COMPUTATIONAL INTELLIGENCE TECHNIQUES WITH APPLICATION TO CRUDE OIL PRICE PROJECTION: A LITERATURE SURVEY FROM 2001–2012

Haruna Chiroma[*], *Sameem Abdulkareem*[*], *Adamu Abubakar*[†], *Mohammed Joda Usman*[‡]

Abstract: This paper is an attempt to survey the applications of computational intelligence techniques for predicting crude oil prices over a period of ten years. The purpose of this research is to provide an exhaustive overview of the existing literature which may assist prospective researchers. The reviewed literature covers a spectrum of publications on the proposed model, source of experimental data, period of data collection, year of publication and contributors. The overall trend of the publications in this area of research issued within the last decade is also addressed. The existing body of research has been analyzed and new research directions have been outlined that have been previously ignored. It is expected that researchers across the globe may thus be encouraged to re–direct their attention and resources in order to keep on searching for an optimum solution.

Key words: *Crude oil price, computational intelligence techniques, west Texas intermediate, factors affecting crude oil price fluctuation, uncertainty events*

Received: July 25, 2012
Revised and accepted: October 25, 2013

[*]Haruna Chiroma – Corresponding author, Sameem Abdulkareem
Faculty of Computer Science and Information Technology, Department of Artificial Intelligence, University of Malaya, Kuala Lumpur, Malaysia, E-mail: freedonchi@yahoo.com, chiromamalaya@siswa.um.edu.my, Phone: +60149576297

[†]Adamu Abubakar
Faculty of Information and Communication Technology, Department of Computer Science, International Islamic University Malaysia, Gombak, Kuala Lumpur, Malaysia

[‡]Mohammed Joda Usman
School of Electronic and Information Engineering, Liaoning University of Technology, Jinzhou, China

1. Introduction

Crude oil prices significantly influence the global economy and determine the economic policies of both government and commercial sectors. Since crude oil is a major player in decision making at various levels of management, a better understanding of future prices is necessary before prudent and informed decisions can be made [1]. The decisive influence of crude oil prices in global economic affairs has given rise to an increased research interest in the crude oil market [2]. Chen et al. [3] concluded that the fluctuation of crude oil prices in the global market at present has triggered a growing curiosity in scrutinizing current models and proposing new ones and identifying improved approaches in order to evade the effects of crude oil price instability. For instance, West Texas Intermediate crude oil sold for less than \$70/barrel on the New York Mercantile Exchange (NYMEX) in August 2007, by July 2008 it was sold at \$148/barrel and after another six months the price plummeted again to just above \$35/barrel [4]. As revealed by Jinliang et al. [5], crude oil prices shoot up to \$78.4/barrel in July 2006 and suddenly crashed down to almost \$60/barrel at the end of the same year. Chevillon and Rifflart [6] pointed out that the market forces were responsible factors for deciding the price of crude oil in the last 20 years. However, the crude oil market remains defective in terms of competition and subjected to demand and supply like any other tradable good but is also influenced by political apprehension.

Regular short term crude oil price movements are caused by normal market forces including but not restricted to the US refinery capacity, OPEC crude oil production ceiling, global demand and supply. The volatility of the oil market is caused by uncertain events, such as political conflicts and regional instability, natural disaster, strike and even the act of taking hostages [7]. Consensus among empirical and theoretical studies shows that shockwaves of changing oil prices exert a detrimental effect on microeconomic pointers by inflating cost of production and operation. In general, crude oil price volatility retards business investments due to uncertainty or re-assignment of resources which increase expenses [8].

Successfully applied neural networks paved the way for computational intelligence (CI) techniques which have begun to play a major role in the energy (oil and gas) sector as claimed by [9]. Several CI techniques emerged inspired by original biological systems, for instance artificial neural networks, evolutional computation, simulated annealing and swarm intelligence which copied the functions of the nervous system and applied the principles of natural selection and thermodynamics and even mimicked the specific behavior of insects. Despite the restrictions associated to each of these techniques, they are applied in solving real life problems in the realm of science, technology, business and commerce .The hybridization of two or more of the techniques has proven to eliminate such constrains and lead to a better solution [10]. Recent studies that involved hybridized CI techniques in search for better solutions include but are not limited to Deng et al. [11] who hybridized genetic algorithms, particle swarm optimization and ant colony optimization. Fazel and Gamasaee [12] integrated fuzzy logic and expert systems. Jia et al. [13] fussed particle swarm optimization, chaotic and Gaussian local search. Khashei et al. [14] combined neural networks and fuzzy logic. Guvenc et al. [15] hybridized genetic algorithms and particle swarm optimization. Acampora et al.

[16] aggregated type–2 fuzzy inference mechanisms, ontologies and fuzzy markup language. Patruikic et al. [17] hybridized a support vector machine and particle swarm optimization. Various CI techniques were proposed in the body of existing literature and applied to crude oil price projection in search for a more reliable prediction system that determines realistic future prices of crude oil which can be utilized by both government and private sector decision makers.

A review of the existing literature hitherto published on the application of CI techniques in energy sectors reveals several overall tends. For example, the support vector machine is countable among the most widely used CI tools applied in energy sector. An extensive review of the existing research completed on support vector machine application in oil refining can be found in Saybani et al. [18] who reviewed academic journal articles published between 1998 and 2010. Research into the support vector machine applied within the oil refining process follows the key terms of algorithm used, type of support vector machine kernel, usage of support vector machine within oil refining process, purpose and contributors. The authors stress that the support vector machine can be used as a guide for assessing the success of the support vector machine in the refining process and other domains within the refineries. Soldo [19] presented a survey on the application of artificial neural networks, simulated annealing and other tools in forecasting gas consumption covering the period between 1949 and 2004. His articles are grouped according to the following major areas: forecasting area, forecasting horizon, consumption data used and forecasting tools. The analysis showed that altogether 29 articles were published in the first 55 years and 47 in the last seven years which suggested to the author that this particular area of research was developing.

To the best knowledge of the authors, no literature survey has yet been conducted on the application of CI techniques in crude oil price projection. This is astonishing especially since applied CI techniques have proven so successful. This paper attempts to fill this existing void by conducting an extensive literature survey covering the academic publications of the previous ten years. It focuses on integration of different CI techniques applied to crude oil price projection, the sources and range of experimental data, the results obtained and the overall trend.

This literature survey differs from past literature surveys conducted on the application of CI techniques in the energy sector in three ways. Firstly, it focuses on the application of CI techniques in crude oil price projection whereas that of [18, 19] focused on the oil refining process and forecasting natural gas consumption. The analysis of results obtained through single CI techniques was also not compared to that of hybrid CI in the previous survey. Furthermore, the sources of experimental data and the range of the collected data were absent in the previous studies. Prospective researchers are intended to use this literature survey as a guide to future studies and as a means to disclose novel research directions. Practitioner may find this survey useful as it provides more insight into the projection of crude oil prices which may render important clues helpful in any future decisions made.

The remaining sections of this paper are organized as follows: Section 2 describes the methodology used for searching articles for inclusion in this survey. Section 3 presents the factors which generally affect crude oil prices. Section 4 provides with the basic concept of CI techniques used for crude oil price projection. Section 5 addresses the applications of CI techniques in projecting crude oil prices.

Section 6 contains the analysis of various CI techniques for oil price prediction. The conclusions and proposed new research directions are presented in section 7 and 8 respectively.

2. Methodology

In order to identify published articles on CI techniques which were specifically applied in either hybrid or individual form to predict crude oil prices, research was completed in two distinct phases. Firstly, the body of existent literature was retrieved online through the ACM Digital Library, IEEEXplore, Science Direct, Scopus, Springer Link, Web of Science, Google Scholar, Direct Open Access Journals (DOAJ), Microsoft Academic Search, ProQuest and CiteSeerX (see Table 1). The literature searched covered the period of publications from 1980 to 2012. This initial online search was based on such keywords as neural networks, genetic algorithms, support vector machine, genetic programming, expert systems, fuzzy logic, particle swarm optimization, ant colony, rough set, etc and each of the keyword is attached with crude oil price. The next stage of literature research focused on the publications accessible through the University Malaya (UM) Digital Library containing a thesis database and a repository of e–journals, e–books, web resources and selected bibliographies. However, this search did not yield the expected number of publications required for the literature survey, as only 35 relevant publications could be identified. During the second search phase, every article initially retrieved was scrutinized and reviewed to ensure it met the required criteria for selection before it was included in the literature survey. Each article had to contain an empirical description of applied CI techniques in hybrid or individual form to predict crude oil prices. The criteria set up for screening led to the rejection of a number of articles resulting in the inclusion of a total of 61 articles. These 61 articles were published between 2001 and 2012 but no articles matching the said criteria were found for the years 2002 and 2003.

Engines	URL
Scopus	http://www.scopus.com
Science Direct	http://www.sciencedirect.com
CiteSeerX	http://citeseerx.ist.psu.edu/index
Google Scholar	http://scholar.google.com
DOAJ	http://www.doaj.org/
IEEExplore	http://ieeexplore.ieee.org
MS Academic Search	http://academic.research.microsoft.com/
UM Library	http://www.umlib.um.edu.my/
ProQuest	http://search.proquest.com/pqdtft/index
Springer Link	http://www.springerlink.com/
Web of Science	http://webscience.org/

Tab. I *URL for the search engines through which research literature were identified.*

Furthermore, articles describing the factors affecting crude oil prices and other relevant literature were also included in the survey. In total, 98 publications (journal articles, conference papers, theses, web resources and monographs) were successfully selected for the purpose of this study.

