Energy Economics 30 (2008) 2623–2635

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm

Lean Yu [a,*], Shouyang Wang [a], Kin Keung Lai [b]

[a] Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[b] Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 12 December 2007
Received in revised form 3 May 2008
Accepted 3 May 2008
Available online 17 May 2008

JEL classification:
C45
C53
Q49

Keywords:
Empirical mode decomposition
Ensemble learning
Feed-forward neural network
Adaptive linear neural network
Crude oil price prediction

ABSTRACT

In this study, an empirical mode decomposition (EMD) based neural network ensemble learning paradigm is proposed for world crude oil spot price forecasting. For this purpose, the original crude oil spot price series were first decomposed into a finite, and often small, number of intrinsic mode functions (IMFs). Then a three-layer feed-forward neural network (FNN) model was used to model each of the extracted IMFs, so that the tendencies of these IMFs could be accurately predicted. Finally, the prediction results of all IMFs are combined with an adaptive linear neural network (ALNN), to formulate an ensemble output for the original crude oil price series. For verification and testing, two main crude oil price series, West Texas Intermediate (WTI) crude oil spot price and Brent crude oil spot price, are used to test the effectiveness of the proposed EMD-based neural network ensemble learning methodology. Empirical results obtained demonstrate attractiveness of the proposed EMD-based neural network ensemble learning paradigm.

1. Introduction

Crude oil has been playing an increasingly important role in the world economy since nearly two-third of the world's energy demands is met from crude oil (Alvarez-Ramirez et al., 2003). It is said that crude oil is also the world's largest and most actively traded commodity, accounting for over 10% of total world trade (Verleger, 1993). As a special commodities, crude oil is traded internationally among many different players—oil producing nations, oil companies, individual refineries, oil importing nations, and speculators. Like most commodities, crude oil price is also basically determined by its supply and demand (Hagen, 1994; Stevens, 1995), but it is strongly influenced by many irregular past/present/future events like weather, stock levels, GDP

* Corresponding author. Tel.: +86 10 62565817; fax: +86 10 62541823.
E-mail address: yulean@amss.ac.cn (L. Yu).

0140-9883/$ – see front matter
doi:10.1016/j.eneco.2008.05.003

growth, political aspects and even people's psychological expectations. Furthermore, since it takes considerable time to ship crude oil from one country to another, oil prices vary in different parts of the world. These factors lead to a strongly fluctuating crude oil market, which has the characteristics of complex nonlinearity, dynamic variation and high irregularity (Watkins and Plourde, 1994). Unfortunately the fundamental mechanism governing the complex dynamics is not well understood by human beings. In addition, as sharp oil price movements can disturb aggregate economic activity, crude oil price fluctuations may have a significant impact on a nation's economy. From the different perspectives, the impacts of crude oil price fluctuations on a nation's economy are reflected in two ways. On one hand, sharp increases in crude oil prices adversely influence economic growth and accelerate inflation for oil importing economies. On the other hand, a fall in crude oil prices, like the one in 1998, will generate serious budgetary deficit problems for oil exporting countries (Abosedra and Baghestani, 2004). Therefore, volatile oil prices are of considerable interest to many institutions and business practitioners, as well as academic researchers. As such, crude oil price forecasting is a very important topic, albeit an extremely hard one, due to its intrinsic difficulties and high volatility (Wang et al., 2005). As the crude oil spot price series are usually considered a nonlinear and nonstationary time series, which is interactively affected by many factors, predicting crude oil price accurately is rather challenging.

In the past decades, traditional statistical and econometric techniques, such as linear regression (LinR), co-integration analysis, GARCH models, naive random walk, vector auto-regression (VAR) and error correction models (ECM) have been widely applied to crude oil price forecasting. For example, Huntington (1994) applied a sophisticated econometric model to predict crude oil prices in the 1980s. Abramson and Finizza (1995) utilized a probabilistic model for predicting oil prices, and Morana (2001) suggested a semi-parametric statistical method for short-term oil price forecasting based on the GARCH properties of crude oil price. Similarly, Barone-Adesi et al. (1998) suggested a semi-parametric approach for oil price forecasting. Gulen (1998) used co-integration analysis to predict the West Texas Intermediate (WTI) price. Ye et al. (2002, 2005, 2006) presented a simple econometric model of WTI prices, using OECD petroleum inventory levels, relative inventories, and high- and low-inventory variables. Mirmirani and Li (2004) used the VAR model to predict U.S. oil price. Lanza et al. (2005) investigated crude oil and oil products' prices using error correction models (ECM).

Usually, the above models can provide good prediction results when the 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 the traditional statistical and econometric models are built on linear assumptions and they cannot capture the nonlinear patterns hidden in the crude oil price series.

Due to the limitations of the traditional statistical and econometric models, some nonlinear and emerging artificial intelligent (AI) models, such as nonlinear regression, artificial neural networks (ANN), support vector machines (SVM) and genetic programming (GP), provide powerful solutions to nonlinear crude oil price prediction. For example, Abramson and Finizza (1991) used belief networks, a class of knowledge-based models, to forecast crude oil prices. Kaboudan (2001) employed GP and ANN to forecast crude oil price. Tang and Hammoudeh (2002) proposed a nonlinear regression model to forecast OPEC basket price. Mirmirani and Li (2004) used the ANN model with genetic algorithm (GA) to predict crude oil price and compared the results with the VAR model. Xie et al. (2006) proposed a support vector regression (SVR) model to predict crude oil price. Similarly, Shambora and Rossiter (2007) and Yu et al. (2007a) also used the ANN model to predict crude oil price. Many experiments found that the AI-based models often had some advantages over statistical-based models. However, these AI models also have their own shortcomings and disadvantages. For example, ANN often suffers from local minima and overfitting, while other AI models, such as SVM and GP including ANN, are sensitive to parameter selection.

