EURASIA Journal of Mathematics, Science and Technology Education
ISSN: 1305-8223 (online) 1305-8215 (print)
OPEN ACCESS
2017 13(12):7893-7904 DOI: 10.12973/ejmste/77926

Assessing Potentiality of Support Vector Machine Method in Crude Oil Price Forecasting

Lean Yu [1*], Xun Zhang [2], Shouyang Wang [2]

[1] School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, CHINA
[2] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, CHINA

Received 16 May 2017 ▪ Revised 16 August 2017 ▪ Accepted 27 September 2017

ABSTRACT
Crude oil price forecasting is one of the most important topics in the field of energy research. Accordingly, numerous methods such as statistical, econometrical and intelligent approaches are applied for crude oil price forecasting. In this paper, a typical competitive learning algorithm, support vector machine (SVM), is empirically investigated to verify the feasibility and potentiality of SVM in crude oil price forecasting. For this purpose, five different prediction models, feed-forward neural networks (FNN), auto-regressive integrated moving average (ARIMA) model, fractional integrated ARIMA (ARFIMA) model, Markov-switching ARFIMA (MS-ARFIMA) model, and random walk (RW) model are used in the study. Experimental results obtained show that the SVM model outperforms the other five methods, implying that it is a fairly good candidate for crude oil price forecasting in terms of either one-step prediction or multi-step prediction.

Keywords: support vector machines, artificial neural networks, ARIMA model, ARFIMA model, Markov-switching ARFIMA model

INTRODUCTION

The sharp increase in crude oil price between 2004 and 2008, has resulted in problems related to high volatility in oil prices receiving much attention. Usually, high crude oil price influences not only macro-economic development but also quality of human lives. Prediction of future crude oil price can help neutralize, to some extent, impact of fluctuations on macro- and microeconomics. However, crude oil price forecasting is not an easy task due to the fact that crude oil price is formulated by complex factors that have several interactive effects between themselves. As Zhang et al. (2008) revealed, three main types of factors (short-term, medium-term, and long-term) affect crude oil price volatility, which has the characteristics of complex nonlinearity, dynamic volatility and high irregularity (Watkins and Plourde, 1994). Unfortunately, the fundamental mechanism governing the complex dynamics in crude oil markets is not well understood by human beings (Yu et al., 2008). In a sense, crude oil price forecasting is still a rather challenging task for both academia and practitioners.

In the past decades, some attempts have been made for exploring the crude oil price dynamics. Some statistical-based models have been widely used for crude oil prices forecasting. Typical models include the probabilistic model (Abramson and Finizza, 1995), econometric structural models (Huntington, 1994; Ye et al., 2002, 2005, 2006), co-integration analysis (Gulen, 1998), vector auto-regression models (VAR) (Mirmirani and Li, 2004), error correction models (ECM) (Lanza et al., 2005), auto-regressive integrated moving average (ARIMA) (Yu et al., 2008) and semi-parametric approach based on GARCH properties (Morana, 2001). Usually, these models can provide good prediction results when the crude oil price series under study is linear or near linear. However, in real-world crude oil price series, there is a great deal of nonlinearity and irregularity. Numerous experiments have demonstrated that the prediction performance might be very poor if one continued using these traditional statistical and econometric models (Weigend and Gershenfeld, 1994). The main reason leading to this phenomenon is that most statistical-based models are built on linear assumptions and they cannot capture the nonlinear patterns hidden in the crude oil price series.

yulean@amss.ac.cn **(*Correspondence)** zhangxun@amss.ac.cn sywang@amss.ac.cn

Contribution of this paper to the literature

- The empirical results find that among different forecasting models used for the two main crude oil price series (WTI crude oil price and Brent crude oil price), in terms of different criteria, the SVM-based forecasting model performs the best. In all testing cases, *RMSE* is the lowest and D_{stat} is the highest, indicating that the SVM-based forecasting approach can provide a promising alternative and offers advantages in crude oil price time.
- Application of the SVM method in crude oil price forecasting is a good and interesting attempt and it may be worth testing its utility in other fields also. We will look into these issues in the future.

For these reasons, some nonlinear models (e.g., Panas and Ninni, 2000), especially the emerging artificial intelligence (AI) models, are used for crude oil price forecasting. Among AI models, artificial neural networks (ANNs) are often regarded as a class of reliable and cost-effective methods for crude oil price prediction. The neural network model, particularly the multi-layer feed-forward neural network (FNN), can be trained to approximate any smooth and measurable nonlinear function without prior assumptions on the original data (Yu et al., 2007b; Lee & Lee, 2016); it has produced many promising results in this field (Kaboudan, 2001; Mirmirani and Li, 2004; Wang et al., 2004, 2005; Shambora and Rossiter, 2007; Yu et al., 2007a, 2008, Chou, 2016). These studies have shown that ANN models are very effective in simulating and describing the dynamics of non-stationary time series due to its unique non-parametric, noise-tolerant and highly adaptive characteristics.

However, the inherent drawbacks of ANN models, e.g., local minima, over-fitting, poor generalization performance and the difficulty of determining appropriate network architectures, hinder practical applications of ANN models. Support vector machines (SVMs), first proposed by Vapnik (1995), provide a class of competitive learning algorithms to improve generalization performance of neural networks and achieve global optimum solutions simultaneously. SVMs are a very specific type of learning algorithms characterized by the capacity control of the decision function, use of kernel functions, and sparsity of the solution (Vapnik, 1995, 1999; Cristianini and Taylor, 2000). Established on the unique theory of the structural risk minimization (SRM) principle to estimate a function by minimizing an upper bound of the generalization error, SVM is resistant to the over-fitting problem and can simulate nonlinear relations in an efficient and stable way. This property leads to a better generalization than conventional methods. Furthermore, SVMs are trained as a convex optimization problem, resulting in a global solution that in many cases yields unique solutions. Initially, SVMs were developed for classification tasks. With introduction of the ε-insensitive loss function, SVMs have been extended to solve nonlinear regression and time series prediction problems, and have exhibited excellent performance (Muller et al., 1997; Vapnik et al., 1996; Cao and Tay, 2001; Tay and Cao, 2001a, 2001b, 2002).