3. The crude oil market

3.1 Factors affecting crude oil price fluctuation

Ji [20] classified the major factors responsible for influencing crude oil price fluctuation into six categories which are as follows: Macroeconomics, speculation, stock market, supply and demand, exchange rate, and commodity market. Yi and Qin [21] identified the factors politics and military interference as the major contributors to volatile crude oil prices. In a study conducted by Zhang et al. [22] prices of crude oil before, during and after Gulf War of 1991 and the Iraq War of 2003 were analyzed using empirical mode decomposition based events analysis. Historical data were collected from West Texas Intermediate and Brent crude oil prices dating from 30 March 1990 to 31 May 1991. The aim was to assess the impact of unpredictable events on crude oil prices. Empirical evidence from this study indicated that crude oil prices hiked up in the face of unpredictable and uncertain events and dropped to normal figures afterwards. In a related study, Ortiz–Cruz et al. [23] depicted daily crude oil prices on a graph for a period of 25 years extracted from the Energy Information Administration of the US Department of Energy (EIAUSDE). The uncertain events affecting the oil prices were indicated on the graph as shown in Fig. 1. Zhang et al. [24] used crude oil prices of West Texas Intermediate covering the period between 1946 and 2006 which were graphically represented as shown in Fig. 2.

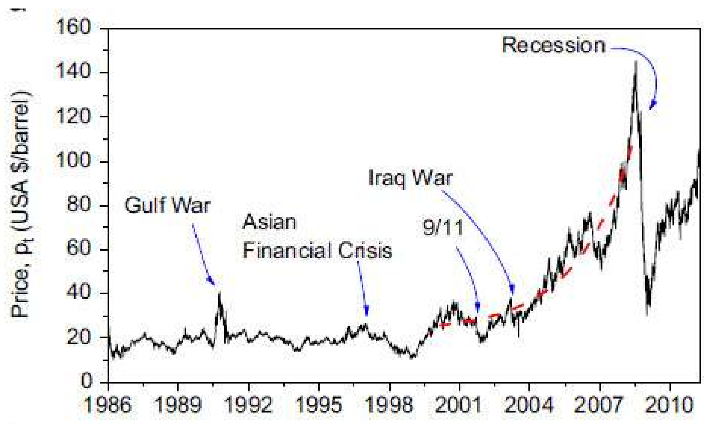

Fig. 1 *West Texas Intermediate crude oil prices over a period covering 1 January 1986 to 15 March 2011. Source: [23].*

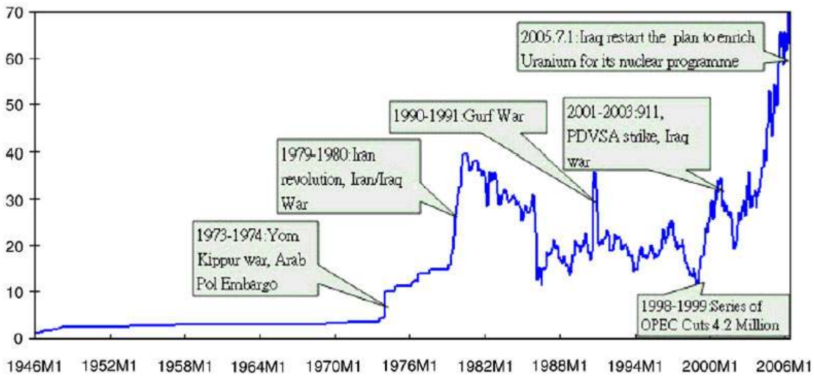

Fig. 2 *West Texas Intermediate crude oil prices from January 1946 to May. Source: [24].*

4. Basic concept of the CI techniques

4.1 Artificial neural network

The original aim of the creation of artificial neural network was to mathematically represent the processing of information in biological systems [25]. The artificial neural network is a system that processes information similar to the human brain and constitutes a general mathematical representation of human reasoning. These networks are built on the following assumptions [26]:

1. Information is processed by neurons
2. Signals are communicated between neurons through established links
3. Every connection between neurons is associated with weight; transmitted signals between neurons are multiplied by the weight.
4. Every network neuron applies an activation function to its input signals so as to regulate its output signal.

4.2 Genetic algorithms

The evolutionary computation community of researchers has not achieved a consensus on the definition of genetic algorithms [27]. According to Barschi [28], genetic algorithms are evolutionary algorithms meaning that they evolved individually in a population by selection, crossover and mutation. Stochastic and heuristic search evolve from genetic algorithms which operate on any data type according to the operators of such data. Genetic algorithms maintain a fixed number of solutions through repeated procedure at every generation. With each new population, individual fitness is evaluated and another population of solutions is created according to the fitness value of each chromosome [29]. Generated members are formed from these two stages and this process is continuously repeated until no improvement of the settings occurs [30].

4.3 Genetic programming

Genetic programming differs from genetic algorithms in terms of problem representation. In genetic programming, a tree is used instead of string. The procedure for searching optimum solutions in both techniques is almost identical. Initially, computer programs such as Boolean, vector, real, complex, symbolic, integer, or multiple values are generated at random. The fitness of each computer program is then evaluated and the program with the highest fitness value selected as the optimum solution once the criteria set for termination is reached. The old populations are replaced by means of crossover, mutation and reproduction [31].

4.4 Particle swarm optimization

The choreographic behavior of birds and insects inspired [32] to propose particle swarm optimization (PSO). The individuals in PSO refine their knowledge of the given search space and have particles referred to as position and velocity. According to this optimization method, two individuals in a population are not engaged in crossover, mutation and particle substitutions in the course of running PSO. PSO operates by inviting to search for the position of the highest fitness in a search space. A memory function is embedded in every particle, and two pieces of information are responsible for adjusting the particle trajectory, namely the best location stayed at present and the best global location reached by the entire swarm. PSO uses an evaluation function to assign fitness values like other optimization techniques. The global best is the highest fitness value reached by a swarm while best local is the highest fitness value that an individual particle has attained. Global and local best are remembered by each particle [32].

4.5 Fuzzy logic

Certain techniques involving principles of human cognition are being commonly referred to as logic. Classical logic is primarily concerned with statements or propositions that are considered as either to be true or false. Variables that represent propositions are considered part of classical logic since any combination of the variables is either true *or* false and cannot be true *and* false or a combination of both. Learning the rules that permit new logical variables to be yielded as functions constitutes the key ingredient. When n logical variables $x_1, x_2, \ldots, x_n$ are given, then x_1 is true, x_2 is false ... x_n is false. Another logical variable y can be defined by a rule as function $x_1, x_2, \ldots, x_n$. Rule: if x_1 is true *and* x_2 is false *and*... *and* x_n is false, then y is false [33].

4.6 Rough set theory

The mathematical basis of rough set theory is built on the indiscernible relation generated by objects. The elementary set is referred to as the set of similar objects forming similar fundamental granules (atoms) of information about the outside world. The union of elementary sets is called 'crisp', else a set is referred to as 'rough set'. In a rough set, objects without certainty are classified as members of the set or compliment whereas a crisp set possesses no element restrictions. Operations of rough set theory are used for pattern detection in data. The technique used in rough sets is also referred to as machine learning, statistics, knowledge discovery and inductive inference. Results generated from operations of rough set theory are subject to interpretation outside the theory. The indiscernibility relation is formulated in form of an information table, information system or attribute value table. If the conditions and decisions are differentiated in an information table, the table is called a decision table [34].

4.7 Expert systems

Expert systems are computer application programs developed to perform expert tasks which provide a solution to a problem in a specific domain. Knowledge of the domain is coded in the application program and the program uses the coded knowledge and specified control strategy to reach a decision. Expert systems are not referred to as a single program because of the necessary integration of several constituents [35]. The main components of expert systems are according to Merritt [36], knowledge base which consists of the expert representation of declarative frequently in *if then* rules; working storage which means the data associated with a solved problem; inference engine which refers to the codes executing recommendations derived from the knowledge base and the associated data; user interface which consists of the code which provides the point of interaction between the user and the system.

5. Applications of hybridized and single CI techniques in crude oil price projection

The CI techniques as described in Section 4 have also been explored in financial predictions. Hybridized and single CI techniques have been applied to predict crude oil prices using voluminous historical data to build prediction models. Part of the hybrid CI technique is to integrate two or more learning models that operate differently to exploits diverse features. The performance exhibited by hybrid learning models is superior to that of single learning models since their original limitations are eliminated by capitalizing on the strengths of each of the single CI technique in the constituent of the hybrid CI technique [37]. However, the application of a hybrid intelligent model does not automatically guarantee superior performance due to the requirement of choosing the correct and compatible hybridization models as well as selecting optimal parameters which significantly determined the performance ability. Documentary evidence shows that the hybrid CI technique is superior to the single CI technique when properly designed [38]. In some cases, CI techniques are fussed with econometrics or statistical tools such as Autoregressive Integrates Moving Average (ARIMA), Autoregressive Moving Average (ARMA), Autoregressive Conditional Heteroskedasticity (GARCH), etc.

In the work of Fan et al. [39] Pattern Modeling and Recognition System (PMRS), Elman Networks (EN) and Generalize Pattern Matching based on Genetic Algorithms (GPMGA) models were used for creating a multi-step prediction tool applicable to crude oil prices. The prediction results indicated that GPMGA outperformed PMRS and EN with an absolute mean percentage error of approximately 2%. In a related study, Ghaffari and Zare [40] collected data set of crude oil prices from West Texas Intermediate from 2004 to 2007 and adopted an adaptive network-based fuzzy inference system (hybrid of neural networks and fuzzy logic) proposed by Roger Jang in 1993 to project crude oil prices. 68.18% prediction accuracy was achieved in the research. In addition, [41] applied the concept of genetic algorithms to optimized architectural parameters of neural networks. Data of spring crude oil prices were collected for a period covering the years 1983 to 2006 extracted from West Texas Intermediate to create a forecasting hybrid system called a neuro–genetic model in order to project monthly crude oil prices. Results suggested that it attained a considerably good prediction accuracy.