To remedy the above shortcomings, some hybrid methods have been used recently to predict crude oil price and obtain the best performances. For example, Wang et al. (2004) developed a hybrid AI system framework by means of a systematic integration of ANN and rule-based expert system, with web text mining, to predict crude oil price. Wang et al. (2005) proposed a TEI@I methodology for crude oil price forecasting and obtained good prediction performance. Likewise, Amin-Naseri and Gharacheh (2007) proposed a hybrid AI approach integrating feed-forward neural networks, genetic algorithm, and *k*-means clustering, to predict the monthly crude oil price and obtain satisfactory results. The basic idea of the above hybrid methods is to overcome the drawbacks of individual models and to generate a synergetic effect in forecasting.

Motivated by hybrid and TEI@I methodologies, this study attempts to apply the "divide-and-conquer" principle to construct a novel crude oil spot price forecasting methodology. In this study, the generic idea of "divide-and-conquer" principle can be understood as "decomposition-and-ensemble". The main aim of decomposition is to simplify the forecasting task, while the goal of ensemble is to formulate a consensus forecasting on original data. In terms of the generic idea, an empirical mode decomposition (EMD) based neural network ensemble learning paradigm is proposed for crude oil spot price forecasting. In this proposed methodology, the difficult crude oil price forecasting task is divided into several relatively easy subtasks, by decomposing the original crude oil price series into some independent sub-series or components. In concrete terms, the original crude oil spot price series, with characteristics of nonlinearity and nonstationarity, were first decomposed into a finite and often small number of intrinsic mode functions (IMFs), using EMD technique (Huang et al., 1998). After these simple IMF components are adaptively extracted via EMD, from a nonlinear and nonstationary time series, each IMF component is modeled by an independent three-layer feed-forward neural network (FNN) model (Hornik et al. 1989; White, 1990), such that the tendencies of these IMF components can be accurately predicted. Finally, prediction results of all IMF components are aggregated, using an adaptive linear neural network (ALNN) (Hagan et al., 1996) to produce an ensemble forecasting result for the original crude oil price series. The main advantage of selecting EMD as a decomposition tool is that the EMD technique, based on Hilbert–Huang Transform (HHT), is very suitable for decomposing nonlinear and nonstationary time series; it has been reported to have worked better, in describing the local time scale instantaneous frequencies, than the wavelet decomposition and Fourier decomposition (Huang et al., 1999; Li, 2006). In wavelet decomposition, it needs to determine a filter function before decomposition (Li, 2006), while in Fourier decomposition, a time series (either linear or nonlinear) can be decomposed into a set of linear components. As the degree of nonlinearity and nonstationarity in a time series increases, the Fourier decomposition often produces large sets of physically meaningless harmonics, when it is applied to nonlinear time series decomposition (Huang et al., 1999).

The main motivation of this study is to propose an EMD-based neural network ensemble learning approach for crude oil spot price prediction and compare its prediction performance with some existing forecasting techniques. The rest of this study is organized as follows. Section 2 describes the formulation process of the proposed EMD-based neural network ensemble learning paradigm in detail. For illustration and verification purposes, two main crude oil price series, West Texas Intermediate (WTI) crude oil spot price and Brent crude oil spot price are used to test the effectiveness of the proposed methodology, and the corresponding results are reported in Section 3. Finally, some concluding remarks are drawn in Section 4.

2. Methodology formulation

In this section, the overall process of formulating the EMD-based neural network ensemble paradigm is presented. First of all, the EMD technique and neural network-based forecasting method are briefly reviewed. Then the EMD-based neural network ensemble learning methodology is proposed. Finally, the overall steps of the EMD-based neural network ensemble learning methodology are summarized.

2.1. Empirical mode decomposition (EMD)

The empirical mode decomposition (EMD) technique, first proposed by Huang et al. (1998), is a form of adaptive time series decomposition technique using the Hilbert–Huang transform (HHT) for nonlinear and nonstationary time series data. 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 the following two prerequisites:

(1) In the whole data series, the number of extrema (sum of maxima and minima) and the number of zero crossings, must be equal, or differ at most by one, and
(2) The mean value of the envelopes defined by local maxima and minima must be zero at all points.

With these two requirements, some meaningful IMFs can be well defined. Otherwise, if one blindly applied the technique to any data series, the EMD may result in a few meaningless harmonics (Huang et al., 1999). Usually, an IMF represents a simple oscillatory mode, compared with the simple harmonic function. Using the definition, any data series $x(t)$ (t=1, 2,…, n) can be decomposed, according to the following sifting procedure.

1) Identify all the local extrema, including local maxima and local minima, of $x(t)$,
2) Connect all local extrema by a cubic spline line to generate its upper and lower envelopes $x_{\text{up}}(t)$ and $x_{\text{low}}(t)$.
3) Compute the point-by-point envelope mean $m(t)$ from upper and lower envelopes, i.e. $m(t)=(x_{\text{up}}(t)+x_{\text{low}}(t)/2)$.
4) Extract the details, $c(t)=x(t)-m(t)$.
5) Check the properties of $c(t)$: (i) if $c(t)$ meets the above two requirements, an IMF is derived and meantime replace $x(t)$ with the residual $r(t)=x(t)-c(t)$; (ii) if $c(t)$ is not an IMF, replace $x(t)$ with $c(t)$.

Repeat Steps 1)–5) until the stop criterion is satisfied. According to the above prerequisites, a typical stop criterion can be defined by the following three conditions: (a) at each point, (mean amplitude)< (threshold2×envelope amplitude), (b) mean of Boolean array ((mean amplitude)/(envelope amplitude)> threshold)<tolerance, and (c) the number of zero crossings and the number of extrema is less than or equal to one, i.e. |#zeros−#extrema|<=1. In the three conditions, threshold, threshold2 and tolerance must be specified before the sifting procedure is performed (Huang et al., 1998).