The main purpose of this study is not to propose a new SVM-based machine learning algorithm but to present an empirical study of application of the SVM model to crude oil price time series prediction, so as to investigate the potentiality of SVM in crude oil price forecasting. For this purpose, five different models - feed-forward neural network (FNN) model, auto-regressive integrated moving average (ARIMA) model, fractional integrated ARIMA (ARFIMA) model, Markov-switching ARFIMA (MS-ARFIMA) model, and random walk (RW) model - are examined in the study. The rest of the article is organized as follows. Section 2 describes the SVM-based method for crude oil price prediction. In order to evaluate the potentiality of SVM in crude oil price forecasting, an empirical study and its computational results are reported in Section 3. Section 4 concludes the article.

SVM-BASED CRUDE OIL PRICE FORECASTING APPROACH

In this section, the overall process of formulating the SVM-based crude oil price forecasting paradigm is presented. First, the theory of SVM for regression tasks is introduced, and then the SVM-based crude oil price time series prediction method is proposed.

Theory of SVM Regression

Support vector machines (SVMs) have originally been proposed for classification purposes but their principles can be extended to regression and time series prediction problems as well. In this paper, we only focus on support vector regression (SVR) for time series prediction. Excellent general descriptions of SVMs, including support vector classification (SVC) and support vector regression (SVR), can be found in Vapnik (1995, 1998), Burges (1998), and Cristianini and Taylor (2000).

Basically, SVR is a linear learning machine. That is, a linear function is always used to solve the regression problems. When dealing with nonlinear regression, SVR first maps the original data x into a high-dimensional feature space via a nonlinear mapping function φ and then makes linear regression in this high-dimension feature space. Usually, the SVM regression function can be formulated as follows

$$y = f(x) = w \cdot \varphi(x) + b \tag{1}$$

where $\varphi(x)$ is called the nonlinear feature space mapped from input space x and y is the estimated value in terms of input data x. Coefficients w and b are estimated by minimizing

$$R_{reg}(C) = C \cdot \frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(d_i, y_i) + \frac{1}{2}\|w\|^2 \tag{2}$$

$$L_\varepsilon(d_i, y_i) = \begin{cases} |d - y| - \varepsilon, |d - y| \geq \varepsilon, \\ 0, otherwise. \end{cases} \tag{3}$$

where R_{reg} is the regularized risk function, d_i is the actual value in the ith period, and C and ε are user-specified parameters. In Eq. (2), the first term $(C/N)\sum_{i=1}^{N} L_\varepsilon(d_i, y_i)$ is the empirical error (risk), measured by the ε-insensitive loss function given by Eq. (3). This loss function provides the advantage of enabling one to use sparse data points to represent the regression function defined by Eq. (1). The second term, $(1/2)\|w\|^2$, is the regularization term, which measures the flatness of the function. C is referred to as the regularized constant and it determines the trade-off between empirical risk and the regularization term. Increasing the value of C will result in relatively higher importance of the empirical risk and vice versa. ε is called the tube size, and it is equivalent to the approximate accuracy placed on the training data points. Introducing the positive slack variables ξ and ξ^*, which represent the distance from the actual values to the corresponding boundary values of ε-tube. For this setting, Eq. (2) can be transformed into the following optimization problem

$$\begin{cases} \text{Minimize } \frac{1}{2}ww^T + C^*\left(\sum_{i=1}^{N} \xi + \xi *\right) \\ \text{Subject to: } w \cdot \varphi(x_i) + b_i - d_i \leq \varepsilon + \xi_i^* \\ d_i - w \cdot \varphi(x_i) + b_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, \text{for all } i = 1,2,\cdots,N. \end{cases} \tag{4}$$

By introducing Lagrangian multipliers and maximizing the dual function of (4), the (4) can be changed into the following dual form

$$\begin{cases} \text{Minimize } R(\lambda_i - \lambda_i^*) = \sum_{i=1}^{N} d_i(\lambda_i - \lambda_i^*) - \varepsilon\sum_{i=1}^{N}(\lambda_i - \lambda_i^*) \\ -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\lambda_i - \lambda_i^*) \cdot (\lambda_i - \lambda_i^*) \cdot K(x_i, x_j) \\ \text{Subject to: } \sum_{i=1}^{N}(\lambda_i - \lambda_i^*) = 0, \\ 0 \leq \lambda_i \leq C, \\ 0 \leq \lambda_i^* \leq C, \\ \text{for all } i = 1,2,\cdots,N. \end{cases} \tag{5}$$

In Eq. (5), λ_i and λ_i^* are called Lagrangian multipliers, which satisfy the equality $\lambda_i \cdot \lambda_i^* = 0$. Finally, the solution of the original problem can be represented as below:

$$f(x, \lambda_i, \lambda_i^*) = \sum_{i=1}^{M}(\lambda_i - \lambda_i^*)\, K(x, x_i) + b \tag{6}$$

In Eq. (6), $K(\bullet)$ is the so-called kernel function which simplifies the use of a mapping. Representing the mapping by simply using a kernel is called the kernel trick, and the problem is reduced to finding kernels that identify families of regression formulae. It can be shown that any symmetric kernel function $K(\bullet)$ satisfying Mercer's condition corresponds to a dot product in some feature spaces (Vapnik, 1995). The most used kernel functions are the Gaussian RBF with a width of σ: $K(x, x_i) = \exp(\|x - x_i\|/2\sigma^2)$, and the polynomial kernel, with an order of d and constants a_1 and a_2, represented as $K(x, x_i) = (a_1 x x_i + a_2)^d$.

SVM-based Crude Oil Price Time Series Forecasting

Since time series prediction can be seen as an auto-regressive process in time, a regression method can be used for this task. When time series prediction is conducted by SVMs, input vector $\{\boldsymbol{x}\}$ to the SVM is a finite set of consecutive measurements of the series $\boldsymbol{x} = (x(t), x(t-1), \ldots, x(t-s))$, with time-delay s, which is a sliding window for the input vector. The output of the regression is $x(t+h)$ where h is the prediction horizon and it is a user-specified parameter. The procedure of developing a SVM-based time series prediction is illustrated in **Figure 1**.