Additionally, Mingming and Jinliang [42] built a multiple-wavelet recurrent neural networks model incorporating wavelet and recurrent neural networks. Wavelet was used to capture pattern in the crude oil price historical dataset, and recurrent neural networks was used to forecast crude oil prices at each scale. The data covered the period from 1946 to 2010 acquired from Brent and West Texas Intermediate crude oil representing European and US oil markets. The study showed that the created multi-wavelet recurrent neural networks model possessed the ability to accurately forecast the subsequent year's world crude oil prices.

On the other hand, Yu et al. [43] adopted an empirical mode of decomposition-based neural networks as proposed earlier in 1998 by Huang et al. and integrated it with ensemble learning to form the Empirical Mode Decomposition-Based Neural Network Ensemble Learning Pattern (EMDNNELP). Historical exemplars for crude oil prices were acquired from the EIAUSDE website published by the US Department of Energy for a period of 1986 to 2006. The dataset was used to model EMDNNELP, and its effectiveness was evaluated with historical data from Brent and West Texas Intermediate crude oil for one-step-ahead prediction. ARIMA, empirical mode decomposition–feed forward neural networks–adaptive linear neural networks, empirical mode decomposition- ARIMA, single-feed-forward neural networks and single ARIMA models were also subjected to the process of forecasting crude oil prices for the purpose of evaluating the performance of EMDNNELP. Results from the experiments indicated that the EMDNNELP model outperformed all other models in terms of prediction accuracy.

In another study, Malik and Nasereddin [44] used several models to forecast quarterly growth of Gross Domestic Product (GDP) based on crude oil prices. The models consisted of simple random walk, autoregressive, linear with lagged oil and GDP, cascaded artificial neural networks with GDP only, cascaded artificial neural networks with GDP and oil and conventional artificial neural networks with GDP and oil. Crude oil data for modeling were collected from the Bureau of Labor Statistics and GDP data were collected from Bureau of Economic Analysis for a period from 1947 to 2004. All models were exposed to one-quarter-ahead-forecast and the results indicated that cascaded neural networks outperformed all models considered in this study.

Crude oil price data were obtained from the EIAUSDE for a period from 1985 to 2007 by Azadeh et al. [45] who proposed neural networks–fuzzy regression algorithms used to estimate long term crude oil spot prices. Fuzzy regression and neural networks models were used to predict the prices whereby the neural networks proved superior to the fuzzy–regression but a hybrid model of the two produced better results than each individual model.

In a related study, Bao et al. [46] sourced historical data of crude oil spot prices of Brent and West Texas Intermediate crude oil as recorded by the EIAUSDE for a period from May 1987 to July 2007 and January 1991 to July 2007 respectively. Wavelet and least-square-support vectors were utilized for the price prediction and the experimental results suggested that wavelet transformation gave a better prediction outcome than least-square-support vectors. However, the hybrid model yielded the most superior forecast accuracy. Also, Alaxandridis and Livanis [47] performed a prediction experiment using data collected from the EIAUSDE on West Texas Intermediate crude oil prices for a period from January 1986 to October 2007 and applied wavelet neural networks to predict the crude oil prices of the following one, three and six months.

Ma [48] collected data covering New York Harbor residual oil prices from November to April 2007 and deployed two neural networks models of symbol evolutionary immune clustering neural networks and radial basis function neural networks to forecast crude oil prices. Their compared performance showed that symbol evolutionary immune clustering neural networks performed better than radial basis function neural networks. However, a study conducted by Tehrani and Khodayar [49] which applied genetic algorithms to optimized neural networks and used experimental sample data collected from West Texas Intermediate crude oil spot prices available at EIAUSDE covering a period from 1994 to 2008 showed that the optimized neural networks deployed to predict crude oil prices and predict performance was better than that of conventional neural networks.

Additionally, a neuro–genetic and rule-based expert system was proposed in [50] and applied in this research to predict spot feature prices of gas with sample data retrieved from the EIAUSDE covering a period from January 2004 to December 2008. Enhanced prediction accuracy was attained in this study.

In a separate study, Liu et al. [51] proposed a fuzzy neural network wherein radial basis function neural networks, a Markov chain-based semi-parametric model and wavelet analysis using forecast results served as input to fuzzy neural networks and the target output constituted the actual crude oil prices. Using sample data of crude oil prices for a period from May 1987 to August 2006 acquired from Brent crude oil, a forecasting model was built which was able to capture a pattern in the historical data whose sample test results indicated high estimate correctness.

Mehrara et al. [52] proposed two models in their work for crude oil price prediction, namely group method data handling neural networks and multi-layer feed-forward neural networks. Each prediction model was built based on historical data covering a period from January 2002 to July 2009 sourced from the EIAUSDE. Out-of-sample prediction accuracy of the models was contrasted and seem to confirm that GMDHNN is superior to MLFFNN.

Pan et al. [53] developed multi–layer feed-forward neural networks with four different prediction models which represent from spot price to spot price, from future price to spot price, from spot price and future price to spot price, from spot price and lead markets to spot price in order to predict short term spot prices of crude oil with different indicators. All the models were tested with out-of-sample data of crude oil prices. The prediction results obtained for time (t) +1 was 79.95% accurate, followed by 69.74% and 60.64% for t+2 and t+3 respectively. The experimental data for crude oil prices were collected from West Texas Intermediate and future contrast traded in NYMEX was secured from the EIAUSDE covering the period from January 1996 to August 2007.

Alizadeh and Mafinezhad [7] proposed a Generalized Regression Neural Networks (GRNN) model for crude oil price prediction. Six factors were used as input to the model including US refinery capacity, US GDP, US dollar nominal effective exchange rate, OPEC total liquid capacity, OPEC crude oil production ceiling allocation and US gasoline ending stocks. 12 months average of Brent crude oil prices were used to model the volatility of oil prices as a result of sudden price change within 12 months out of 20 months of experimental data. This was considered in the study as crises index which also served as input to GRNN in order to reflect unexpected price changes. The simulated results generated in this work showed that GRNN possessed a good prediction accuracy. Also, the model possessed the capability of including unexpected changes in crude oil prices when carefully defined and selected. In the same year, Assareh et al. [54] forecasted the total Iranian oil demand from 2006 to 2030 using data collected for Iranian socio–economic indicators comprising population size, GDP, oil consumption, import and export data covering a period from 1981 to 2005. Linear and exponential form of particle swarm optimization and genetic algorithms models were developed in three different scenarios. This research was able to forecast Iranian oil consumption until 2030 in three different scenarios. The results indicated an annual average growth rate of 4.6%, 5.3% and 5.18% for scenario I, II and III respectively as generated by the best performing model (particle swarm optimization demand exponential). Generally, all models presented in this work proved a good forecasting power for oil demand.

In the work of Abdel–Aal [55], a hybrid model of abductive and neural networks was used to forecast one year's energy demand using historical data covering the period from January 1985 to October 1992 obtained from El–Sharkawi (University of Washington, USA). The model was able to forecast a one-year-ahead energy demand with a high level of accuracy.

Wang and Yang [56] examined the probability of predicting crude oil, heating oil, gasoline and natural gas future markets within a day by conducting an experiment with different models (neural networks, semi parametric function coefficient, nonparametric kernel regression and GARCH) with data collected for 30 minutes intraday prices and returns of the four energy future contracts sourced from NYMEX. For each of the four individual future contracts, 15 to 20 year prices of future contracts were analyzed, a period during which prices were low and on steady decline (bear market), another period where prices were high and on a steady increase (bull market) were identified. The result indicated that only heating oil and natural markets possessed the feature of being predictable within a day, especially under bull market conditions.

Yu et al. [57] proposed a CI agent-based fuzzy ensemble prediction model which integrated a support vector machine, radial basis function networks and back-propagation neural networks. The sample data for their study were collected from West Texas Intermediate and Brent crude oil spot prices covering a period from January 2000 to December 2007. Each of the individual models (including ARIMA) were subjected to crude oil forecasting, and the results showed that the CI agent-based fuzzy ensemble prediction model outperformed individual models in terms of prediction accuracy.

Mehdi [58] proposed a fuzzy neural networks model and gathered crude oil weekly prices from January 1989 to January 2009 collected from the Mediterranean Sidi Kerir Iranian light spot price Freight on Board (FOB). The proposed prediction model in this research was able to predict Iranian light spot prices with a high level of accuracy. Uncertainty events influencing prices of crude oil were translated into knowledge-based elements from which were derived the fuzzy rules so as to integrate expert judgment in forecasting crude oil prices and thus complement the proposed model.

Wang et al. [59] proposed a hybrid model of neural networks, rule based expert system and web–based text mining called hybrid CI system. The system used historical data of monthly spot prices of crude oil collected from EIAUSDE for a period from January 1970 to December 2002. Information on uncertainty events affecting crude oil prices were extracted through web–based text mining techniques. The proposed system operated by collecting information and comparing it to pre–defined patterns. If the information was based on irregular events affecting crude oil prices, the rule-based expert system was executed to forecast spot prices. Otherwise neural networks forecasting was executed using historical data. Simulation results showed that the performance of hybrid forecast yielded more accurate results than single neural networks forecasting.

He et al. [60] proposed a Wave Decomposition Network Value at Risk (WD-NEVaR) model to estimate crude oil market values. Historical data in the form of daily closing prices of crude oil were collected over a period from April 1983 to Jun 2006, May 1987 to Jun 2006 and January 1997 to January 2005 sourced from West Texas Intermediate, Brent and Dubai respectively. ARMA–GARCH, Wavelet Decomposition Value at Risk (WDVaR) and WDNEVaR models were used to forecast oil prices in all three markets. The findings demonstrated that the WDNEVaR model (hybrid of wavelet analysis and neural networks) produced considerably superior forecasting results as compared to the WDVaR and ARMA – GARCH models.