The EMD extracts the next IMF by applying the above sifting procedure to the residual term $r_1(t)=x(t)-c_1(t)$, where $c_1(t)$ denotes the first IMF. The decomposition process can be repeated until the last residue $r_n(t)$ has at most one local extremum or becomes a monotonic function, from which no more IMFs can be extracted. The above sifting procedure can be implemented using Matlab software.

At the end of this sifting procedure, the data series $x(t)$ can be expressed by

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

where n is the number of IMFs, $r_n(t)$ is the final residue, which is the main trend of $x(t)$, and $c_j(t)$ (j=1, 2,…, n) are the IMFs, which are nearly orthogonal to each other, and all have nearly zero means. Thus, one can achieve decomposition of the data series into n-empirical mode functions and one residue. The IMF components contained in each frequency band are different and they change with variation of time series $x(t)$, while $r_n(t)$ represents the central tendency of data series $x(t)$.

Relative to the traditional Fourier and wavelet decompositions, the EMD technique has several distinct advantages. First of all, it is relatively easy to understand and implement. Second, the fluctuations within a time series are automatically and adaptively selected from the time series, and it is robust for nonlinear and nonstationary time series decomposition. Third, it lets the data speak for themselves. EMD can adaptively decompose a time series into several independent IMF components and one residual component. The IMFs and the residual component displaying linear and nonlinear behavior depend only on the nature of the time series being studied. Finally, in wavelet decomposition, a filter base function must be determined beforehand, but it is difficult for some unknown series to determine the filter base function. Unlike wavelet decomposition, EMD is not required to determine a filter base function before decomposition. In terms of the above merits the EMD can be used as an effective decomposition tool.

2.2. Artificial neural networks (ANNs)

Artificial neural networks (ANNs) are a class of typical intelligent learning paradigm, widely used in some practical application domains. In this study, a standard three-layer feed-forward neural network (FNN) (Hornik et al., 1989; White, 1990), based on error back-propagation algorithm, is selected for modeling the decomposed IMFs and the residual component, while the adaptive linear neural network is chosen for aggregating the results produced by FNN, based on individual IMFs.

Usually, a FNN-based forecasting model can be trained by the in-sample dataset and applied to out-of-sample dataset for prediction. The model parameters (connection weights and node biases) are adjusted iteratively by a process of minimizing the forecasting error function. Basically, the final output of the FNN-based forecasting model can be represented as

$$f(x) = a_0 + \sum_{j=1}^{q} w_j u\left(a_j + \sum_{i=1}^{p} w_{ij} x_i\right) \tag{2}$$

where x_i (i=1, 2,..., p) represents the input patterns, $f(x)$ is the output, a_j (j=0, 1, 2,..., q) is a bias on the jth unit, and w_{ij} (i=1, 2,..., p; j=1, 2,..., q) is the connection weight between layers of the model; φ(· is the transfer function of the hidden layer, p is the number of input nodes, and q is the number of hidden nodes. Actually, the FNN model in Eq. (2) performs a nonlinear functional mapping from past observation (x_{t-1}, x_{t-2},..., x_{t-p}), to the future value x_t, i.e.

$$\mathrm{x}_t = \varphi\left(x_{t-1}, x_{t-2}, \ldots, x_{t-p}, w\right) + \xi_t \tag{3}$$

where w is a vector of all parameters and φ is a function determined by neural network training. Thus, from some perspectives, the FNN model is equivalent to a nonlinear autoregressive (NAR) model (Yu et al., 2005, 2007b).

The main reason of selecting FNN as a predictor is that it is often viewed as a "universal approximator" (Hornik et al., 1989). Hornik et al. (1989) and White (1990) found that a three-layer feed-forward neural network (FNN) with an identity transfer function in the output unit and logistic functions in the middle-layer units can approximate any continuous function arbitrarily well, given a sufficient amount of middle-layer units. That is, neural networks have the ability to provide a flexible mapping between inputs and outputs. In this study, we utilize the three-layer FNN forecasting method for modeling the decomposed IMFs and the residual component.

For prediction results produced by FNNs, we use an adaptive linear neural network (ALNN) to fuse the results. ALNN is a single-layer neural network, where the transfer function is a pure linear function, and the learning rule is Widrow–Hoff (i.e. least mean squared (LMS) rule) (Hagan et al., 1996). Typically, the mathematical representation can be represented by

$$\mathrm{f}(x) = \varphi\left(\sum_{i=1}^{m} w_i x_i + b\right) \tag{4}$$

where x_i (i=1, 2,..., m) represents the input variables, $f(x)$ is the output, b is a bias, w_i (i=1, 2,..., m) is the connection weight, m is the number of input nodes, and $\varphi(\cdot)$ is the transfer function of the single-layer ALNN. Usually, the LMS rule adjusts the weights and bias so as to minimize the mean square error (MSE) as follows:

$$\mathrm{MSE}=\frac{1}{N}\sum_{j=1}^{N}\left[e_j(x)\right]^2 = \frac{1}{N}\sum_{j=1}^{N}\left[T_j(x)-P_j(x)\right]^2 \tag{5}$$

where N is the number of the data samples and $e_j(x)$ is the error between the target output $T_j(x)$ and the network output $P_j(x)$. The change to the weight and bias for multiple neurons can be written as

$$w(x+1) = w(x) + 2\eta e(x) T^T(x) \tag{6}$$

$$b(x+1) = b(x) + 2\eta e(x) \tag{7}$$

where η is a learning rate. To obtain a good convergence to the optimal weight and bias, the learning rate must be less than the reciprocal of the largest eigenvalue of the correlation matrix $\boldsymbol{x}^T\mathbf{x}$ of the input vectors.