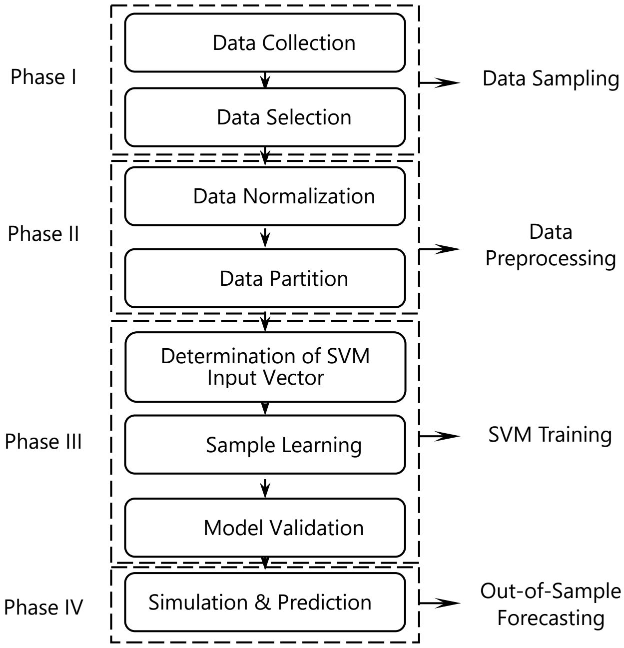

Figure 1. A procedure of SVM-based time series forecasting system

As can be seen from **Figure 1**, the procedure of SVM-based time series prediction model can be divided into four phases, briefly described as follows.

Phase I: Data Sampling. In situations where there are vast volumes of data to sift through, a process called data sampling can help minimize data processing and significantly reduce computational costs. Data sampling is a process whereby a statistically representative portion of the information is examined to determine if it contains responsive data. Using data sampling can help narrow the research focus, for example, by determining whether there are time periods in which relevant events do not exist; this makes it unnecessary to process or review that particular part of the data set. To develop a SVM-based model for forecasting, different data should be collected, and data collected from various sources must be selected in terms of some specific criteria.

For crude oil price, there are a variety of data used for this research. West Texas Intermediate (WTI) and Brent crude oil prices are two main crude oil price benchmarks. From the viewpoint of data type, spot prices and futures prices are available. From the point of data frequency, daily, weekly, monthly, quarterly, and yearly data can be used. The main purpose of data sampling is to select a representative data for further processing and analysis.

Phase II: Data Preprocessing. After data sampling, the next task is data preprocessing. It includes two steps: data normalization and data division. In any model development process, familiarity with the available data is of the utmost importance. SVM models are no exception. Data normalization can have a significant effect on SVM model's performance. After that, normalized data should be divided into two subsets: in-sample data and out-of-sample data, to be used for model estimation and model evaluation and verification respectively.

Phase III: SVM Training. After the data is preprocessed, SVM training can be performed using the processed data. In this phase, there are three main tasks: determination of SVM input vector, sample learning, and model validation. Usually, the SVM input vector is determined by time-delay s via the trial and error method. In sample learning, regularization constant C, suitable kernel functions $K(\bullet)$, and associated kernel parameters in kernel functions should be determined. Often they are determined by trial and error because there are no universal criteria for deciding the parameters. As an alternative, some search-based methods, such as grid search and direct search methods, can also be used to determine the SVM parameters. After training, model validation must be performed so as to guarantee the generalizability of SVM. After validation, a SVM predictor with optimal parameters can be obtained.

Phase IV: Out-of-Sample Forecasting. Using the optimal SVM predictor, the trained SVM can be used for out-of-sample time series prediction.

It is worth noting that the proposed SVM-based crude oil price forecasting model is constructed on a sliding window or rolling windows data basis. The estimation window is with a fixed size and it recursively changes as forecasting moves forward in time (West, 1996). In a sense, the proposed model is actually a SVM-based crude oil price rolling forecasting model.

To evaluate the forecasting ability of the SVM predictor, this study investigates its performance by comparing it with FNN, ARIMA, ARFIMA, MS-ARFIMA, and RW models. ARIMA model proposed by Box and Jenkins (1976) is a mixture of auto-regressive (AR) model and moving average (MA) models to describe time series. One of the distinct disadvantages of ARIMA is that it cannot describe the long-range dependence of complex time series (e.g. financial time series) because it implies an extreme form of persistence. But for a pure unit root series, the impact of shocks does not vanish even in the infinite horizon. For this a fractional integrated ARIMA (ARFIMA) model (Granger and Joyeux, 1980) is proposed. An ARFIMA model is a natural extension of ARIMA that tolerates fractional integration. However, both the ARIMA and ARFIMA models cannot capture the nonlinear patterns if nonlinearity exists in time series. In order to remedy the shortcoming, the Markov switching ARFIMA (MS-ARFIMA) model (Hamilton, 1994), which is a combination of Markov process and ARFIMA model, is used to assess the impact of both long memory and non-linearity on forecasting. As a popular intelligent model, a three-layer feed-forward neural network (FNN) incorporating the Levenberg-Marquardt algorithm is adopted for comparison purpose. The major merit of FNN models is their flexible nonlinear modeling capability. They can capture the nonlinear characteristics of time series well. However, FNN does not lead to one global or unique solution due to differences in initial weights. Interested readers can refer to Box and Jenkins (1976), Granger and Joyeux (1980), Hamilton (1994) and White (1990) for more details about these models.

EMPIRICAL STUDY

In this section, we first describe the data, and then define some evaluation criteria for prediction purposes. Finally, empirical results and explanations are presented.

Research Data and Evaluation Criteria

As Section 2.2 reveals, there are many crude oil price data series. For the purpose of investigation and analysis, the two benchmark crude oil price series, West Texas Intermediate (WTI) crude oil spot price and Brent crude oil spot price are chosen as experimental targets. Both price series are used widely as the basis of many crude oil price formulae (Yu et al., 2008).

