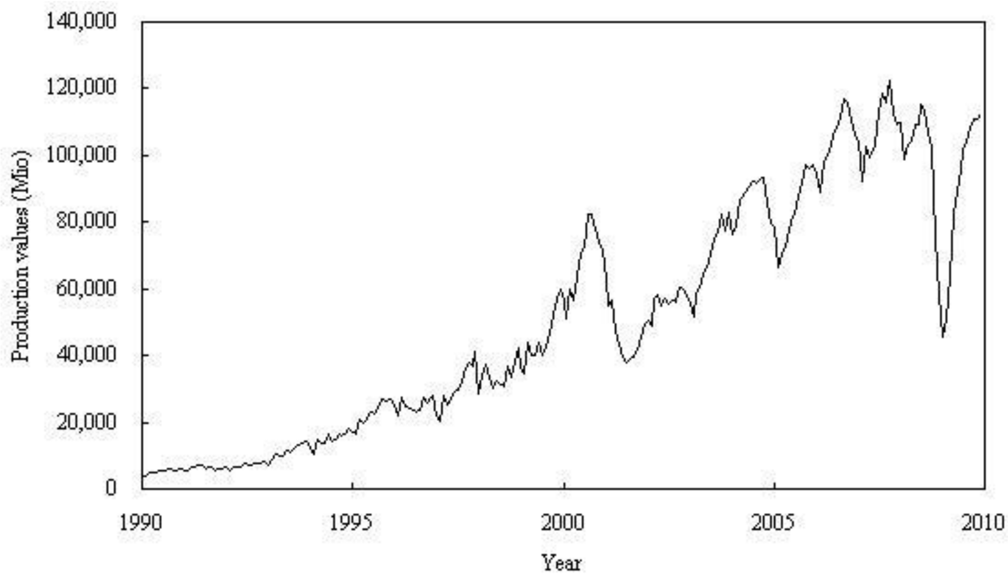

Figure 1. Variations of the monthly production values for Taiwan's IC industry.

Table 1. Performance indices and their calculations.

Metrics	Calculation
MAPE*	$\frac{1}{n} \times \sum_{i=1}^{n} \left\lvert \frac{F_i - A_i}{F} \right\rvert \times 100\%$
RMSE	$\sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (F_i - A_i)^2}$
MAD	$\frac{1}{n} \times \sum_{i=1}^{n} \lvert F_i - A_i \rvert$

study. It is well known that the selection of these parameters, C, *ε*, and *σ* of a SVR model is important to the accuracy of forecasting (Cherkassky and Ma, 2004; Lin et al., 2003). There are no general rules for the choice of these parameters. The grid search proposed by Lin et al. (2003) is a common and straightforward method using exponentially growing sequences of C, *ε*, and *σ* to identify good parameters (for example, $C=2^{-15}$, 2^{-13}, 2^{-11},…, 2^{15}). The parameter set of C, *ε*, and *σ* which generate the minimum forecasting mean square error (MSE) is considered as the best parameter set. In this study, the grid search is used in each IMF to determine the best parameter set for training an optimal SVR forecasting model. Finally, the forecasting results can be obtained by the sum of the forecasts.

EXPERIMENTAL RESULTS AND ANALYSIS

Data sets

To evaluate the performance of the proposed methodology, monthly production values of IC industry in Taiwan from January 1990 to December 2009 are selected. The collected data is divided into two sets, training data and testing data. This study uses 70% of the data from the data set as the training set. The remaining 30% are used as the testing set. Therefore, the training samples are selected data from January 1990 to December 2005, while the data from January 2006 to December 2009 are used to check the prediction performance of the model. There are totally 240 data points in the dataset and the variations of the historical monthly data from the IC industry are shown in Figure 1. The plot exhibits a permanent deterministic pattern of long-term upward trend with short-term fluctuations that are independent from one time period to the next. From Figure 1, it can be seen that the production values of the Taiwanese IC industry appear to be non-stationary in that the mean is increasing over time.

Evaluating indices

In this study, the one-step-ahead prediction is performed in the experiments. Although multi-step forecasting may capture some system dynamics, the performance will be quite poor due to the accumulation of errors. In practice, short-term forecasting results are more useful as they provide timely information for the correction of forecasting value.