The main goal of the ALNN is to integrate the prediction results produced by three-layer FNNs, for ensemble purpose. It is worth noting that the FNN and ALNN are implemented using neural network toolbox of Matlab software, and the version number is 5.0 (R2006a), which is released on 03 February, 2006.

2.3. Overall process of the EMD-based neural network ensemble paradigm

Suppose there is a time series $x(t)$, $t=1, 2,\ldots, N$, in which one would like to make the l-step ahead prediction, i.e. $x(t+l)$. For example, $l=1$ means one single-step ahead prediction and $l=30$ represents 30-step ahead prediction. Depending on the previous techniques and methods, an EMD-based neural network ensemble paradigm can be formulated, as illustrated in Fig. 1.

As can be seen from Fig. 1, the proposed EMD-based neural network ensemble forecasting paradigm is generally composed of the following three main steps:

1) The original time series $x(t)$, $t=1, 2,\ldots, N$ is decomposed into n IMF components, $c_j(t)$, $j=1, 2,\ldots, n$, and one residual component $r_n(t)$ via EMD.
2) For each extracted IMF component and the residual component, the three-layer FNN model is used as a forecasting tool to model the decomposed components, and to make the corresponding prediction for each component.
3) The prediction results of all extracted IMF components and the residue produced by FNNs in the previous step are combined to generate an aggregated output using an ALNN model, which can be seen as the final prediction result for the original time series.

To summarize, the proposed EMD-based neural network ensemble forecasting paradigm is actually an "EMD–FNN–ALNN" ensemble learning approach. That is, it is an "EMD (Decomposition)–FNN (Prediction)–ALNN (Ensemble)" methodology. In order to verify the effectiveness of the proposed EMD-based neural network ensemble methodology, two main crude oil price series, West Texas Intermediate (WTI) crude oil spot price and Brent crude oil spot price are used for testing purpose in the next section.

3. Experiments

3.1. Research data and evaluation criteria

As is known to all, there are a great number of crude oil price series. In this study, two main crude oil price series, West Texas Intermediate (WTI) crude oil spot price and Brent crude oil spot price are chosen as experimental samples. The main reason of selecting these two oil price indicators is that these two crude oil prices are the most famous benchmark prices, which are used widely as the basis of many crude oil price

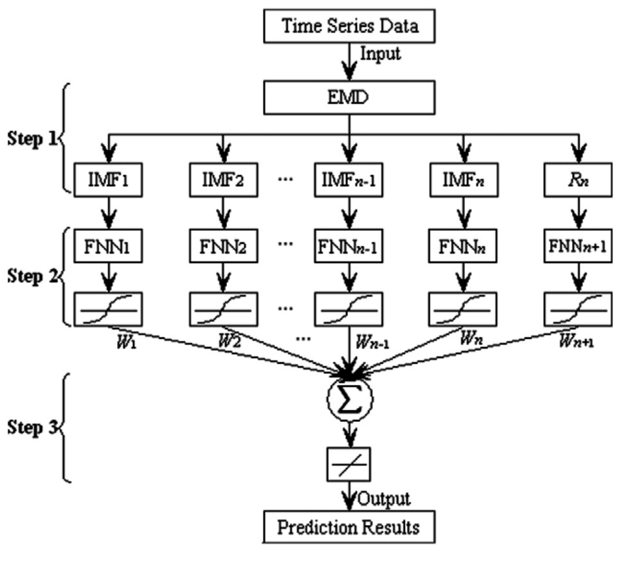

Fig. 1. The overall process of the EMD-based neural network ensemble methodology.

formulae. The two crude oil price data used in this study are daily data, and are freely obtainable from the energy information administration (EIA) website of Department of Energy (DOE) of USA (http://www.eia.doe.gov/).

For WTI crude oil spot price, we take the daily data from January 1, 1986 to September 30, 2006, excluding public holidays, with a total of 5237 observations. For convenience of neural network modeling, the data from January 1, 1986 to December 31, 2000 is used for the training set (3800 observations), and the remainder is used as the testing set (1437 observations).

For Brent crude oil spot price, the sampling data covers the period from May 20, 1987 to September 30, 2006 with a total of 4933 observations. The main reason of different starting points is that the EIA website only provides the Brent data since May 20, 1987. Similarly, we take the data from May 20, 1987 to December 31, 2002 as in-sample (training periods) training set (3965 observations), and take the data from January 1, 2003 to September 30, 2006 as out-of-sample (testing period) testing set (968 observations), which is used to evaluate the performance of prediction, based on evaluation criteria. Note that only one-step-ahead prediction is performed in the experiments. Actually, multi-step-ahead prediction, e.g., step size is equal to 30, can also be performed, but the prediction performance in such cases is unsatisfactory. For this purpose, a long-term forecasting is also performed to predict the crude oil price from 2008 to 2012 based on monthly data before 2008. That is, we utilize the data before 2008 to estimate the model and then make the prediction from 2008 to 2012.

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

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{x}(t)-x(t)\right)^2} \tag{8}$$

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 directions, 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., 2005, 2007b), which can be expressed as

$$_{\mathrm{Dstat}} = \frac{1}{N}\sum_{t=1}^{N} a_t x 100\% \tag{9}$$

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

In order to compare the forecasting capability of the proposed EMD-based neural network ensemble learning methodology with other popular forecasting approaches, the autoregressive integrated moving average (ARIMA) model and the single FNN model are used as the benchmark models.