In this study, we take monthly data from January 1990 to July 2008, with a total of 223 observations, as shown in **Figure 2**. The data are freely available from the Energy Information Administration (EIA) website of Department of Energy (DOE) of the United States (http://www.eia.doe.gov/). For ensuring rolling forecasting modeling, the estimation window is fixed at 180 observations and the forecasting horizon is set at one step. This means that the first training sample for model estimation is from January 1990 to December 2004, and then it repeatedly moves forward. The second training sample is from February 1990 to January 2005, and the last one is from July 1993 to June 2008. There are, in total, 43 observations in the testing set (from January 2005 to July 2008) used to evaluate the performance of prediction. Only one-step-ahead prediction is performed in the experiments. Actually, multi-step-ahead prediction, e.g., step size or prediction horizon is larger than one, can also be performed, but the prediction performance in such cases is unsatisfactory. For this reason, this study focuses only on one-month-ahead forecasting.

To measure the forecasting performance, two main criteria are used for evaluation of level prediction and directional forecasting. First, we select the root mean squared error (*RMSE*) as the criterion of evaluation of the level of accuracy. Typically, the *RMSE* can be defined by

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{x}(t) - x(t))^2} \tag{7}$$

where $x(t)$ is the actual value, $\hat{x}(t)$ is the predicted value, and N is the number of predictions.

Clearly, accuracy is one of the most important criteria for forecasting models, the other being the decision improvements generated from directional predictions. From the business point of view, the latter is more important than the former. For business practitioners, the aim of forecasting is to support or improve decisions, so as to make more money. But in crude oil price forecasting, improved decisions usually depend on correct forecasting of direction, of actual price, and predicted price, $x(t)$ and $\hat{x}(t)$. The ability to predict movement direction can be measured by a directional statistic (D_{stat}) (Yu et al., 2007b, 2008), which can be expressed as

$$D_{stat} = \frac{1}{N}\sum_{t=1}^{N} a_t \times 100\% \tag{8}$$

where a_t=1 if$(x_{t+1} - x_t)(\hat{x}_{t+1} - x_t) \geq 0$, and a_t=0 otherwise.

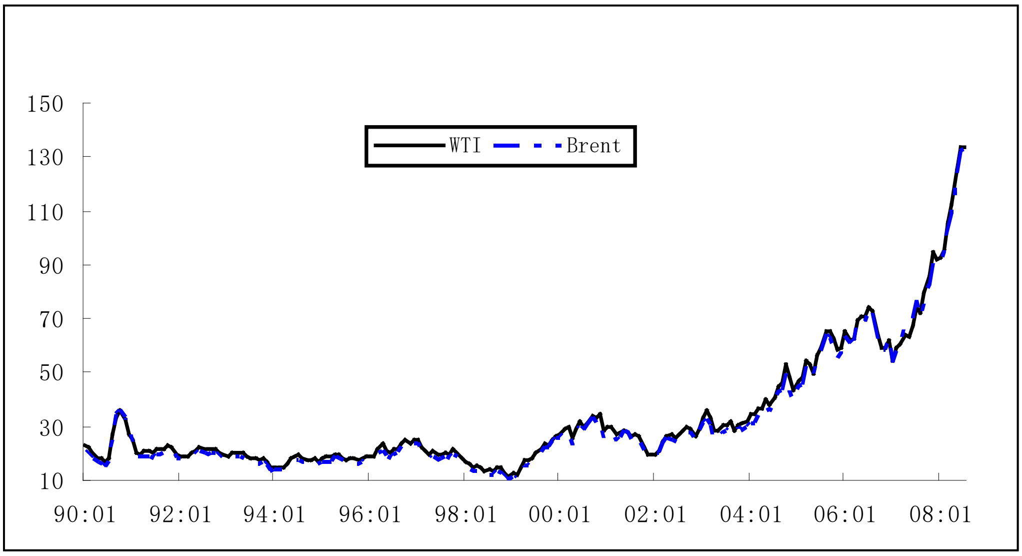

Figure 2. Monthly data of WTI and Brent oil prices for the period 1990-2008

Experimental Results

Before applying the ARIMA model, an augmented Dickey-Fuller (ADF) test should be done. The results of the ADF test show that the two crude oil price time series follows the unit root process. In order to utilize the ARIMA model, the first-order difference is necessary. Thus, ARIMA (p, 1, d) is used. Note that ARIMA (p, 1, q) has p autoregressive terms and q moving-average terms after the first-order difference. As previously mentioned in Section 2.2, we use rolling forecasting with sliding windows data and, therefore, values of p and q are different for each prediction of WTI and Brent oil price. The best ARIMA model for each training sample is selected, based on Schwarz Criterion (SC) minimization, but the maximum p and q is restricted to only 3.

From the experience of the ARIMA modeling, a fractional integrated ARIMA (p, 1, q) is estimated for each training sample. The selection criterion is also SC minimization since Schmidt and Tschernig (1995) proved that the SC criterion is the most robust model selection criteria for ARFIMA model identification. But the maximum p and q are set to 3, as in case of the ARIMA model. The estimation method is exact maximum likelihood (EML), proposed by Sowell (1992). Both ARIMA and ARFIMA models are applied by the ARFIMA package 1.1 based on OX console 5.1.

For the MS-ARFIMA model, the number of regimes is set to two. The WTI price prediction model is built on the basis of SC minimization but Brent price forecasting model is built in terms of AIC minimization because for WTI crude oil price prediction the model based on SC outperforms the model based on AIC, and for Brent crude oil price prediction AIC outperforms SC. But the maximum p is restricted to 3 and q is to 0. The MS-ARFIMA model is implemented by MSVAR 1.3.2 package (Krolzig, 1998) (This package is implemented on OX console 3.4), combined with ARFIMA package 1.0.4, based on OX console 5.1.

For neural network models, FNN (I-H-O) models are used, where FNN (I-H-O) denotes that the FNN model has I input neurons, H hidden nodes, and O output neurons. In this study, we use 5 input neurons, 9 hidden neurons and 1 output neuron in terms of the results of trial and error method. For SVM models, the Gaussian RBF kernel is used. Regularization constant C and the kernel parameter σ^2 are determined by the trial and error method. Due to the changing estimation window in rolling forecasting, these parameter values of SVM models will change for each prediction. In this study, 43 experiments are made for each crude oil price series and thus we will have 172 parameter values for the two crude oil price series (i.e. 43 C and 43 σ^2 for WTI and Brent crude oil price respectively, because of rolling estimation window). Due to space constraints, these values are not provided here. In addition, time delay parameter s is set to 5.