In this study, three popular indices to measure the forecast are mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute difference (MAD). MAPE, RMSE and MAD were used to measure the correctness of a prediction in terms of levels and the deviation between the actual and predicted values. The smaller the values, the closer the predicted values are to the actual values. Table 1 shows these performance metrics and their calculations.

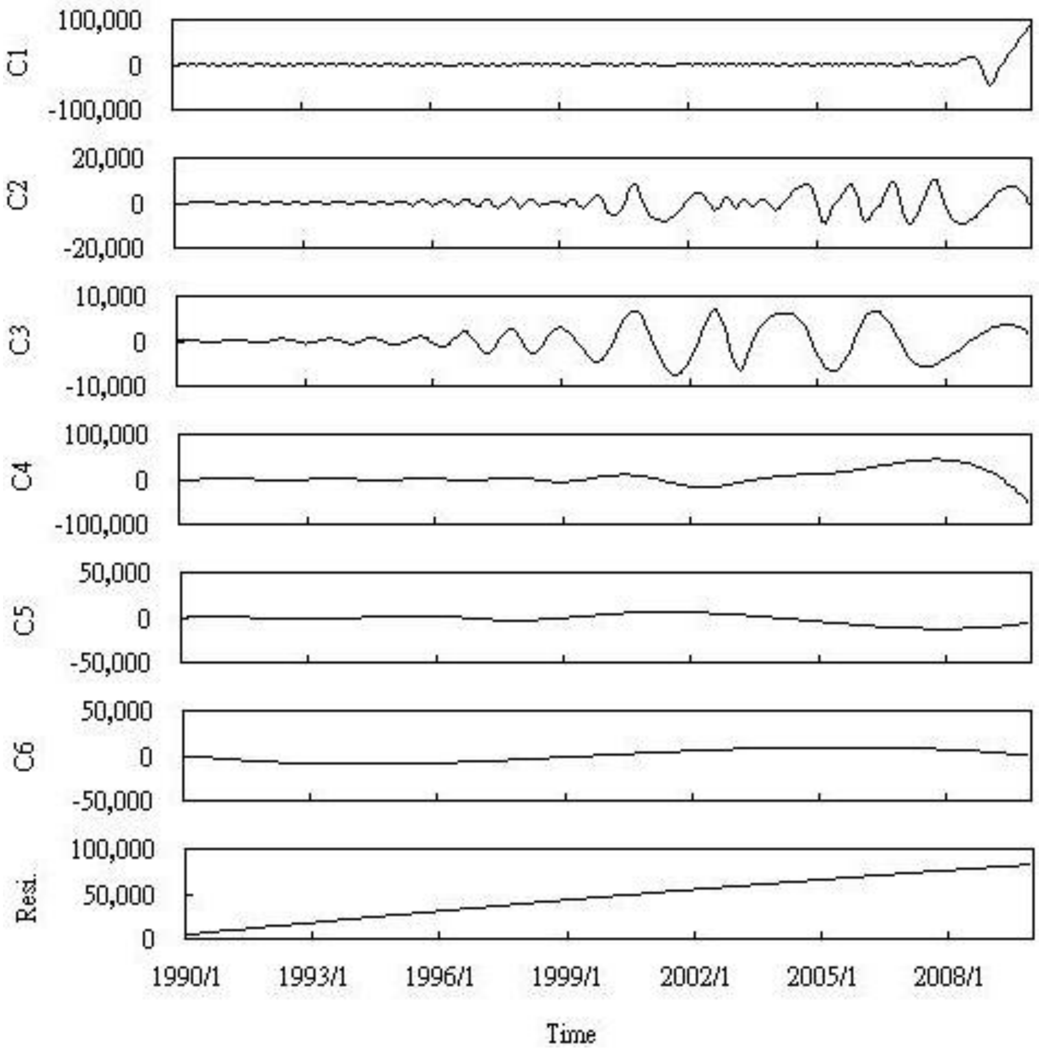

Figure 2. EMD for production values.

Forecasting results

According to methodology proposed in this paper, observed data should be mapped into several time sequence domains by EMD method with the sifting threshold *SD* = 0.3. Figure 2, shows the decomposition results for the production values. Clearly, the data is decomposed into six IMFs and one residue. For the decomposition, it is observed that the dominant mode of the observed data is not any IMF but the residue. As Huang et al. (1998) mentioned the residue is often treated as the deterministic long term behavior. So, the IMFs of IC industry can be decomposed into the sequences including c_1, c_2, c_3, c_4, c_5, c_6, and r_7, just as shown in Figure 2, which can be applied to train SVR, respectively.