In an ARIMA model (Box and Jenkins, 1970), the future value of a variable is assumed to be a linear function of several past observations and random errors. That is, the underlying process that generates the time series takes the form

$$\phi(B)y_t = \theta(B)e_t \tag{10}$$

where y_t and e_t are the actual value and random error at time t respectively; B denotes the backward shift operator, i.e. By_t=$y_{t-1,}$ ^2y_t B^2y_t= y_{t-2} and so on; and $\phi(B)$=1−ϕ_1B− −ϕ_pB^p, $\theta(B)$=1−θ_1B− −θ_qB^q, where p, q are integers and often referred to as lag orders of the model. Random errors, e_t, are assumed to be independently and identically distributed with a mean of zero and a constant variance of σ^2, i.e. e_t~$IID(0, \sigma^2)$. If the dth difference of $\{y_t\}$ is an ARIMA process of order p and q, then y_t is called an ARIMA (p–d–q) process.

In addition, three variants of the ensemble model, the EMD–ARIMA–ALNN model, the EMD–FNN–Averaging model and the EMD–ARIMA–Averaging model, are also used to predict crude oil prices for comparison purposes.

3.2. Experimental results

In this study, all ARIMA models are implemented via the Eviews software package, which is produced by Quantitative Micro Software Corporation. Individual FNN models and the EMD-based neural network ensemble learning model are built using the neural network toolbox (Version 5.0) of Matlab software package, which is produced by Mathworks Laboratory Corporation. The EMD–ARIMA–ALNN model uses ARIMA model to predict IMFs extracted by EMD, and applies ALNN for combination. The EMD–FNN–Averaging model applies the FNN to predict all IMFs and integrates the predicted results, using a simple averaging method, while the EMD–ARIMA–Averaging model utilizes ARIMA method to predict the IMFs, and then combines the prediction results with a simple averaging strategy. Thus, ARIMA, FNN, EMD–ARIMA–ALNN ensemble, EMD–FNN–Averaging ensemble, EMD–ARIMA–Averaging ensemble, and the proposed EMD–FNN–ALNN ensemble models are all used to predict the two main crude oil prices for comparison purposes.

According to previous steps shown in Section 2.3, we start to perform the prediction experiments. First, using the EMD technique, the two typical crude oil price series can be decomposed into several independent IMFs and one residue. Before decomposition, the threshold and tolerance levels of the stop criterion are specified by [threshold, threshold2, tolerance] = [0.05, 0.5, 0.05]. Using EMD, we can get graphical representations of the decomposed results for two main crude oil prices, as illustrated in Figs. 2 and 3. Particularly, Fig. 2 shows the decomposition results for the WTI crude oil spot price series. Clearly, the WTI crude oil spot price series is decomposed into ten IMFs and one residue. Similarly, Fig. 3 illustrates the results of decomposition of the Brent crude oil price series using EMD, into nine IMFs and one residue. These results help us to improve the prediction performance by using the “divide-and-conquer” strategy.

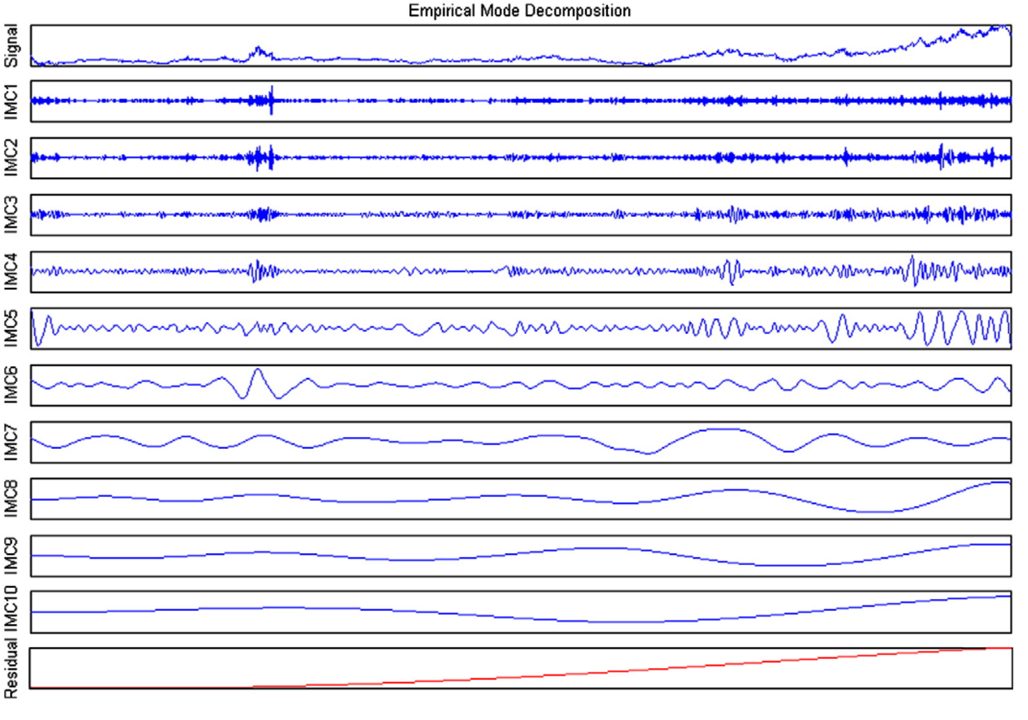

Fig. 2. The decomposition of WTI crude oil spot price.