Using the above settings, prediction results for the two crude oil price series are computed. **Figures 3-4** graphically depict the price forecasting results of the two crude oil prices using different models. From the figures, we can roughly see that the performance of the SVM-based crude oil price forecasting approach is better than of other crude oil price forecasting models listed in this study. For the FNN model, the neural network toolbox (Version 5.0) of Matlab software package is used. In the SVM model, LS-SVMlab1.5 toolbox (Pelckmans et al., 2003) is adopted, based on the Matlab platform.

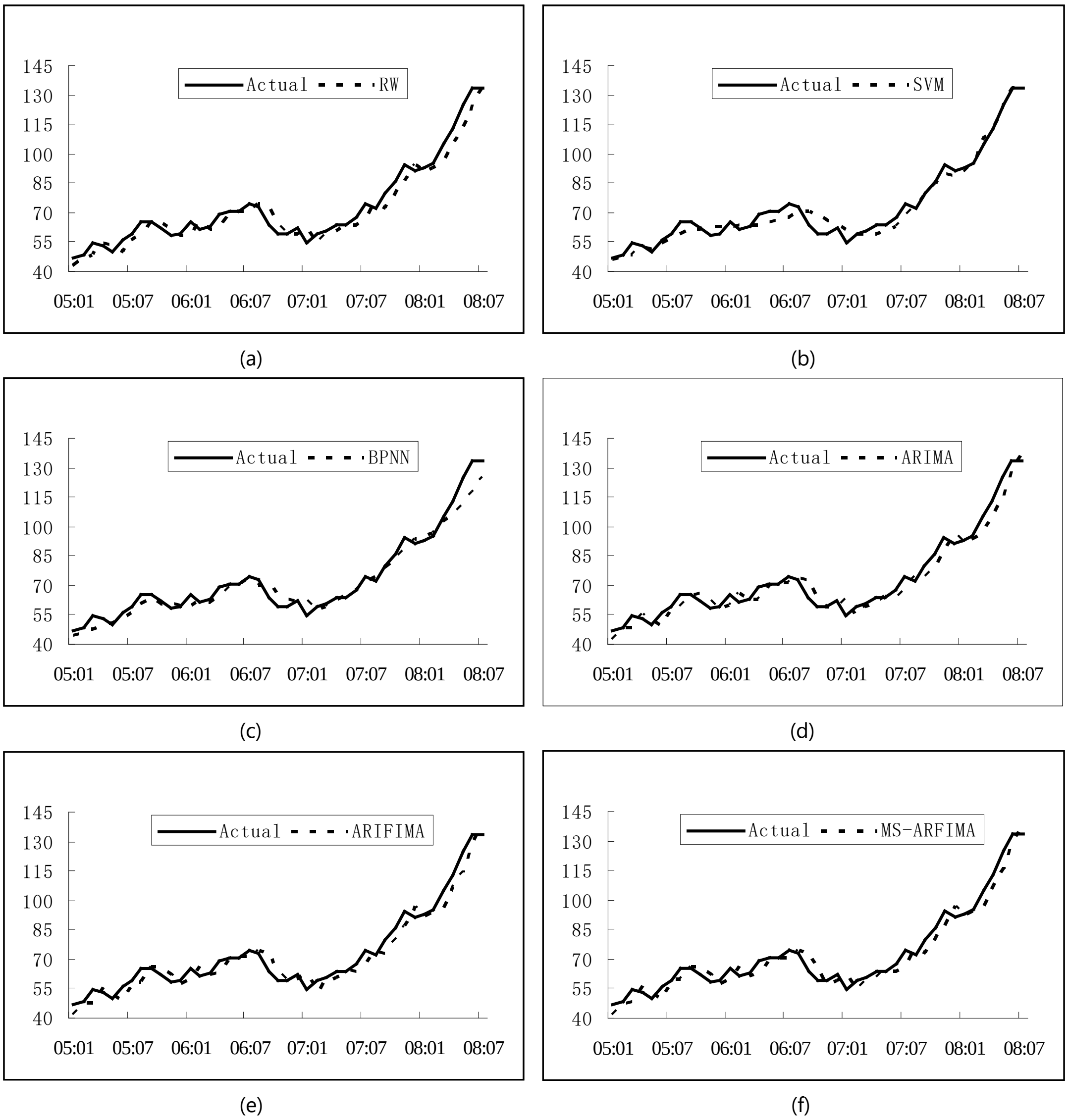

Figure 3. WTI crude oil price prediction results of each model

Tables 1 to **2** show the forecasting performance of different models from different perspectives. **Table 1** shows comparison of *RMSE* of WTI crude oil price prediction by the SVM-based forecasting model and other models listed in this study. Similarly, **Table 2** provides comparison of the directional prediction (D_{stat}) of Brent crude oil price prediction by different models. From the tables, it is easy to find that the SVM-based forecasting approach is very promising for all crude oil price series forecasting under study, whether the measurement of forecasting performance is goodness of fit, such as *RMSE* (refer to **Table 1**), or it is D_{stat} (refer to **Table 2**), indicating that the proposed SVM-based forecasting model has good potentiality in crude oil price forecasting, relative to other models listed in this study.