Jammazi and Aloui [1] built a hybrid CI model called Harr a Trous Wavelet multilayer back- propagation neural networks (HTW-MBPNN). Monthly crude oil prices were collected from January 1988 to March 2010 in the form of experimental data sourced from the EIAUSDE. The experiment was aimed at forecasting the oil prices for the next 19 months (June 2011 to December 2012). When the HTW–MBPNN forecasted results were compared to conventional back-propagation neural networks, EIAUSDE predictions and West Texas Intermediate future predicted prices indicated that the results generated by HTW–MBPNN were overall more precise than those generated by other models.

Zhang et al. [61] introduced fuzzy time series into prediction of crude oil prices. West Texas Intermediate spot prices data were obtained from the EIAUSDE covering a period from January 1991 to December 2009. Results showed that fuzzy time series could be considered as a good short term forecasting tool.

Kulkar and Haidar [62] collected West Texas Intermediate crude oil prices and future contract traded in NYMEX from September 1996 to August 2007 sourced from the EIAUSDE. The two models of recurrent networks and multilayer feed-forward neural networks were considered and the multilayer feed-forward was selected for the study. Three day moving average was used to remove noise in the data so as to improve forecast accuracy. Crude oil spot prices were forecasted for three days ahead using multilayer feed- forward neural networks. The one day forecast accuracy achieved was 78%, the two days accuracy reached 60%, and the three days accuracy generated 53%.

Chen and Qu [63] examined the prediction of oil field production using production data over the period of ten years (1996 – 2010) using back-propagation neural networks and polybasic linear regression to predict oil production until 2015. The results generated by both models indicated the superiority of neural networks in terms of prediction accuracy. However, Qunli et al. [64] proposed wavelet transform and radial basis function neural networks in their study. Historical data were collected from UK Brent blend crude oil spot price FOB covering a period from January 1997 to October 2008. Wavelet transform was applied on the experimental sample data to improve forecast accuracy. Radial basis function neural networks was also subjected to experimental demonstration to forecast future crude oil prices. Results from the experiment showed that a higher forecast accuracy was achieved by the radial basis function neural networks model.

Wang et al. [65] embedded a jump stochastic time effective function into neural networks in order to improve its forecast accuracy. Data used in the study were extracted from five different sources, namely Brent, West Texas Intermediate, Dubai, Daqing and Shengli. The proposed model was applied to predict fluctuating crude oil prices. The result analysis showed that the smaller the fluctuation in prices was, the more accurate the prediction became whereas highly volatile prices led to poor prediction results. The authors of this study concluded that there existed a positive correlation between predicted and original data. In the same year, [66] used multi–layer perceptron neural networks with historical data of US gas prices collected from the EIAUSDE for a period from 1949 to 2010. Prediction results indicted better prediction accuracy was attained in the study.

Fernandez [67] adopted and applied the three models ARIMA, neural networks and support vector machine to predict crude oil and natural gas spot prices using historical data for a period from 1994 to 2005 sourced from DataStream. All models were subjected to crude oil and natural gas spot price prediction. Results of the experimental analysis showed that ARIMA outperformed the neural networks and support vector machine in short-term forecasting while neural networks and support vector machine performed better in long-term forecast accuracy. However, the hybridization of neural networks and support vector machine produced superior results to any of those achieved by the three individual models considered in this study.

Lai et al. [68] proposed Wavelet Decomposition Nonlinear Ensemble Value at Risk (WDNEVaR), a hybrid of wavelet analysis and neural networks to project value at risk in crude oil market prices. Historical data on West Texas Intermediate crude oil prices used in this research were collected from Global Financial Data ranging from April 1983 to June 2006. WDNEVaR was able to measure value at risk in crude oil market prices with higher accuracy than the predicted results produced by other models in comparison (ARMA–GARCH).

Haidar and Wolff [69] used a neural networks model to predict one-step and multi-step ahead prices of crude oil using crude oil future data and market data. Simulated results favored multi–step forecast over one-step. In addition, Pang et al. [70] used wavelet neural networks, linear relative inventory and nonlinear relative inventory models to predict one, two and three months ahead prices of crude oil based on data collected for the Organization of Economic Cooperation and Development inventory and West Texas Intermediate crude oil prices. Both data were obtained from the EIAUSDE covering a period from January 1992 to August 2006. The subsequent analysis of experimental results revealed that the wavelet neural networks model beat the linear relative inventory and nonlinear relative inventory models in terms of their prediction accuracy.

Movagharnejad et al. [71] used neural networks to forecast price differences in crude oil prices among five selected countries in the Persian Gulf region. Data for the study were collected from the American Petrochemical Institute and the EIAUSDE for a period from January 2000 to April 2010. Results proved that it was possible to accurately forecast varying crude oil prices in the Persian Gulf region.

Shambora and Rossiter [72] developed a crude oil trading system for generating buy and sell signals based on predictions produced by neural networks incorporated into the system. Historical data used in this study consisted of crude oil future contracts of NYMEX extracted from Datastream covering a period from April 1991 to December 2003. Among the selected trading strategies were buy and hold, technical analysis, naïve random walk and returns on treasury bills. The experimental results suggested that the neural networks model was superior to all other trading strategies in relation to yielding profits.

Raudys [73] proposed multi-layer perceptron to extract visual features from crude oil prices. The data used to build this forecasting model dated from November 1993 to January 2005 and were collected from West Texas Intermediate, Natural Gas–Henry, Fuel Oil No. 2 (NY), Gasoline, Unld Reg. Non–Oxy (NY) and the American Stock Exchange. The results indicated that the highest upsurges and declines in crude oil prices could be successfully and accurately predicted using multi-layer perceptron.

Panella et al. [74] considered three different models for the prediction of the dynamics in crude oil, natural gas and electricity prices in European and US markets. Sample data were collected from Europe (Brent crude oil) and the US (West Texas Intermediate crude oil) covering a period from 2001 to 2010. The models were radial basis function neural networks, adaptive neuro-fuzzy inference system networks and least-square approximation, the best prediction accuracy being achieved by adaptive neuro-fuzzy inference system networks.

Quek et al. [75] suggested novel recurrent neural networks to project prices of gold, crude oil and currency. Sample data were collected from Bloomberg and Datastream databases covering a period from 2000 to 2002. Multi-layer perceptron feed forward neural networks and Elman recurrent networks were also used to predict the prices of the mentioned commodities for the purpose of comparison. Novel recurrent neural networks was able to outperform all other models although all of them possessed the ability to capture data patterns and project future prices with good accuracy.

Gholamian et al. [76] designed a hybrid intelligent system composed of fuzzy rule base and neural networks. The data ranging from June 1998 to November 2000 and July 1988 to December 2000 were collected for experimental simulations from Sahand Naftiran. The analysis of experimental results showed an impressive projection accuracy of crude oil prices by the system.

Kaynar et al. [77] presented neural networks and neuro–fuzzy system to predict weekly gas consumption in Turkey based on historical data retrieved from Batas A.S for a period from January 2002 to April 2006. The study revealed a strong prediction ability of the adaptive neuro-fuzzy inference system compared to ARIMA, multi-layer perceptron and radial basis function networks but their MAPE was closer to each other indicating they were also good contenders for forecasting weekly gas consumption. In an earlier study executed by [78], crude oil prices were forecasted with fuzzy neural networks using crude oil time series extracted from NYMEX. Simulated results pointed to good forecast accuracy.

The support vector machine was proposed by Xie et al. [79] to forecast crude oil prices using monthly crude oil spot prices of West Texas Intermediate for a period from January 1970 to December 2003. ARIMA and back-propagation neural networks were also used to predict crude oil prices and evaluate the performance of the support vector machine. The accuracy of the forecasted results obtained from the experimental simulations placed the support vector machine ahead of ARIMA and back-propagation neural networks.

Zimberg [80] began his research work by conducting a series of experiments with Elman networks, feed-forward neural networks and the adaptive neuro–fuzzy inference system based on West Texas Intermediate and Brent crude oil prices covering the period from1991 to 2003. After preliminary experiments, the adaptive neuro–fuzzy inference system was selected as the best performing model and adopted for implementation in the study. The selected model and an econometric model were used to predict crude oil prices in which simulated outcomes of the sample test showed that the forecast accuracy of the adaptive neuro–fuzzy inference system was higher than that of the econometrics model.

Lacks et al. [81] aimed at forecasting crude oil prices in short, medium and long term periods using neural networks and historical data collected over a period from 1999 to 2006. The experiment result established that neural networks performed poorly in short term forecasting but could efficiently predict future prices of crude oil in middle and longer periods of time.

Qi and Zhang [82] applied a conventional support vector machine and improved cluster support vector machine for predicting crude oil prices using 1000 historical datasets. Both models were used to predict crude oil prices and the results revealed that the improved cluster support vector machine produced superior forecasted results.

Malliaris and G. Malliaris [83] targeted to forecast one-month-ahead spot prices of crude oil, heating oil, gasoline, natural gas and propane, their spot prices in market being interrelated. Spot price historical data of crude oil, heating oil, gasoline, natural gas and propane were extracted from Barchart (www.barchart.com) for a period from January 1994 to December 2002. The data of December 1997 to November 2002 were considered as experimental sample data for building the forecasted models. A multi-linear regression, a neural networks model and a simple model were applied in each of the energy markets and a one-month-ahead forecast was generated by each model for each market. The neural networks performed better in all markets except in the propane market.