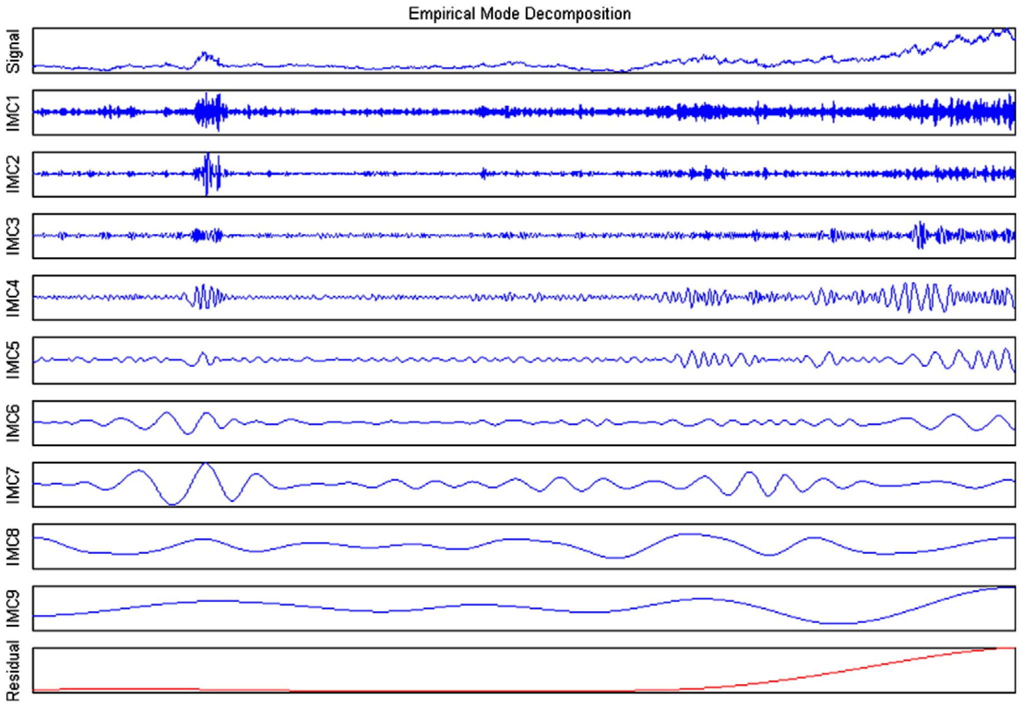

Fig. 3. The decomposition of Brent crude oil spot price.

Using the above mentioned decomposed IMFs and the residue, different ensemble methods listed above can be used for forecasting purposes. It is worth noting that two single models (ARIMA and FNN) utilize the original time series for prediction. In the experiments, ARIMA (*p*–*d*–*q*) and FNN (*I*–*H*–*O*) are used for multiple testing. Note that ARIMA (*p*–*d*–*q*) suggests that ARIMA has *p* autoregressive terms and *q* moving average terms, after *d*-order difference, while FNN (*I*–*H*–*O*) denotes that the FNN model has *I* input neurons, *H* hidden nodes, and *O* output neurons. Particularly, the logistic and linear functions are used for the transfer function in the hidden and output layers respectively, as Hornik et al. (1989) and White (1990) indicated. The above parameters can be determined by trail-and-error in the two single models, and final parameters estimation results are shown in Tables 1 and 2 for WTI and Brent respectively.

For the four ensemble models, we use the decomposed IMFs to perform the experiments. In the EMD–FNN–ALNN model, *k* FNN models with different topological structures are used for prediction, with the extracted *k* IMFs produced by EMD, while the ALNN (*k*, linear) are used for aggregation (Note that ALNN (*k*,

Table 1
Parameters estimation results of the ARIMA for WTI series

Variable	Coefficient	Std. error	*t*-statistic	Prob.
C	19.86757	1.463476	13.57560	0.0000
AR(1)	0.988271	0.016174	61.10103	0.0000
AR(2)	−0.039110	0.022779	−1.716960	0.0861
AR(3)	−0.048284	0.022779	−2.119670	0.0341
AR(4)	0.092879	0.016177	5.741555	0.0000
R-squared	0.985620	Mean dependent var		19.82458
Adjusted *R*-squared	0.985604	S.D. dependent var		4.692276
S.E. of regression	0.562987	Akaike info criterion		1.690195
Sum squared resid	1201.573	Schwarz criterion		1.698417
Log likelihood	−3202.991	*F*-statistic		64,957.86
Durbin–Watson stat	1.996748	Prob (*F*-statistic)		0.000000

Table 2
Parameters estimation results of the ARIMA for Brent series

Variable	Coefficient	Std. error	t-statistic	Prob.
AR(1)	−1.558157	0.016578	−93.98989	0.0000
AR(2)	−0.955386	0.015950	−59.89976	0.0000
MA(1)	1.556909	0.014629	106.4242	0.0000
MA(2)	0.965582	0.013660	70.68464	0.0000
R-squared	0.002873	Mean dependent var		0.008223
Adjusted R-squared	0.002266	S.D. dependent var		0.645857
S.E. of regression	0.645124	Akaike info criterion		1.962064
Sum squared resid	2050.129	Schwarz criterion		1.967340
Log likelihood	−4832.487	Durbin–Watson stat		1.972227

linear) represents that the ALNN model has k input neurons and one linear transfer function). In the EMD–FNN–Averaging model, k FNN models are used to predict the k IMFs extracted by EMD, similar to the process of modeling the EMD–FNN–ALNN, and k prediction results are integrated by simple averaging. In the EMD–ARIMA–ALNN methodology, k ARIMA with different autoregressive terms and moving average terms are used for prediction; the ALNN (k, linear) model is used for integration, while the EMD–ARIMA–Averaging methodology adopts a simple averaging method to fuse the prediction results produced by k ARIMA models. Note that k is the number of decomposed IMF components and one residual component. In this study, k is equal to 11 for WTI and k is equal to 10 for Brent series, in terms of decomposition results shown in Figs. 2 and 3.

Using the above settings, the evaluation of prediction results for the two main crude oil price series are shown in Tables 3 and 4 via RMSE and D_{stat}. From the two tables, we can generally see that the forecasting results of the proposed EMD-based neural network ensemble learning approach are very promising for all crude oil prices under study, either where the measurement of forecasting performance is goodness-of-fit, such as RMSE (refer to Table 3), or where the forecasting performance criterion is D_{stat} (refer to Table 4), indicating that the prediction performance of the proposed EMD-based neural network ensemble forecasting model is better than that of other models listed in this study.