In case of the *RMSE* criterion, the SVM-based crude oil price forecasting approach performs the best in all cases. Furthermore, *RMSE* results, in case of WTI, are slightly better than in case of Brent crude oil price. The possible reason could be that WTI crude oil prices have larger volatility than Brent crude oil prices due to the fact that WTI is the most important price indicator in international crude oil markets. The other five methods produce some interestingly mixed results. The FNN model gives *RMSE* results for WTI that are better than those for Brent. There are two possible reasons. On one hand, the FNN model is in a class of unstable predictors because of initial random weights, thus making the prediction results unstable in crude oil price forecasting. On the other hand, there may be some local minima in FNN while forecasting Brent crude oil price. In ARIMA, ARFIMA and MS-ARFIMA models, we find that MS-ARFIMA is generally better than the other two. The ARFIMA model fails to show forecasting ability superior to ARIMA in case of Brent crude oil price. As we know, the advantage of the ARFIMA model, compared to ARIMA, is the ability to capture the long memory characteristic, but the advantage of MS-

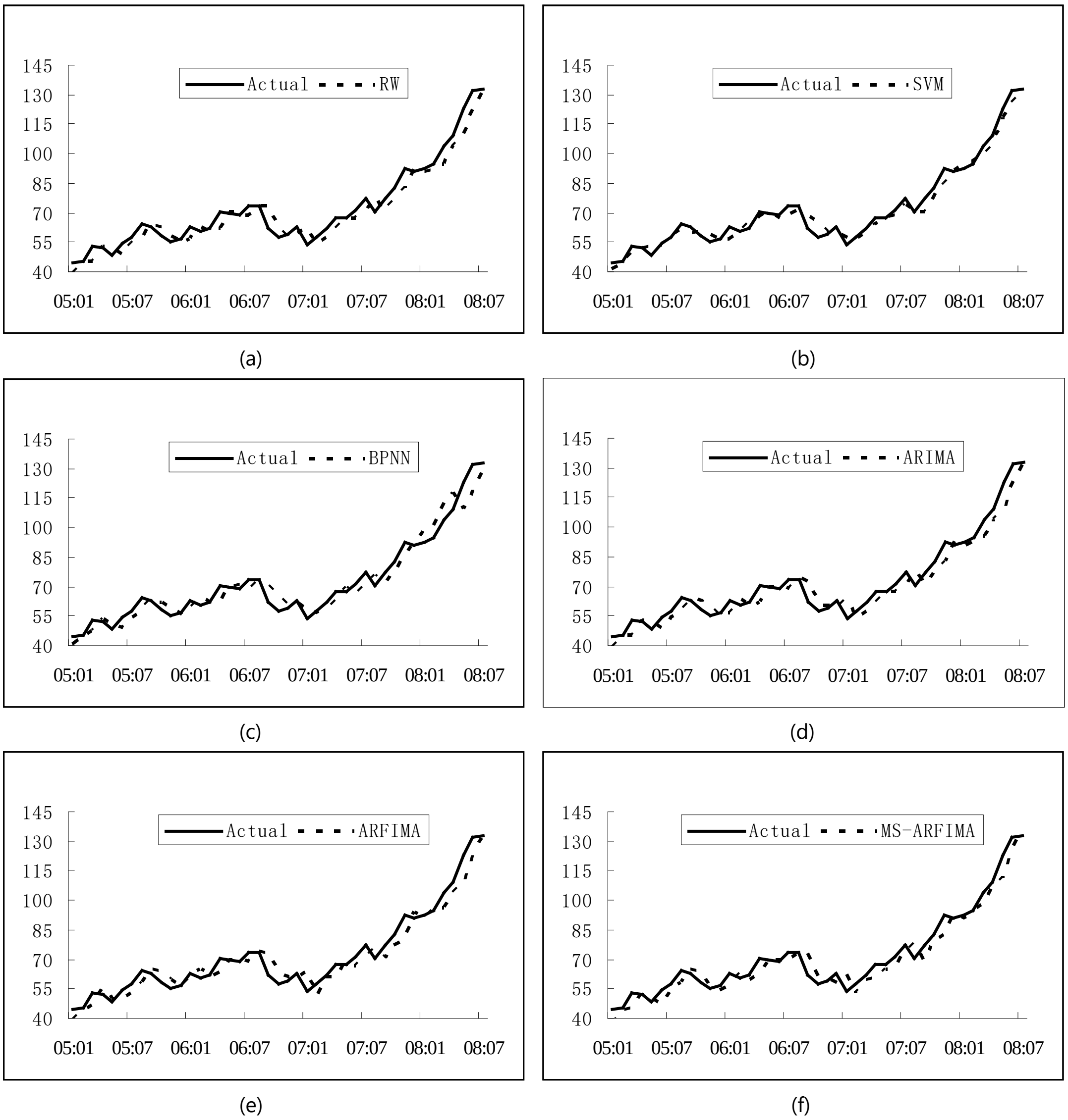

Figure 4. Brent crude oil price prediction results of each model

Table 1. The *RMSE* comparisons for different models

Models	WTI		Brent	
	RMSE	**Rank**	***RMSE***	**Rank**
RW	5.1545	6	5.5195	6
SVM	3.9337	1	3.6042	1
FNN	4.8682	3	5.4572	5
ARIMA	5.0493	5	5.4167	3
ARFIMA	4.9956	4	5.4420	4
MS-ARFIMA	4.8591	2	5.2786	2

ARFIMA over ARFIMA lies in the ability to effectively capture the nonlinearity hidden in crude oil price series. This may suggest that nonlinearity is more important for Brent crude oil price prediction than the long memory characteristic. RW models show the worst results, implying that crude oil prices are not unpredictable. This confirms the results of Ye et al. (2005).

However, the low *RMSE* does not necessarily mean that there is a high hit rate in forecasting of crude oil price movement direction, which is more important for business practitioners. Thus, the D_{stat} comparison is necessary. In **Table 2**, we find that in terms of D_{stat} also, the proposed SVM-based forecasting model performs much better than other models. Furthermore, from the business practitioners' point of view, D_{stat} is more important than *RMSE* because D_{stat} is an important decision criterion for investments in crude oil market. With reference to **Table 2**, variations between different models are very significant. For example, in case of WTI forecasting, D_{stat} for the RW

Table 2. The D_{stat} comparisons for different models

Models	WTI		Brent	
	D_{stat} (%)	Rank	D_{stat} (%)	Rank
RW	58.14	6	55.81	5
SVM	74.42	1	76.74	1
FNN	65.12	2	69.77	2
ARIMA	65.12	2	65.12	3
ARFIMA	60.47	5	58.14	4
MS-ARFIMA	62.79	4	53.49	6

Table 3. Diebold-Mariano test between different models for WTI crude oil price forecasting