IMF c_1 is given as training samples firstly, in the selection of parameters for modeling SVR, through grid search proposed by Lin et al. (2003), it can be obtained that the parameter set ($C=2^1$, $\varepsilon=2^{-11}$) gives the best forecasting result (minimum testing MSE). The similar procedures can be with other SVR models, which are established in the same way as for the IMF of c_2, c_3, c_4, c_5, c_6, and r_7. Then, more accurate results of forecast can be derived through the algebraic sum of the forecasted IMF c_1, c_2, c_3, c_4, c_5, c_6, and r_7. The actual production values in the IC industry and predicted values from the single SVR and EMD-SVR models are illustrated in Figure 3. It can be observed from Figure 3 that the predicted values obtained from the proposed EMD-SVR model are closer to the actual values than those obtained from the single SVR model.

Table 2 compares the results obtained with the EMD-SVR and single SVR models for production values in the IC industry. Table 2 depicts that the MAPE, RMSE and MAD of the proposed EMD-SVR model are, respectively, 0.44%, 640.19 and 373.10. It can be observed that these values are smaller than SVR model.

This indicates that there is a smaller deviation between

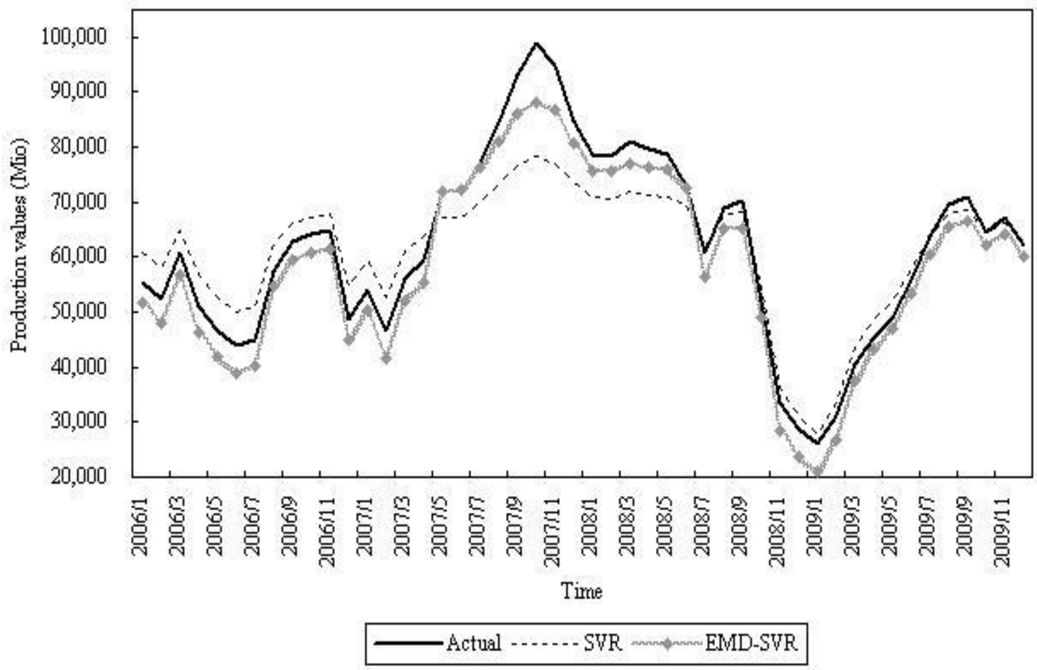

Figure 3. Forecasted production values from different models for IC industry.

Table 2. Comparison of different models for production values.

Metrics \ Models	SVR	EMD-SVR
MAPE	6.77%	0.44%
RMSE	8372.88	640.19
MAD	6006.78	373.10

the actual and predicted values using the proposed EMD-SVR model. Thus, the proposed EMD-SVR model provides a better forecasting result than the single SVR model based on MAPE, RMSE and MAD.

Conclusions

Due to the specific characteristics in the IC industry, production values are inherently nonlinear and non-stationary, this variability makes it difficult for a single model including SVR to capture such a dynamic input-output relationship inherent in the data. This paper has presented a production values forecasting model by integrating EMD and SVR. The main contribution of the paper is to propose a hybrid method as well as a simple approach for a stable prediction of non-stationary data. The proposed EMD-SVR method pre-processes the production values of IC industry in Taiwan and decomposes them into more stationary and regular components (IMF or residue) using the EMD technique. Furthermore, the corresponding SVR model for each divided component is easier to build. After each IMF components and residue are modeled and forecasted using the SVR model, the forecasting values are then summarized as the production values forecasting results. The experiments have evaluated the production values of IC industry in Taiwan.