Mingming et al. [84] proposed a wavelet Boltzmann cooperative neural networks and kernel density estimation (WBNNK) model to predict crude oil prices in relation to international gas prices. International crude oil and natural gas prices were collected for a period from 1976 to 2009 as sample data. The prediction task was fragmented into Botzman neural networks used to predict trend of oil prices and Gaussian kernel density used to predict randomization. The work analyzed the transform coefficient for international natural gas and crude oil prices and WBNNK was subsequently employed to predict crude oil prices, the experimental results suggesting good forecast accuracy.

Abdullah and Zeng [85] proposed an Artificial Neural Network–Quantitative (ANN-Q) model. Quantitative data were derived from online news and monthly crude oil prices were obtained from West Texas Intermediate for a period from January 1984 to February 2009. Two other models were also employed to evaluate the performance of the proposed model, namely Mining + Econometrics + Intelligence (intelligent algorithms) + Integration (TEI@I) nonlinear integration model and the Empirical Mode Decomposition Feed-Forward Neural Networks Adaptive Linear Neural Networks (EMD–FNN–ALNN) model. All three models were used to forecast future crude oil prices and their performances were compared. The results deduced from the following performance analysis showed that the TEI@I non-linear integration model generated the most accurate results followed by the ANN-Q and EMD–FNN–ALNN models.

Haidar et al. [86] used ARIMA, GARCH and artificial neural networks models in their study. Daily crude oil/return spot prices of West Texas Intermediate were collected for a period from January 1986 to March 2010 sourced from the EIAUSDE. A series of nonlinearity tests was performed on the observations and identified crude oil time series as being non-linear and therefore their concentration was duel on artificial neural networks because ARIMA and GARCH were not suitable tools. Crude oil prices were forecasted using artificial neural networks and the results showed a reasonably high forecast accuracy. However, the ARIMA and GARCH models possessed the possibility of forecasting crude oil prices provided that the noise in the data was smoothened.

Jinliang et al. [87] used a three step approach to predict international crude oil prices. Firstly, Botzman neural networks was used to predict prices of crude oil using historical data collected from 1975 to 2009. Secondly, random portions of the historical data were predicted using Gaussian kernel density estimation. Lastly, a hybrid of the Botzman neural networks and Gaussian kernel density estimation model was used to predict prices of crude oil which yielded high prediction accuracy as suggested by the experimental results.

Kaboudan [88] built genetic programming and neural networks forecasting models using monthly closing prices of crude oil covering a period from January 1993 to December 1998 and sourced from the EIAUSDE. The prediction accuracies of both models was then compared using naïve random walk fit forecasting accuracy and the results showed that the genetic programming model was superior to the multi-layer feed forward neural networks model.

Yu et al. [89] gathered historical data for crude oil prices from West Texas Intermediate and Brent spot prices of the EIAUSDE covering a period from January 2000 to March 2008. The generalized intelligent–agent–base fuzzy group model incorporated back-propagation neural networks, support vector machine regression and radial basic function neural networks. All individuals and hybrid models were used to forecast crude oil prices, the computation based on trapezoidal fuzzyfication. As result, the generalized intelligent–agent–base fuzzy group proved to be a reliable instrument for predicting crude oil prices which outperformed all other individual models.

Moshiri and Foroutan [90] obtained the historical data of crude oil prices required to build a prediction model from NYMEX covering a period from 1983 to 2003. ARIMA, GARCH and artificial neural networks models were used to predict crude oil prices. The forecasted results yielded by the models confirmed the superiority of the artificial neural networks model over ARIMA and GARCH in terms of forecast accuracy, especially in a chaotic time series.

Torben [91] compared the prediction performance of ARIMA, structural vector error correction and artificial neural networks regression. The data used to build forecast models were quarterly crude oil prices for FOB covering a period from the first quarter of 1986 to the fourth quarter of 2009. All three models were applied to forecast crude oil prices and their performances compared. The results indicated that the structural vector error correction performed the worst, ARIMA yielded good results and the artificial neural networks regression model produce the best results.

Shouyang et al. [92] proposed TEI@I methodology for predicting prices of crude oil using both quantitative and qualitative historical data to build a forecast model. This approach integrated several modules: a man–machine interface which provided an interaction point between the system and user; web–based text mining to extract information on uncertainty events influencing price of crude oil from the internet and served as a rule-based expert system which received the information from the web–based text mining module and determined its effect on crude oil prices based on stored knowledge; the econometric module which applied ARIMA to model the linear component of crude oil prices and artificial neural networks which captured a non-linear pattern in the time series of crude oil prices. The data used to build the ARIMA and artificial neural networks model were monthly spot prices of West Texas Intermediate covering the period from January 1970 to December 2003. The forecast results yielded by ARIMA and artificial neural networks, the effects of irregular events as determined by web–based text mining and a rule-based expert system, were fussed to produce an integrated forecast of crude oil prices. The forecast results produced by the artificial neural networks and ARIMA were then contrasted with the integrated forecast and the experimental analysis suggested that the integrated forecast results were superior. Khashman and Nwulu [93] designed a support vector machine model using weekly spot prices of West Texas Intermediate crude oil sourced from the EIAUSDE covering the period from January 1986 to December 2009. The model was applied to predict crude oil prices and a rate of 81.3% accuracy was achieved.

Xu et al. [94] proposed a rough set and wavelet neural networks (RSWNN) hybrid model for analyzing the factors that affected crude oil prices and predict future prices. Their approach consisted of text mining, rough set and wavelet neural networks. Text mining retrieved the necessary data on the factors affecting crude oil prices from the EIAUSDE, Reuters and Brent crude oil. Rough set further refined the document and extracted the main factors influencing crude oil prices while wavelet neural networks classified the factors according to their degree of significance. Historical data of these main factors were collected from 1970 to 2005 sourced from Brent crude oil. The simulation results showed that the model served as a good forecasting instrument for projecting oil prices.

A study of Chinese crude oil production and consumption was conducted by Ma and Wu [95] using production and consumption data from 1995 to 2008. A Grey forecasting model and Grey neural networks forecasting model were built and applied to predict national crude oil production and consumption from 2009 up to 2015. The subsequent comparison of the forecast results generated by both models indicated that Grey neural networks possessed a superior ability to project crude oil production and consumption.

Khashman and Nwulu [96] conducted a comparative analysis of the support vector machine and back-propagation for the prediction of crude oil prices. Both models were build based on West Texas Intermediate crude oil prices data collected from the EIAUSDE covering the period from January 1986 to December 2009. The prediction results generated by each model were compared, the back-propagation model proving superior to the support vector machine. Phichhang and Wang [97] proposed a hybrid model of neural networks and GARCH to forecast the crude oil price volatility in Chinese and US crude oil markets. US spot price FOB and Chinese Daqin spot price FOB were collected from the EIAUSDE, the sample data covering the period from 1997 to 2010. The comparative analysis of hybrid and GARCH indicated a definite superiority of the hybrid model.

Yu et al. [98] proposed a knowledge-based forecasting system integrating text mining and rough set. Text mining was responsible for searching the internet and the internal file system to collect both regular factors and uncertainty events influencing crude oil prices, creating a metadata repository and generating pattern and rules. Rough set further refined the pattern and rules generated by text mining and predicted future crude oil prices. ARIMA, linear regression model, auto–regression integration moving average and back-propagation neural networks were used to evaluate the effectiveness of the proposed technique. The data required for building the comparative models were collected from the EIAUSDE and West Texas Intermediate covering the period from January 1970 to October 2004. The prediction results generated by all models were then compared and the results proved the superiority of the proposed model to the other techniques.

6. Analysis of Various CI techniques for oil price prediction

Tab. II offers an overview of research works reviewed in this paper, the first column containing different CI models proposed by several researchers, the majority proposing hybridized CI techniques to predict future crude oil prices.

All methods proposed by the authors in their separate studies superseded chosen techniques for comparison as established in each work, except in the study conducted by [85] in which the performance of the proposed model proved inferior. Hybridized CI models generally generated superior results to individual models. The second column lists the sources of the experimental data used to build the respective forecasting models. Most research was based on data of West Texas Intermediate and Brent crude oil prices representing US and UK crude oil markets, mostly extracted from the EIAUSDA. Only a few studies sourced their experimental data elsewhere or did not disclose their sources at all. The data gathered were extracted at different intervals (daily, weekly, monthly and quarterly) depending on the respective research objective. For example, [60] collected on daily bases, [93, 96] weekly, [1] monthly and [91] quarterly. The third column specifies the coverage period of the sources data, the most extensive being [42] followed by [66]. Ma [48] limited their data period to that of only five months which constitutes the shortest period of data used in all included studies. The fourth column indicates the trend of publications. The highest number of ten publications was recorded in 2008 and lowest of one publication in 2004. Overall, the trend shows an increasing number of publications with 44 of 61 research papers published within the last five years. The last column lists the names of the respective contributors.