	FNN	MS-ARFIMA	ARFIMA	ARIMA	RW
SVM	-0.6472	-1.4733	-1.5686	-1.4554	-1.4217
	(0.2588)	(0.0703)	(0.0584)	(0.0728)	(0.0776)
FNN		0.0088	-0.1335	-0.2049	-0.3627
		(0.5035)	(0.4469)	(0.4188)	(0.3584)
MS-ARFIMA			-1.3086	-0.9135	-0.9294
			(0.0953)	(0.1805)	(0.1763)
ARFIMA				-0.3992	-0.6559
				(0.3449)	(0.2560)
ARIMA					-0.8806
					(0.1893)

model is 58.14%, whereas for FNN and ARIMA models, D_{stat} is 65.12%, and for the SVM-based forecasting model, D_{stat} reaches 74.42%. In the six listed models, SVM and FNN models perform better than the other four statistical-based models. The main reason is that SVM and FNN are two typical nonlinear intelligent models which can capture the nonlinearity in crude oil price series, indicating that the intelligent models have stronger prediction abilities than the statistical-based models. It is interesting that the D_{stat} rank of ARIMA model outperforms ARFIMA and MS-ARFIMA models.

Both the intelligent predictors (SVM and FNN) can, in principle, describe the nonlinear dynamics of crude oil price series. Designed on the basis of the unique theory of the structural risk minimization principle to estimate a function by minimizing an upper bound of the generalization error, SVM is resistant to the over-fitting problem, and can model nonlinear relations in an efficient and stable way. Furthermore, the SVM is trained as a convex optimization problem that produces a global solution that in many cases yields unique solutions. Compared with the SVM's merits, FNN tends to suffer from an over-fitting problem and does not lead to a single global or unique solution, owing to differences in initial weights. Therefore, SVM generally outperforms FNN; empirical results further confirm such an analysis, indicating that SVM has stronger forecasting capabilities than other methods listed in this study.

In order to further compare the predictive accuracy of different forecasting models, the Diebold-Mariano statistic (Diebold and Mariano, 1995) is used to test the statistical significance of forecasts of different models. When comparing two forecasts, the question arises whether predictions of a given model, A, are significantly more accurate, in terms of a loss function, than those of the competing model, B. The Diebold-Mariano test aims to test the null hypothesis of equality of expected forecast accuracy against the alternative of different forecasting abilities across models. In this study, the loss function is set to mean square prediction error (*MSPE)* and the null hypothesis is that the *MSPE* of a specific model is less than that of another model. The statistical testing results are shown in **Tables 3** and **4**. Note that the results listed in **Tables 3** and **4** are the Diebold-Mariano test values, and p values are in brackets.

From figures in **Tables 3** and **4**, we can draw the following conclusions. First of all, for WTI crude oil price prediction, the proposed SVM-based forecasting model outperforms MS-ARFIMA, ARFIMA, ARIMA and RW models at 10% statistical significance level. However, the SVM-based forecasting model does not significantly outperform the FNN-based forecasting model.

Table 4. Diebold-Mariano test between different models for Brent crude oil price forecasting

	FNN	MS-ARFIMA	ARFIMA	ARIMA	RW
SVM	-2.1141	-6.9339	-4.9006	-3.8811,	-3.7139
	(0.0173)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
FNN		0.2813	0.0310	0.0936	-0.1702
		(0.6108)	(0.5124)	(0.5373)	(0.4324)
MS-ARFIMA			-0.8342	-0.5425	-0.5480
			(0.2021)	(0.2937)	(0.2918)
ARFIMA				0.2738	-0.7847
				(0. 6079)	(0.2163)
ARIMA					-1.2923
					(0.0981)

Second, for Brent crude oil price forecasting, the proposed SVM-based prediction approach performs much better than MS-ARFIMA, ARFIMA and RW model at 1% statistical significance level. At the same time, the SVM-based forecasting model outperforms the FNN-based forecasting model at 5% statistical significance level. This statistically demonstrates the potentiality of SVM approach in crude oil price forecasting.

Third, in WTI crude oil price forecasting, MS-ARFIMA model can perform better than ARFIMA model at 10% statistical significance. The main reason is that MS-ARFIMA model can capture nonlinear patterns hidden in crude oil price series while ARFIMA cannot. Other than this, there is no significant difference between other models in WTI crude oil price forecasting, according to the results in **Table 3**.

Finally, in Brent crude oil price forecasting, ARIMA model can significantly outperform the RW model at 10% statistical significance level. In addition, there is no statistically significant difference between other models in WTI crude oil price forecasting in terms of results in **Table 4**.

CONCLUSIONS

In this study, the potentiality of applying SVM model to crude oil price prediction is assessed by empirical investigation. Experimental results obtained confirm the potentiality of SVM model in crude oil price forecasting, which is explained by two competitive advantages of SVM model over the conventional statistical-based forecasting models and ANN-based forecasting models. On the one hand, due to adoption of the structure risk minimization (SRM) principle, the SVM model provides better generalization capability than the conventional statistical models and ANN-based models. On the other hand, the SVM model can eliminate the typical drawbacks of conventional models, e.g., local minima and over-fitting problems, in terms of empirical risk minimization (ERM) principle, and thus obtain more stable and robust generalization results, relative to conventional methods. In addition, the SVM model has fewer free parameters than the ANN model. Regularization constant C and kernel parameter σ^2 are the two factors that need to be considered in the SVM model in this study. While in ANN model, the network architecture, learning parameters estimation, and the network training algorithm greatly affect prediction performance and thus extra care is required during simulation (Tay and Cao, 2001).

In summary, empirical results find that among different forecasting models used for the two main crude oil price series (WTI crude oil price and Brent crude oil price), in terms of different criteria, the SVM-based forecasting model performs the best. In all testing cases, *RMSE* is the lowest and D_{stat} is the highest, indicating that the SVM-based forecasting approach can provide a promising alternative and offers advantages in crude oil price time series forecasting.

However, in crude oil price forecasting, many issues deserve to be studied further. For example, irregular events have an important impact on crude oil price fluctuations; how to incorporate effects of irregular events on crude oil price volatility into predictions is a major issue. In addition, application of the SVM method in crude oil price forecasting is a good and interesting attempt and it may be worth testing its utility in other fields also. We will look into these issues in the future.