In this study, the EMD-SVR model is proposed to produce lower prediction error in the datasets. It outperformed the single SVR model. According to the experiments, it can be concluded that EMD, which can fully capture the local fluctuations of data, can be used as a pre-processor to decompose the complicated raw data into a finite set of IMFs and a residue. By this pre-processing, we can advance the simplification of SVR modeling based on MAPE, RMSE and MAD. Therefore, the proposed method is very suitable for prediction with nonlinear and non-stationary data, and is an efficient method for production values forecasting.

The favorable results obtained in this work reveal that the proposed model is a valid alternative for use in high technological production values forecasting. Future studies may aim at combining EMD and other advancement of methods, like extreme learning machine (ELM) and grey theory, in evaluating the ability of the proposed forecasting scheme. Second, in this paper EMD has been introduced and issues with its application for production values forecasting. However, two key decisions in the EMD application process, the rule for deciding when to stop sifting for an IMF and the choice of cubic spline end condition rule have to be reviewed and discussed in detail for the future studies. Finally, results should be compared with other methods such as wavelet transform or neural network.

ACKNOWLEDGEMENT

The authors acknowledge the financial support from the National Science Council of Taiwan, R.O.C., under project number 100WFAB500033.

REFERENCES

Blanco-Velasco M, Weng B, Barner KE (2008). ECG signal denoising and baseline wander correction based on the empirical mode decomposition. Comput. Biol. Med. 38(1):1-13.

Balocchi R, Menicucci D, Santarcangelo E, Sebastiani L, Gemignani A, Ghelarducci B, Varanini M (2004). Deriving the respiratory sinus arrhythmia from the heartbeat time series using empirical mode decomposition. Chaos Soliton Fract. 20(1):171-177.

Cao LJ (2003). Support vector machines experts for time series forecasting. Neurocomputing 51:321-339.

Cao LJ, Tay FEH (2001). Financial forecasting using support vector machines. Neural Comput. Appl. 10:184-192.

Cheng JS, Yu DJ, Yang Y (2004). Energy operator demodulating approach based on EMD and its application in mechanical fault diagnosis. Chin. J. Mech. Eng. 40(8):115-118.

Cherkassky V, Ma Y (2004). Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw. 17:113-126.

Dong Y, Li Y, Xiao M, Lai M (2008). Analysis of earthquake ground motions using an improved Hilbert–Huang transform. Soil Dyn. Earthq. Eng. 28(1):7-19.

Flandrin P, Rilling G, Goncalves P (2004). Empirical mode decomposition as a filter bank. IEEE Signal Process. Lett. 11(2):112-114.

Guajardo J, Weber R, Miranda J (2006). A forecasting methodology using support vector regression and dynamic feature selection. J. Inform. Knowl. Manage. 5:329-335.

Guo K, Zhang X, Li H, Meng G (2008). Application of EMD method to friction signal processing. Mech. Syst. Signal Proc. 22(1):248-259.

Hsu LC (2003). Applying the grey prediction model to the global integrated circuit industry. Technol. Forecast. Soc. Chan. 70(6):563–574.

Hsu LC, Wang CH (2007). Forecasting the output of integrated circuit industry using a grey model improved by the Bayesian analysis. Techn. Forecast. Soc. Chan. 74(6):843-853.

Huang NE (2001). Review of empirical mode decomposition. Proc. SPIE, pp. 71-80.

Huang NE, Shen Z, Long SR (1999). A new view of nonlinear water waves: the Hilbert spectrum. Ann. Rev. Fluid Mech. 31(1):417-457.

Huang NE, Shen Z, Long SR, Wu MC, Shin HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London, Series A: Mathematical, Phys. Eng. Sci. 454:903-995.

Hussain AJ, Knowles A, Lisboa PJG, El-Deredy W (2008). Financial time series prediction using polynomial pipelined neural networks. Expert. Syst. Appl. 35(3):1186-1199.

Hwang PA, Huang NE, Wang DW (2003). A note on analyzing nonlinear and non-stationary ocean wave data. Appl. Ocean Res. 25(4):187-193.