Focusing on the RMSE indicator, the EMD-based neural network ensemble learning approach (EMD–FNN–ALNN method) performs the best in all cases, followed by EMD–FNN–Averaging model, EMD–ARIMA–ALNN model, individual BPNN model, EMD–ARIMA–Averaging model and the individual ARIMA model. Interestingly, the RMSEs of the single BPNN model for predictions of the two crude oil price series are better than those of the EMD–ARIMA–ALNN and EMD–ARIMA–Averaging methodology, as well as the single ARIMA models. The possible reasons could be two-fold. On one hand, the BPNN model is a class of nonlinear forecasting model, which can capture nonlinear patterns hidden in the crude oil prices. On the other hand, the crude oil market is a high-volatility market and the crude oil prices often show nonlinear and nonstationary patterns, while the EMD–ARIMA–ALNN model, EMD–ARIMA–Averaging model and the single ARIMA model are linear models, which are not suitable for predicting crude oil price with high volatility and irregularity. In addition, we also find that the EMD–ARIMA–ALNN model and EMD–ARIMA–Averaging model perform much better than the single ARIMA model. Likewise, prediction performances of EMD–FNN–ALNN and EMD–FNN–Averaging models are better than that of the single FNN model. The possible reason could be that the EMD decomposition impacts the prediction performance.

Table 3
The RMSE comparisons for different methods

Methodology	WTI		Brent	
	RMSE	Rank	RMSE	Rank
EMD–FNN–ALNN	0.273	1	0.225	1
EMD–FNN–Averaging	0.509	2	0.457	2
EMD–ARIMA–ALNN	0.975	4	0.872	4
EMD–ARIMA–Averaging	1.769	5	1.392	5
Single FNN	0.841	3	0.743	3
Single ARIMA	2.035	6	1.768	6

Table 4
The D_{stat} comparisons for different methods

Methodology	WTI		Brent	
	D_{stat} (%)	Rank	$D_{stat}D_{stat}$ (%)	Rank
EMD–FNN–ALNN	86.99	1	87.81	1
EMD–FNN–Averaging	77.45	2	79.44	2
EMD–ARIMA–ALNN	62.98	4	68.91	4
EMD–ARIMA–Averaging	60.13	5	62.29	5
Single FNN	69.03	3	75.41	3
Single ARIMA	52.47	6	57.75	6

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. Focusing on D_{stat} of Table 4, we find that the proposed EMD-based neural network ensemble learning model also performs much better than other models, according to the ranks. Furthermore, from the business practitioners' point of view, D_{stat} is more important than RMSE because the D_{stat} can be seen as an important decision criterion in investments in crude oil market. With reference to Table 4, the differences among the different models are very significant. For example, for the WTI test case, the D_{stat} for the individual ARIMA model is 52.47%, for the individual FNN model it is 69.03%, and for the EMD–FNN–Averaging model, D_{stat} is 77.45%, while for the EMD-based neural network ensemble forecasting model, D_{stat} reaches 86.99%. Ranks of the other three ensemble models, in terms of forecasting accuracy, is in the middle range, for any of the crude oil price series. The main reason of this phenomenon is that the bad performance of individual models and the ensemble strategy (i.e. simple averaging) have an important effect on the overall forecast efficiency. Similarly, the single FNN can model nonlinear time series, such as crude oil price series, well, and the D_{stat} rank is also third, for the two crude oil price series. In the same way, we also notice that the D_{stat} ranking for ARIMA is the lowest. The main reason is that high noise, nonlinearity and complex factors are characteristic of crude oil price series, while the ARIMA model is a class of linear models.

In addition, to show the predictability of the proposed EMD-based neural network ensemble learning model, an out-of-sample forecasting will be performed for WTI and Brent from 2008 to 2012. Based on the previous experiment designs and model descriptions, the out-of-sample forecasting results are illustrated in Fig. 4. In particular, the data from 2004 to 2007 are also illustrated in Fig. 4 for comparison purposes.

From Fig. 4, an interesting finding is that the overall tendency of crude oil price is upward since 2004 and after 2008 the trend of crude oil price will be downward. That is, the year 2008 will be a turning point for crude oil price evolution process. The main reasons leading to such predictions reflect three aspects. First of all, as the blood of national economy, strong growth of world economy (particularly the economic

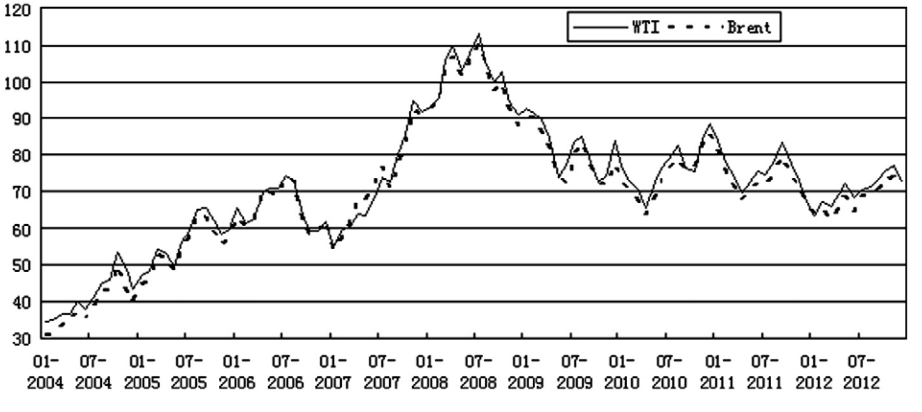

Fig. 4. The out-of-sample prediction for WTI and Brent from 2008 to 2012.

development of China, India and other developing economies) largely increases the demand of crude oil and thus pushing the crude oil price upward since 2001. Second, the intense political situations in main oil production districts such as Mid-East and Africa increase the uncertainties of crude oil supply, therefore making the crude oil price higher and higher in recent years. Third, with the constant increase of crude oil price, some substitute energies will be created and innovated. Thus the current crude oil price will lower step by step, but overall price level will be still in a high position, which is consistent with our predictions and “mean-reversion” principle.