ACKNOWLEDGEMENTS

This work is partially supported by grants from the National Natural Science Foundation of China (NSFC No. 71433001, 71622011 and 71301006), the National Program for Support of Top Notch Young Professionals, and Beijing Advanced Innovation Center for Soft Matter Science and Engineering.

REFERENCES

Abramson, B., & Finizza, A. (1995). Probabilistic forecasts from probabilistic models: a case study in the oil market. *International Journal of Forecasting, 11*(1), 63-72.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*, 121-167.

Cao, L. J., & Tay, F. E. H. (2001). Financial forecasting using support vector machines. *Neural Computing Applications, 10*, 184-192.

Cristianini, N., & Taylor, J. S. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, New York.

Chou, J. R. (2016). An Empirical Study of User Experience on Touch Mice. *Eurasia Journal of Mathematics, Science & Technology Education*, *12*(11), 2875-2885.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*, 253-263.

Doornik, J. A., & Ooms, M. (2001). *A package for estimating, forecasting and simulating ARFIMA models: ARFIMA Package 1.1 for Ox*. Working paper, Nuffield College, Oxford.

Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis, 1*, 15-29.

Gulen, S. G. (1998). Efficiency in the crude oil futures market. *Journal of Energy Finance & Development, 3*, 13-21.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Huntington, H. G. (1994). Oil price forecasting in the 1980s: what went wrong? *The Energy Journal, 15*(2), 1-22.

Kaboudan, M. A. (2001). Compumetric forecasting of crude oil prices. *The Proceedings of IEEE Congress on Evolutionary Computation*, 283-287.

Krolzig, H. M. (1998). *Econometric modelling of markov-switching vector autoregressions using MSVAR for Ox*, Working Paper, Nuffield College, Oxford.

Lanza, A., Manera, M., & Giovannini, M. (2005). Modeling and forecasting cointegrated relationships among heavy oil and product prices. *Energy Economics, 27*, 831-848.

Lee, D. K., & Lee, E. S. (2016). Analyzing team based engineering design process in computer supported collaborative learning. *Eurasia Journal of Mathematics, Science & Technology Education*, *12*(4), 767-782.

Mirmirani, S., & Li, H. C. (2004). A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. *Advances in Econometrics, 19*, 203-223.

Morana, C. (2001). A semiparametric approach to short-term oil price forecasting. *Energy Economics, 23*(3), 325-338.

Muller, K. R., Smola, J. A., & Scholkopf, B. (1997). Prediction time series with support vector machines. *Proceedings of International Conference on Artificial Neural Networks*, Lausanne, 999-1004.

Panas, E., & Ninni, V. (2000). Are oil markets chaotic? A non-linear dynamic analysis. *Energy Economics, 22*, 549-568.

Pelckmans, K., Suykens, J. A. K., Van Gestel, T., De Brabanter, J., Lukas, L., Hamers, B., De Moor, B., & Vandewalle, J. (2003). *LS-SVMlab Toolbox User's Guide (version 1.5), ESAT-SCD-SISTA Technical Report 02-145*, Katholieke Universiteit Leuven.

Schmidt, C. M., & Tschernig, R. (1995). The identification of fractional ARIMA models. *Sonderforschungsbereich, 373*, Humboldt Universitaet Berlin.

Shambora, W. E., & Rossiter, R. (2007). Are there exploitable inefficiencies in the futures market for oil? *Energy Economics, 29*, 18-27.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics, 53*, 165-188.

Tay, F. E. H., & Cao, L. J. (2001a). Applications of support vector machines in financial time series forecasting. *Omega, 29*, 309-317.

Tay, F. E. H., & Cao, L. J. (2001b). Improved financial time series forecasting by combining support vector machines with self-organizing feature map. *Intelligent Data Analysis, 5*, 339-354.

Tay, F. E. H., & Cao, L. J. (2002). Modified support vector machines in financial time series forecasting. *Neurocomputing, 48*, 847-861.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. New York: Springer.

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks, 10*, 988-999.

Vapnik, V. N., Golowich, S. E., & Smola, A. J. (1996). Support vector method for function approximation, regression estimation, and signal processing. *Advances in Neural Information Processing Systems, 9*, 281-287.

Wang, S. Y., Yu, L., & Lai, K. K. (2004). A novel hybrid AI system framework for crude oil price forecasting. *Lecture Notes in Computer Science, 3327*, 233-242.

Wang, S. Y., Yu, L., & Lai, K. K. (2005). Crude oil price forecasting with TEI@I methodology. *Journal of Systems Sciences and Complexity, 18*(2), 145-166.

Watkins, G. C., & Plourde, A. (1994). How volatile are crude oil prices? *OPEC Review, 18*(4), 220-245.

Weigend, A. S., & Gershenfeld, N. A. (1994). *Time Series Prediction: Forecasting the Future and Understanding the Past*. Addison-Wesley, Reading, MA.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica, 64*, 1067-1084.

White, H. (1990). Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings. *Neural Networks, 3*, 535-549.

Ye, M., Zyren, J., & Shore, J. (2002). Forecasting crude oil spot price using OECD petroleum inventory levels. *International Advances in Economic Research, 8*, 324-334.

Ye, M., Zyren, J., & Shore, J. (2005). A monthly crude oil spot price forecasting model using relative inventories. *International Journal of Forecasting, 21*, 491-501.

Ye, M., Zyren, J., & Shore, J. (2006). Forecasting short-run crude oil price using high and low-inventory variables. *Energy Policy, 34*, 2736-2743.

Yu, L., Lai, K. K., Wang, S. Y., & He, K. J. (2007a). Oil price forecasting with an EMD-based multiscale neural network learning paradigm. *Lecture Notes in Computer Science, 4489*, 925-932.

Yu, L., Wang, S. Y., & Lai, K. K. (2007b). *Foreign-Exchange-Rate Forecasting with Artificial Neural Networks*. Springer, New York.

Yu, L., Wang, S. Y., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics, 30*(5), 2623-2635.

Zhang, X., Lai, K. K., & Wang, S. Y. (2008). A new approach for crude oil price analysis based on empirical mode decomposition. *Energy Economics, 30*(3), 905-918.

http://www.ejmste.com