Proposed Model	**Source of Data**	**Covered period**	**Year**	**Authors**
Neural networks	American Petrochemical Institute and EIAUSDA	January, 2000 to April, 2010	2011	Movagharnejad et al.
Adaptive neuro fuzzy inference system	Batas A.S	January, 2002 to April, 2006	2011	Kaynar et al.
Novel recurrent neural networks	Bloomberg and datastream	2000 to 2002	2008	Quek et al.
Adaptive neurofuzzy inference system networks	EIAUSDA	2001 to 2010	2011	Panella et al.
Wavelet recurrent neural network	Undisclosed by the authors	1945 to 2010	2012	Mingming and Jinliang
Fuzzy neural networks	EIAUSDA	May 20, 1987 to August 30, 2006	2007	Liu et al.
Jump stochastic time effective function neural networks	Brent, West Texas Intermediate, Dubai, Daqing and Shengli	Not specified by authors	2012	Wang et al.
Cascaded neural networks	Bureau of Labor Statistics & Bureau of Economic Analysis	January, 1947 to December, 2004.	2006	Malik and Nasereddin
Hybridization of neural networks and support vector machine	DataStream	1994 to 2005	2006	Fernandez

Neural networks	DataStream	April 16, 1991 to December 31, 2003	2007	Shambora and Rossiter
ARIMA, structural vector error correction and artificial neural networks regression	FOB	First quarter of 1986 to fourth quarter of 2009	2010	Torben
WDNEVaR	Global Financial Data	April 4, 1983 to Jun 30, 2006	2007	Lai et al.
Fuzzy – neural networks	Mediteranean sidi kerir Iran light spot price FOB	January, 1989 up to January, 2009	2009	Mehdi
Symbol evolutionary Immune clustering neural networks	New York harbor residual	November 29, 2006 to 24 of April, 2007	2009	Ma
TEI@I	EIAUSDA	January, 1970 to December, 2003	2005	Shouyang et al.
Support vector machine	Undisclosed by the authors	January, 1970 to December, 2003	2006	Xie et al.
Neural networks	Undisclosed by the authors	1999 to 2006	2007	Lacks et al.
Neural networks	http://finance.yahoo.com	Not specified by authors	2008	Haidar et al.
Adaptive neuro – fuzzy inference system	EIAUSDA	1991 to 2003	2008	Zimberg
Improved cluster support vector machine	Undisclosed by the authors	Not specified	2009	Qi and Zhang
Hybrid of Bozeman neural networks and Gaussian kernel density estimation model	Undisclosed by the authors	1975 to 2009	2009	Jinliang et al.
Generalized regression neural networks	Undisclosed by the authors	20 months	2010	Alizadeh and Mafinezhad
Particle swamp optimization and genetic algorithms	Energy Balance Annual Report. Tehran: Ministry of Energy, 2005	1981 to 2005	2010	Assareh et al.
WBNNK	Undisclosed by the authors	1976 to 2009	2009	Mingming et al.
Grey neural networks	Undisclosed by the authors	1995 to 2008	2010	Ma and Wu
Back- propagation neural networks	Undisclosed by the authors	1996 to 2010	2011	Chen and Qu
Fuzzy neural networks	Undisclosed by the authors	Not specified by authors	2001	Rast

Artificial neural networks	EIAUSDA	1983 to 2003	2006	Moshiri and Foroutan
Neural networks	Undisclosed by the authors	Not specified by authors	2010	Wang and Yang
ANN- Q	Online news & West Texas Intermediate	January, 1984 to February, 2009	2010	Abdullah and Zeng
Abductive and neural networks	Professor A.M El – Sharkawi of the University of Washington, USA	January 1, 1985 to October 12, 1992	2008	Abdel – Aal
Hybrid intelligent system	Sahand Naftiran Company	June 2, 1998 to November 30, 2000	2005	Gholamian et al.
Redial basis function neural networks	EIAUSDA	January, 1997 to October, 2008	2009	Qunli et al.
Genetic programming and neural networks	EIAUSDA	January, 1993 to December, 1998	2001	Kaboudan
Wavelet and least square support vector	EIAUSDA	January, 1991 to July, 2007 & May, 1987 to July, 2007	2007	Bao et al.
EMDNNELP	EIAUSDA	January 1, 1986 to September 30, 2006	2008	Yu et al .
Wavelet neural network	EIAUSDA	January, 1986 to October, 2007	2008	Alexandridis and Livanis
Generalized intelligent – agent – base fuzzy group	EIAUSDA	January 1, 2000 to March 31, 2008	2008	Yu et al.
Multilayer feed forward neural networks	EIAUSDA	September, 1996 to August, 2007	2009	Kulkarni and Haidar
Fuzzy time series	EIAUSDA	January 2, 1991 to December 31, 2009	2010	Zhang et al.
hybrid of neural networks and GARCH	EIAUSDA	1997 to 2010	2010	Phichhang and Wang
Group method data handling neural networks	EIAUSDA	January 1, 2002 to July 13, 2009	2011	Mehrara et al.

Wavelet neural networks	EIAUSDA	January, 1992 to August, 2006	2011	Pang et al.
Artificial neural networks	EIAUSDA	January 2, 1986 to March 2, 2010	2011	Haidar and Wolff
Back-propagation	EIAUSDA	January, 03 1986 to December 25, 2009	2011	Khashman and Nwulu
Neural networks – fuzzy regression algorithms	EIAUSDA	1985 to 2007	2012	Azadeh et al.
Neuro – genetic	EIAUSDA	January 1, 2004 to December 31, 2008	2012	Abrishami and Varahrami
HTW – MBPNN	EIAUSDA	January, 1988 to March, 2010	2012	Jammazi and Aloui
Multi – layer perceptron neural networks	EIAUSDA	1949 to 2010	2012	Sotoudeh and Farshad
knowledge base forecasting system	EIAUSDA	January, 1970 to October, 2004	2005	Yu et al.
RSWNN	EIAUSDA, Reuters & Brent.	1970 to 2005	2007	Xu et al
Hybrid AI system	EIAUSDA	January, 1970 to December, 2002	2004	Wang et al
Neuro – genetic	EIAUSDA	January, 1983 to December, 2006	2007	Amin-Naseri and Gharacheh
Adaptive network base fuzzy inference system	Undisclosed by authors	January 5, 2004 to April 30^{th}, 2007	2009	Ghaffari and Zare
Genetic algorithms – neural networks	EIAUSDA	1994 to 2008	2011	Tehrani and Khodayar
Generalize pattern matching based on genetic algorithms	EIAUSDA	March 20, 1987 to July 26, 2005 and January 2, 1986 to July 26, 2005	2008	Fan et al.
Multi – layer feed forward neural networks	EIAUSDA	January, 1996 to August, 2007	2009	Pan et al .

AI agent based fuzzy	EIAUSDA	January 1, 2000 to December 31, 2007	2008	Yu et al.
WDNEVaR	GFD	April 4, 1983 to Jun 30, 2006, May 20, 1987 to Jun 30, 2006 & January 20, 1997 to January 30, 2005	2009	He et al.
Multilayer perceptron	West Texas Intermediate, Natural Gas – Henry, Fuel oil No. 2 (NY), Gasoline, Unld Reg. Non – oxy (NY) and American Stock Exchange	November 1, 1993 to January 12, 2005	2005	Raudys
Neural networks	www.barchart.com	December, 1997 to November, 2002	2008	Malliaris and G. Malliaris

Tab. II *Summary of reviewed literature.*

Tab. III summarizes the studies that used hybrid CI technique for the projection of crude oil prices and compared the results with single CI techniques. All 18 studies proposed a hybrid model, compared it with single CI techniques and demonstrated that the hybrid technique generated better results than single techniques in terms of their prediction accuracy. Therefore, it can be deduced that a hybrid CI technique is preferable for the projection of crude oil prices.

References	Results
1, 39, 43, 44, 46, 48, 49, 52, 55, 57, 59, 67, 74, 77, 80, 89, 92, 98	Hybrid models are superior over individual CI techniques

Tab. III *Comparing performances of hybrid and single CI technique.*

7. Conclusions

This paper has presented a concise review of the CI techniques applied in projecting crude oil prices. Hybridized models proposed in the discussed studies demonstrated their superiority to singular models which constitutes a trend which may likely continue in the near future. Most of the experimental data were extracted from the EIAUDA and the most patronized oil markets of West Texas Intermediate and Brent. This is probably due to the fact that their influence in the global oil markets is most pronounced. Most of the literature surveyed applied a high data volume for building their CI prediction models. It is evident that this particular area of research is rapidly developing because 44 out of 58 research papers have been published within the last five years. This concise survey is hoped to provide a better insight and understanding of such prediction models, range and source of data that have been used in several literature for future researchers.

8. New research direction

The discussed authors concentrated on building an efficient CI model for predicting crude oil prices using regular market factors which generally affect crude oil prices. However, none of the authors applied artificial neural networks despite its demonstrated success in building a model that can also capture the effects of uncertainty events such as armed conflict, political unrest and economic instability. This paper aims at suggesting a new research direction for researchers across the globe to channel their abilities and resources towards creating the optimum neural networks model that is able to provide a promising and reliable prediction model inculcating both regular and uncertainty factors. More research in this direction is still lacking but highly recommendable considering the volatile nature of the Gulf region which produces significant amount of the world crude oil and remains under constant threat from terrorist attacks, natural disasters, extreme weather conditions, oil company mergers and strikes. The authors of this research is presently working on such an artificial neural networks model that has the ability to capture the effects of both regular and uncertainty factors on crude oil prices and generate reliable and more accurate crude oil price projections than can meet practical needs than existing models.

References

[1] Jammazi R., Aloui C.: Crude oil forecasting: experimental evidence from wavelet decomposition and neural network modeling. Energ. Econ., **34**, 2012, pp. 828–841.

[2] He K., Yu L., Lai K. K. Crude oil price analysis and forecasting using wavelet decomposed ensemble Model. Energy **46**, 2012, pp. 564–574.

[3] Chen A., Leung M. T., Wang L.: Application of polynomial projection ensembles to hedging crude oil commodity risk. Expert Syst. Appl., **39**, 2012, pp. 7864–7873.

[4] Kaufmann R. K.: The role of market fundamentals and speculation in recent price changes for crude oil. Energ. Policy, **39**, 2011, pp. 105–115.

[5] Jinliang Z., Mingming T., Mingming T.: Effect simulation of international gold prices on crude oil prices. In: Proc. Int. Colloquium Comput. Commun. Control Manage., Wuhan, China 2009, pp. 459–463.

