An Intelligent Model Framework for Handling Effects of Uncertainty Events for Crude Oil Price Projection: Conceptual Paper

Haruna Chiroma *Member, IAENG*, Sameem Abdulkareem, and Abdulsalam Ya'u Gital

***Abstract*— This proposal is a novel intelligent model framework that can learn new patterns from new datasets which might have been distorted by uncertainty events. During the first phase of the research, historical data of Brent crude oil prices were collected from the Energy Information Administration of the US Department of Energy. The data were cleaned and normalized. The second phase involves genetic optimization of neural network to build an intelligent model using the preprocessed data. The intelligent model will be periodically retrained with distorted data, such as data from the 1991 Gulf War, the 1997 Asian financial crisis, the 2002 Venezuelan unrest, the second Gulf War of 2003, the 2001 US twin tower attack, and the 2007 global financial recession. Retraining could allow the model to learn and capture new patterns on the basis of the distorted data to predict crude oil prices during uncertainty events. This novel approach to crude oil price prediction is expected to produce more accurate results than the results discussed in existing literature, and subsequently, provide more realistic Predicted prices of crude oil for proper planning by governments, intergovernmental organizations such as the Organization of Petroleum Exporting Countries (OPEC), and private businesses. This in turn will help to avoid the negative effects of crude oil price volatility.**

***Index Terms*—Genetic Algorithm, Neural Network, Uncertainty Events, Crude Oil Price**

I. Introduction

GLOBALIZATION implies that interdependent products and their prices will fluctuate when there is a change in product. Crude oil is perhaps the commodity that exhibits such characteristics more than any other commodity. Distortion in oil markets has serious effects on other goods and services. Crude oil is perhaps the most significant commodity traded across the globe, although there is no consensus about that. Almost all sectors of the world economy depend on crude oil. Therefore, any fluctuation in crude oil prices will have a multiplier effect on the global economy [1]. Many products use oil derivatives (e.g., rubber, medicines) or depend on energy sources that use oil derivatives for either their manufacture or operation (e.g., motors).

Instability of crude oil prices, partly triggered by political crises in oil producing countries, has a direct effect on almost all sectors in global capital markets. Hence, other markets are also affected if there is a crisis in oil markets [2]. Crude oil production may be limited by unexpected events such as war or revolution in the Middle East, which are generally viewed as exogenous (external factors) with respect to world macroeconomic conditions. Hamilton [3] describes the first historical shock on crude oil markets that occurred in 1862–1864 as a result of the US Civil War, where crude oil prices were destabilized. Kilian [4] discusses that the effects of the decline in oil production due to unexpected events make crude oil a unique commodity.

Regular short term crude oil price movements are caused by normal market forces such as US refinery capacity, Organization of Petroleum Exporting Countries (OPEC) crude oil production ceilings, and global demand and supply; on the other hand, volatility in the oil market is prompted by unexpected events such as war, revolution, earthquake, oil workers' strike, and hostage taking [5]. Because of the chaotic nature of the crude oil market, several computational intelligent models have been proposed in the literature to forecast prices of crude oil in order to counter the negative impact. Several researchers and institutions perceive crude oil volatility as a source of great concern, making crude oil price prediction a key issue, but a very difficult one to solve [6] because of the interactive effects of several factors, nonlinearity, and the complex, dynamic nature of crude oil prices in general [7].

Accurate projection of crude oil prices is very important to governments, private institutions, and businesses in strategic decision making. Therefore, researchers in the computational intelligence community have proposed several models for crude oil price projection. Most previous studies focus primarily on applying computational intelligent techniques to build effective models for forecasting crude oil prices using historical data of standard economic laws (demand/supply) governing crude oil trading. Very few studies have developed models that consider both regular and unexpected events. [8 – 9, 6] considered both regular factors and the effects of uncertainty events and applied expert systems to manage the effects of unexpected events.