Ince H, Trafalis TB (2006). A hybrid model for exchange rate prediction. Decis. Support Syst. 42:1054-1062.

Jiang R, Yan H (2008). Studies of spectral properties of short genes using the wavelet subspace Hilbert–Huang transform (WSHHT). Physica A, 387(16-17):4223-4247.

Kim KJ (2003). Financial time series forecasting using support vector machines. Neurocomputing 55:307-319.

Li HG, Meng G (2006). Detection of harmonic signals from chaotic interference by empirical mode decomposition. Chaos Soliton Fract. 30(4):930-935.

Li QS, Wu JR (2007). Time–frequency analysis of typhoon effects on a 79-storey tall building. J. Wind Eng. Ind. Aerod. 95(12):1648-1666.

Liang H, Lin QH, Chen JDZ (2005). Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease. IEEE Trans. Biomed Eng. 52(10):1692-1701.

Lin CJ, Hsu CW, Chang CC (2003). A practical guide to support vector classification. Department of Computer Science and Information Engineering, National Taiwan University.

Lu CJ, Lee TS, Chiu CC (2009). Financial time series forecasting using independent component analysis and support vector regression. Decis. Support Syst. 47:115-125.

Mitani Y, Tsutsumoto K, Kagawa N (2003). Time Series Prediction of Acoustic Signals Using Neural Network Model and Wavelet Shrinkage. Proceedings of the 10th International Congress on Sound and Vibration, Stockholm, Sweden: IIAV, pp. 4189-4196.

Mohandes MA, Halawani TO, Rehmam S, Hussain AA (2004). Support vector machines for wind speed prediction. Renew. Energy 29: 939-947.

Pai PF, Hong WC (2005). Support vector machines with simulated annealing algorithms in electricity load forecasting. Energy. Convers. Manage. 46:2626-2669.

Pai PF, Lin CS (2005). Using support vector machines in forecasting production values of machinery industry in Taiwan. Int. J. Adv. Manuf. Tech. 27:205-210.

Pai PF, Yang SL, Chang PT (2009). Forecasting output of integrated circuit industry by support vector regression models with marriage honey-bees optimization algorithms. Expert Syst. Appl. 36:10746-10751.

Rai VK, Mohanty AR (2007). Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform. Mech. Syst. Signal Proc. 21(6):2607-2615.

Su ZY, Wang CC, Wu T, Wang YT, Tang FC (2008). Instantaneous frequency–time analysis of physiology signals: the application of pregnant women's radial artery pulse signals. Physica A, 387(2–3): 485-494.

Tao Q, Chen Q, Li L (2005). Analytic unit quadrature signals with nonlinear phase. Physica D, 203(1-2):80-87.

Tay FEH, Cao LJ (2001). Application of support vector machines in financial time series forecasting. Omega 29:309-317.

Tay FEH, Cao LJ (2003). Support vector machine with adaptive parameters in financial time series forecasting. IEEE Trans. Neural Netw. 14:1506-1518.

Thissen U, Van Brakel R, De Weijer AP, Melssen WJ, Buydens LMC (2003). Using support vector machines for time series prediction. Chemometr. Intell. Lab. 69:35-49.

Vapnik VN (1995). The nature of statistical learning theory, Springer-Verlag, New York.

Vapnik VN (1999). An overview of statistical learning theory. IEEE Trans. Neural Netw. 10:988-999.

Vincent HT, Hu SLJ, Hou Z (1999). Damage detection using empirical mode decomposition method and a comparison with wavelet analysis. Proceedings of the Second International Workshop on Structural Health Monitoring, Stanford, pp. 891-900.

Wang CH, Hsu LC (2008). Using genetic algorithms grey theory to forecast high technology industrial output. Appl. Math. Comput. 195(1):256-263.

Wu MC, Hu CK (2006). Empirical mode decomposition and synchrogram approach to cardiorespiratory synchronization. Phys. Rev. E 73:051917.

Xie H, Wang Z (2006). Mean frequency derived via Hilbert–Huang transform with application to fatigue EMG signal analysis. Comput. Meth. Prog. Biol. 82(2):114-120.

Yaser SAM, Atiya AF (1996). Introduction to financial forecasting. Appl. Intell. 6:205-213.

Yu DJ, Cheng JS, Yang Y (2005). Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings. Mech. Syst. Signal Proc. 19(2):259-270.