[6] Chevillon G., Rifflart C.: Physical market determinants of the price of crude oil and the market premium. Energ. Econ., **31**, 2009, pp. 537–549.

[7] Alizadeh A., Mafinezhad K.: Monthly brent oil price forecasting using artificial neural networks and a crisis index. In: Proc. Int. Conf. Electron. Inf. Eng., Kyoto, Japan, 2010, pp. 465–468.

[8] Rafiq S., Salim R., Bloch H.: Impact of crude oil price volatility on economic activities: An empirical investigation in the Thai economy. Resour. Policy, **34,** 2009, pp. 121–132.

[9] Khoukhi A., Oloso M., Elshafei M., Abdulrahefm A., Al-Majed A.: Support vector regression and functional networks for viscocity and gass/oil ratio curves estimation. Int. J. Comput. Intel. Appl., **3**, 2011, pp. 269–293.

[10] Editorial.: Hybrid learning machine. Neurocomputing, **72**, 2009, pp. 2729–2730.

[11] Deng W., Chen R., He B., Liu Y., Yin L., Guo J.: A novel two-stage hybrid swarm intelligence optimization algorithm and application. Soft Comput., **6**, 2012, pp. 1707–1722.

[12] Fazel M. H. Z., Gamasaee R.: Type-2 fuzzy hybrid expert system for prediction of tardiness in scheduling of steel continuous casting process. Soft Comput., **16**, 2012, pp. 1287–1302.

[13] Jia D., Zhang G., Qu B., Khurram M.: A hybrid particle swarm optimization algorithm for high-dimensional problems. Comp. Ind. Eng., **61**, 2011, pp. 1117–1122.

[14] Khashei M., Bijari M., Reza S. H.: Combining seasonal ARIMA models with computational intelligence techniques for time series forecasting. Soft Comput., **16**, 2012, pp. 1091–1105.

[15] Guvenc U., Duman S., Saracoglu B., Ozturk A.: A hybrid GA–PSO approach based on similarity for various types of economic dispatch problems. Electron Electr. Eng. Kaunas: Techn., **2(108)**, 2011, pp. 109–114.

[16] Acampora G., Lee C., Vitiello A., Wang M.: Evaluating cardiac health through semantic soft computing techniques. Soft Comput., **16**, 2012, pp. 1165–1181.

[17] Patruikic M., Milan R. R., Jakovljevic B., Dapic V.: Electric energy forecasting in crude oil processing using support vector machines and particle swamp optimization. 9$^{\text{th}}$ Symposium on Neural Network Application in Electrical Engineering, Faculty of Electrical Engineering, University of Belgrade, Serbia, 2008.

[18] Saybani M. R., Wah Y. T., Amini A., Yazdi S. R. A. S., Lahsasna A.: Applications of support vector machines in oil refineries: A survey. Int. J. Physical Sci., **6(27)**, 2011, pp. 6295–6302.

[19] Soldo B.: Forecasting natural gas consumption. Appl. Energ., **92**, 2012, pp. 26–37.

[20] Ji Q.: System analysis approach for the identification of factors driving crude oil prices. Comp. Ind. Eng., **63**, 2012, pp. 615–625.

[21] Yi Y., Qin N.: Oil price forecasting based on self – organizing data mining. In: Proc. of IEEE Int. Conf. Grey Syst. Intell. Services, Nanjig, China, 2009, pp. 1386–139.

[22] Zhang X., Yu L., Wang S., Lai K. K.: Estimating the impact of extreme events on crude oil prices: An EMD – based event analysis method. Energ. Econ., **31**, 2009, pp. 768–778.

[23] Ortiz – Cruz A., Rodriguez E., Ibarra-Valdez C., Alvarez-Ramirez J.: Efficiency of crude oil markets: evidence from informational entropy analysis. Energ. Policy, **41**, 2012, pp. 365–373.

[24] Zhang X., Lai K.K., Wang S.: A new approach for crude oil price analysis based on empirical mode decomposition. Energ. Econ., **30**, 2008, pp. 905–918.

[25] McCulloch W. S., Pitts W.: A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys., **5**, 1943, pp. 115–133. Reprinted in Anderson and Rosenfeld, 1988.

[26] Fausett L.V.: Fundamentals of Neural Networks. Architectures, Algorithms, and Applications. Prentice-Hall: Englewood Cliffs, 1994.

[27] Mitchell M.: An introduction to genetic algorithms, MIT Press: Massachusetts, 1996.

[28] Barschi M.W.: A genetic algorithms for resource – constrained scheduling. PhD thesis submitted to mechanical engineering, Massachusetts Institute of technology, 1996.

[29] Grefenstette J. J.: Optimization of control parameters for genetic algorithms. IEEE T. Syst. Man Cy., **16(1)**, 1986, pp. 122–128.

[30] Vose M. D., Liepins G.E.: Punctuated equilibrium in genetic search. Complex. Syst., **5**, 1991, pp. 31–44.

[31] Koza J. R.: Genetic programming, on the programming of computers by means of natural selection. MIT: Cambridge, 1992.

[32] Eberhart R., Kenned J.: A New Optimizer Using Particle Swarm Theory. In: Micro Machine and Human Science, Proc. of the 6th Int. Symposium on, Nagoya, Japan, 1995, pp. 39–43.

[33] Chen G., Tat T. P.: Introduction to fuzzy sets, fuzzy logic and fuzzy control system. CRC Press: Boca Raton, 2001.

[34] Pawlak Z.: Rough set approach to knowledge based decision support. Eur. J. Oper. Res., **99**, 1997, pp. 48–57.

[35] Krishnamoorth C.S., Rajeev S.: Artificial Intelligence and expert system for engineers. CRC Press: LCC, 1996.

[36] Merritt D.: building expert system in prolog. Springer: New York, 1989.

[37] Abraham A.: Intelligent systems: architectures and perspectives, recent advances in intelligent paradigms and applications. In: Abraham A, Jain L, Kacprzyk J, editors. Studies in fuzziness and soft computing. Springer: Germany, 2002, pp. 1–35.

[38] Bahrammirzaee A. A.: comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. Neural Comput. Appl., **19**, 2010, pp. 1165–1195.

[39] Fan Y., Liang Q., Wei Y.: A generalized pattern matching approach for multi – step prediction of crude oil price. Energ. Econ., **30**, 2008, pp. 889–904.

[40] Ghaffari A., Zare S.: A novel algorithms for prediction of crude oil price variation based on soft computing. Energ. Econ., **31**, 2009, pp. 531–536.

[41] Amin-Naseri M. R., Gharacheh E. A.: A hybrid artificial intelligence approach to monthly forecasting of crude oil prices time series. In: The Proc. of the 10^{th} Int. Conf. Eng. Appl. Neural Networ., CEUR-WS284, 2007, pp. 160–167.

[42] Mingming T., Jinliang Z.: A multiple adaptive wavelet recurrent neural network model to analyze crude oil prices. J. Econ. Bus., **64**, 2012, pp. 275–286.

[43] Yu L., Wang S., Wen B., Lai K. K.: Forecasting crude oil price with an EMD- based neural networks ensemble learning paradigm. Energ. Econ., **30**, 2008, pp. 2623–2635.

[44] Malik F., Nasereddin M.: Forecasting output using oil prices: A cascaded artificial neural network approach. J. Econ. Bus., **58**, 2006, pp. 168–180.

[45] Azadeh A., Moghaddam M., Khakzad M., Ebraimipour V.: A flexible neural network-fuzzy mathematical programming algorithm for improvement of oil price estimation and forecasting. Comp. Ind. Eng., **62**, 2012, pp. 421–430.

[46] Bao Y., Zhang X., Yu L., Lai K. K., Wang S.: Hybridizing wavelet and least squares support vector machines for crude oil price forecasting. School of management and economics of UESTC web. http://www.mgmt.uestc.edu.cn/prc/papers/IWIF-II%20Wang%20Shouyang%202.pdf. 2007. Accessed 12 June 2012.

[47] Alexandridis A., Livanis E.: Forecasting crude oil prices using wavelet neural networks. University of Macidonia web. http://users.uom.gr/∼aalex/files/Download/FSDET%202008%20.pdf. 2008. Accessed 10 July 2012.

[48] Ma X.: Fuel oil price forecasting using symbiotic evolutionary immune clustering neural network. In: Proc. of 2^{nd} Int. Conf. Intell. Comput. Technol. Aut., Qingdao, China, 2009, pp. 322– 325.

[49] Tehrani R., Khodayar F.: A hybrid optimization artificial intelligent model to forecast crude oil using genetic algorithms. Afr. J. Bus. Manag., **5(34)**, 2011, pp. 13130–13135.

[50] Abrishami H., Varahrami V.: Survey of a rule based expert system for gas price forecasting. World Appl. Sci. J., **16**, 2012, 4, pp. 493–500.

[51] Liu J., Bai Y., Li B.: A new approach to forecast crude oil price based on fuzzy neural Network In: Proc. IEEE Comput. Society Fourth Int. Conf. on Fuzzy Sys. Knowl. Discovery art., Haikou, China, 2007, pp. 273–277.

[52] Mehrara M., Moeini A., Ahrari M.: Comparison of two approach of neural network for forecasting of oil future prices. Asia J. Bus. Manage. Sci., **1(3)**, 2011, pp. 54–64.

[53] Pan H., Haidar I., Kulkarni S.: Daily prediction of short-term trend of crude oil prices using neural networks exploiting multimarket dynamics. Front. Comput. Sci. Chi., **3(2)**, 2009, pp. 177–191.

[54] Assareh E., Behrang M. A., Assari M. R., Ghanbarzadeh A.: Application of particle swarm optimization and genetic algorithms techniques on demand estimation of oil in Iran. Energ., **35**, 2010, pp. 5223–5229.