Generally speaking, in terms of the experiments presented in this study, we can draw the following several conclusions. (1) The experimental results show that the EMD-based neural network ensemble learning model is superior to the single ARIMA model, the single FNN model, the EMD–ARIMA–Averaging and EMD–ARIMA–ALNN model, as well as the EMD–FNN–Averaging model, for the test cases of the two main crude oil prices, in terms of accuracy level of prediction, as measured by RMSE, and directional prediction statistics (D_{stat}), as can be seen from Tables 3 and 4. (2) The prediction performance of the EMD–ARIMA–ALNN model and EMD–ARIMA–Averaging model are much better than that of the single ARIMA model. Likewise, the EMD–FNN–ALNN and EMD–FNN–Averaging models perform better than the single FNN model. This indicates that the “divide-and-conquer” principle or “decomposition-and-ensemble” strategy can effectively improve the prediction performance, and the results emphasize that the EMD decomposition is meaningful to prediction performance improvement in crude oil price forecasting. (3) The proposed EMD-based neural network ensemble learning approach is able to improve forecasting accuracy significantly; in other words, the performance of the EMD-based neural network ensemble forecasting model is better than other forecasting models, in terms of RMSE and D_{stat}. This leads to the fourth conclusion. (4) The EMD-based neural network ensemble learning model can be used as an alternative solution to world crude oil spot price forecasting.

4. Concluding remarks

This study proposes using an EMD-based neural network ensemble learning model to predict world crude oil spot prices. In terms of empirical results, we find that across different forecasting models, for the two main crude oil prices – WTI crude oil spot price and Brent crude oil spot price – in terms of different criteria, the EMD-based neural network ensemble learning model performs the best. In all testing cases, the RMSE is the lowest and the D_{stat} is the highest, indicating that the EMD-based neural network ensemble forecasting paradigm can be used as a very promising methodology for world crude oil price prediction.

Acknowledgments

The authors would like to thank the Editor-in-Chief, Prof. Dr. Richard Tol, of the Energy Economics Journal, and two anonymous referees for their valuable comments and suggestions. Their comments helped improve the quality of the paper immensely. This work is partially supported by grants from the National Natural Science Foundation of China (NSFC No. 70601029, 70221001), the Knowledge Innovation Program of the Chinese Academy of Sciences, and the Strategic Research Grant of City University of Hong Kong.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.eneco.2008.05.003.

References

Abosedra, S., Baghestani, H., 2004. On the predictive accuracy of crude oil future prices. Energy Policy 32, 1389–1393.
Abramson, B., Finizza, A., 1991. Using belief networks to forecast oil prices. International Journal of Forecasting 7 (3), 299–315.
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.
Alvarez-Ramirez, J., Soriano, A., Cisneros, M., Suarez, R., 2003. Symmetry/anti-symmetry phase transitions in crude oil markets. Physica A 322, 583–596.

Amin-Naseri, M.R., Gharacheh, E.A., 2007. A hybrid artificial intelligence approach to monthly forecasting of crude oil price time series. The Proceedings of the 10th International Conference on Engineering Applications of Neural Networks, CEUR-WS284, pp. 160–167.
Barone-Adesi, G., Bourgoin, F., Giannopoulos, K., 1998. Don't look back. Risk 100–103 August 8.
Box, G.E.P., Jenkins, G., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, CA.
Gulen, S.G., 1998. Efficiency in the crude oil futures market. Journal of Energy Finance & Development 3, 13–21.
Hagan, M.T., Demuth, H.B., Beale, M.H., 1996. Neural Network Design. PWS Publishing Company, Boston.
Hagen, R., 1994. How is the international price of a particular crude determined? OPEC Review 18 (1), 145–158.
Hornik, K., Stinchocombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366.
Huang, N.E., Shen, Z., Long, S.R., 1999. A new view of nonlinear water waves: the Hilbert spectrum. Annual Review of Fluid Mechanics 31, 417–457.
Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences 454, 903–995.
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.
Lanza, A., Manera, M., Giovannini, M., 2005. Modeling and forecasting cointegrated relationships among heavy oil and product prices. Energy Economics 27, 831–848.
Li, X., 2006. Temporal structure of neuronal population oscillations with empirical mode decomposition. Physics Letters A 356, 237–241.
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.
Shambora, W.E., Rossiter, R., 2007. Are there exploitable inefficiencies in the futures market for oil? Energy Economics 29, 18–27.
Stevens, P., 1995. The determination of oil prices 1945–1995. Energy Policy 23 (10), 861–870.
Tang, L., Hammoudeh, S., 2002. An empirical exploration of the world oil price under the target zone model. Energy Economics 24, 577–596.
Verleger, P.K., 1993. Adjusting to Volatile Energy Prices. Institute for International Economics, Washington DC, USA.
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.
White, H., 1990. Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings. Neural Networks 3, 535–549.
Xie, W., Yu, L., Xu, S.Y., Wang, S.Y., 2006. A new method for crude oil price forecasting based on support vector machines. Lecture Notes in Computer Science 3994, 441–451.
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., 2005. A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. Computers & Operations Research 32 (10), 2523–2541.
Yu, L., Wang, S.Y., Lai, K.K., 2007b. Foreign-Exchange-Rate Forecasting with Artificial Neural Networks. Springer, New York.