[55] Abdel-Aal R. E.: Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks. Comp. Ind. Eng., **54**, 2008, pp. 903–917.

[56] Wang T.,Yang J.: Nonlinearity and intraday efficiency test on energy future markets. Energ. Econ., **32**, 2010, pp. 496–503.

[57] Yu L., Wang S., Lai K. K.: A generalized intelligent – agent – based fuzzy group forecasting model for oil price prediction. In: Proc. of IEEE Int. Conf. Syst., man Cy., Singapore, 2008, pp. 489–493.

[58] Mehdi M.Z.: Forecasting momentary price of crude oil and gas fossile energies in the world markets through fuzzy based modeling. International gas union web. http://www.igu.org/html/wgc2009/papers/docs/wgcFinal00056.pdf. 2009. Accessed 20 May 2012.

[59] Wang S. Y., Yu L., Lai K. K.: A novel hybrid AI system framework for crude oil price forecasting. Lect. Notes Comput. Sci., **3327**, 2004, pp. 233–242.

[60] He K., Chi X., Chen S., Lai K. K.: Estimating VaR in crude oil market: A novel multi – scale non-linear ensemble approach incorporating wavelet analysis and neural network. Neurocomputing, **72**, 2009, pp. 3428–3438.

[61] Zhang X., Wu Q., Zhang J.: Crude oil price forecasting using fuzzy time series. In: Proc. Of IEEE 3rd Int. Symp. Knowl. Acquis. Model, Wuhan, China, 2010, pp. 213–216.

[62] Kulkar S., Haidar I.: Forecasting model for crude oil price using artificial neural networks and commodity future prices. Int. J. Comp. Sci. Inf. Secur., **2(1)**, 2009, pp. 81–88.

[63] Chen X., Qu Y.: A prediction method of crude oil output based on artificial neural networks. In: Proc. of the IEEE Int. Conf. Comput. Inf. Sci., pp. 702–704.

[64] Qunli W., Ge H., Xiaodong C.: Crude oil forecasting with an improved model based on wavelet transform and RBF neural network. IEEE Int. Forum Inf. Technol. Appl., Chengdu, China, 2009, pp. 231–234.

[65] Wang J., Huopo P., Fajiang L.: Forecasting crude oil price and stock price by jump stochastic time effective neural network model. J. Appl. Math. 2012, Article ID 646475, pp. 1–15.

[66] Sotoudeh M., Farshad E.: Application of neural network for forecasting gas price in America. J. Math. Comput. Sci., **4(2)**, 2012, pp. 216–226.

[67] Fernandez V.: Forecasting crude oil and natural gas spot prices by classification methods. IDEAS web. http://www.webmanager.cl/prontus_cea/cea_2006/site/asocfile/ASOCFILE1200611281 05820.pdf. 2006. Access 4 June 2012.

[68] Lai K. K., He K., Yen J.: Modeling VaR in crude oil market: A multi scale nonlinear ensemble approach incorporating wavelet analysis and ANN. Lect. Notes Comput. Sci., **4487**, 2007, pp. 554–561.

[69] Haidar I., Wolff C. R.: Forecasting crude oil price (revisted). United State association of energy economics web. http://www.usaee.org/usaee2011/submissions/OnlineProceedings/Forecasting%20Crude%20Oil%20Price%20(Revisited).pdf. 2011. Accessed 10 June, 2012.

[70] Pang Y., Xu W., Yu L., Ma J., Lai K. K., Wang S., Xu S.: Forecasting crude oil spot price by wavelet neural networks using OECD petroleum inventory levels. New Math. Nat. Comput., **2**, 2011, pp. 281–297.

[71] Movagharnejad K., Mehdizadeh B., Banihashemi M., Sheikhi M. K.: Forecasting the differences between various commercial oil prices in the Persian Gulf region by neural networks. Energ., **36**, 2011, pp. 3979–3984.

[72] Shambora W.E., Rossiter R.: Are there exploitable inefficiencies in the futures market for oil?. Energ. Econ., **29**, 2007, pp. 18–27.

[73] Raudys A.: Accuracy of MLP based data visualization used in oil prices forecasting task. Lect. Notes Comput. Sci., **3617**, 2005, pp. 761–769.

[74] Panella M., Barcellona F., D'Ecclesia R.L.: Forecasting Energy Commodity Prices Using Neural Networks. Adv. Decis. Sci., Hindawi, 2012, Article ID 289810, 26 pages, DOI:10.1155/2012/289810.

[75] Quek C., Pasquier M., Kumar N.: A novel recurrent neural network-based prediction system for option trading and hedging. Appl. Intell., **29**, 2008, pp. 138–151.

[76] Gholamian M. R., Ghomi S. M. T. F., Ghazanfari M.: A hybrid systematic design for multiobjective market problems: a case study in crude oil markets. Eng. Appl. Artif. Intell., **18**, 2005, pp. 495–509.

[77] Kaynar O., Yilmaz I., Demirkoparan F.: Forecasting of natural gas consumption with neural networks and neuro fuzzy system. Energ. Educ. Sci. Tech., **26(2)**, (2011), pp. 221–238.

[78] Rast M.: Fuzzy neural networks for modeling commodity markets. In: Proc. Of IFSA World Congress and 20th NAFIPS Int. Conf., **2**, 2001, pp. 952–955.

[79] Xie W., Yu L., Xu S., Wang, S.: A new method for crude oil price forecasting based on support vector machines. Lect. Notes Comput. Sci., **3994**, 2006, pp. 441–451.

[80] Zimberg B.: Crude oil price forecasting with ANFIS. University of Vaasa web. http://lipas.uwasa.fi/∼phelo/ICIL2008TelAviv/24.pdf. 2008. Accessed 3 July 2012.

[81] Lacks R., Borgermann C., Dirkmorfeld M.: Forecasting the price development of crude oil with artificial neural networks. In: Proc. Int. Work – Conf. Artif. Neural Networ. Salamanca, 2009, pp. 248 – 255.

[82] Qi Y., Zhang W.: The improved SVM for forecasting the fluctuation of international crude oil prices. Proc. of IEEE Int. Conf. Electron Commer. Bus. Intell., Beijing, China, 2009, pp. 269–271.

[83] Malliaris M. E., Malliaris S. G.: Forecasting energy product prices. Eur. J. Financ., **14(6)**, 2008, pp. 453–468.

[84] Mingming T., Jinliang Z., Mingxin T.: Effect simulation of international natural gas prices on crude oil prices based on WBNNK model. IEEE Comput. Society 6^{th} Int. Conf. Natural Comput., Sanya, China, 2010, pp. 1643–1648.

[85] Abdullah S. N., Zeng X.: Machine learning approach for crude oil price prediction with artificial neural networks-Quantitative (ANN-Q) model. In: Proc. of the Int. Joint Conf. Neural Networ., Barcelona, Spain, 2010, pp. 1–8.

[86] Haidar I., Kulkarni S., Pan H.: Forecasting model for crude oil prices based on artificial neural networks. Symposium on Machine Learning and Applications (MLA 2008). In: Proc. 4^{th} Int. Conf. Intell. Sensors, Sensor Networks Inf. Process., Sydney, China, 2008, pp. 103–108.

[87] Jinliang Z., Mingming T., Tao M.: Based on wavelet-Boltzmann neural network and kernel density estimation model predict international crude oil prices. In: Proc. IEEE Int. Conf. future Comput. Commun, Wuhan, China, 2009, pp. 150–153.

[88] Kaboudan M. A.: Compumetric forecasting of crude oil prices. In: Proc. IEEE Congress Evol. Comput., Seoul, Korea, 2001, pp. 283–287.

[89] Yu L., Wang S., Wen B.: An AI-agent trapezoidal fuzzy ensemble forecasting model for crude oil prediction. In: Proc. of the IEEE 3rd Int. Conf. Innovative Comput. Inf. Control, Dalian, China, 2008, pp. 327.

[90] Moshiri S., Foroutan F.: Forecasting nonlinear crude oil futures prices. Energ. J., **27**, 2006, pp. 81–95.

[91] Torban F. B.: Models for oil price prediction and forecasting . Msc Thesis, San Diego State University, 2010.

[92] Shouyang W., Lean Y., Lai K. K.: crude oil price forecasting with tei@i methodology. J. Syst. Sci. Complex, **18(2)**, 2005, pp. 145–166.

[93] Khashman A., Nwulu I. N.: Intelligent prediction of crude oil using support vector machines. In: 9^{th} IEEE Int Symposium Appll. Mach. Intell. Inform., Smolenice, Germany, 2011, pp. 165–169.

[94] Xu W., Wang J., Zhang X., Zhang W., Wang S.: A new hybrid approach for analysis of factors affecting crude oil price. Lect. Notes Comput. Sci., **4489**, 2007, pp. 964–971.

[95] Ma H., Wu Y.: A grey forecasting model based on BP neural network for crude oil production and consumption in China. In: Proc. IEEE Comput. Society 3th Int. Symposium Inf., Qingdao, China, 2010, pp. 105–109.

[96] Khashman A., Nwulu I. N.: support vector machine versus back propagation algorithms for oil prediction. Lect. Notes Comput. Sci., **6677**, 2011, pp. 530–538.

[97] Phichhang O. U., Wang H.: Application of neural networks in modeling and forecasting volatility of crude oil markets: Evidence from US and China. Adv. Mater. Res., **230-232**, 2011, pp. 953–957.

[98] Yu L., Wang S., Lai K. K.: A rough set refined text mining approach for crude oil market tendency forecasting. Int. J. Knowl. Syst. Sci., **2(1)**, 2005, pp. 33–46.