Sotoudeh and Farshad [10] claimed that forming accurate market information is a difficult task because human experts

H. C. is a PhD research scholar with the University of Malaya, Kuala Lumpur, Malaysia and Federal College of Education (Technical) Gombe, Nigeria. (corresponding author: +60149576297, +2348035955858; e-mail: freedonchi@yahoo.com).

S. A. She is an Assoc. Professor with the University of Malaya, Department of Artificial Inteligence, Kuala Lumpur, Malaysia. (e-mail: sameem@um.edu.my).

A. Y. G. is a PhD research scholar with the Computer Science Department, University of Technologi Malaysia, Johor Baru, Malaysia and Abubakar Tafawa Balewa University, Bauchi, Nigeria (e-mail: asgital@yahoo.com). Manuscript received 23 December, 2013; revised 13 January, 2014.

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

themselves do not fully understand all factors affecting the oil market. Therefore, expert systems is not a suitable computational technique for forecasting prices because they require complete knowledge to perform well and exhibit poor performance in case of omitted or incomplete information. However, neural networks (NNs) are capable of handling chaotic and partly understood environments. [11] pointed out that the fundamental mechanisms that drive the dynamics in the crude oil market are not fully understood. As such, rule induction techniques, including fuzzy logic, decision trees, and similar algorithms, are not suitable for crude oil price projection.

The handling effects of uncertainty events on crude oil prices, before projecting future prices using NNs is not sufficiently researched. It is highly important to consider factors such as the volatile nature of some regions that produce a significant percentage of world crude oil, threats from terrorist events, natural disasters, war related to crude oil, unexpected extreme weather, oil company mergers, and oil workers' strikes. Statistical tools are the predominant means to select feature subsets in the domain of crude oil price prediction. In addition, the application of genetic algorithms (GAs) for feature subset selection for oil price prediction is scarce despite GA being significantly better than statistical tools to select feature subsets [12]. [13] argued that a successful prediction requires a set of independent variables to form correlation relationships. To our knowledge, none of the literature in this domain has investigated this relationship.

Our approach is to apply a genetically optimized NN to nonlinear and volatile crude oil price projections because the optimized NN performs better than a conventional NN, which face the possibility of being stuck in local minima, and require trial and error searches to produce the optimal structure of the conventional NN. Furthermore, the optimized NN has the ability to model highly complex, dynamic, and nonlinear systems and does not require complete information to perform well unlike expert systems and other rule induction techniques. In the development of our NN-based predictor, a GA is to be applied for feature subset selection, and the correlation among independent variables will be investigated using factor analysis.

The remaining sections of the paper are organized as follows: Section II presents basic concepts of GAs and NNs. Section III explains uncertainty events. Section IV provides the proposed conceptual framework and its description, and Section V provides the conclusions. Finally, further research work is revealed in Section VI.

II. Basic Theories of Genetic Algorithm and Neural Network

A. Genetic Algorithm

The idea of GAs was conceived by [14], as a search method based on the principle of natural selection. In GA operations, an initial population with possible optimal solutions is generated. New populations are then created by selecting parents on the basis of their fitness, parents mingle (crossover) to create new offspring, and the strings that result from crossover are mutated in order to prevent a solution that is a local minimum. The new generated populations are used for further GA search steps, where the production of new generations (evolution) continues until a stoppage point is reached, by testing for some specified condition or set of conditions. If a stoppage condition is satisfied, the best chromosome within the current population is returned as the overall optimal solution [15].

B. Neural Network

NNs are computer models constructed to mimic the functions of the human brain through parallel computation of multiple input vectors. NNs comprise neurons distributed in input, hidden, and output layers. Neurons in the input layer supply inputs to neurons in hidden layers. Signals to each hidden unit consist of the weighted sum of each input unit, and these are transformed to an output value by an activation function such as the sigmoid function. The computed output is weighted and passed forward to neurons in the subsequent layer, thereby creating a feed-forward path to the output layer. Weights connecting two neurons in opposite layers are iteratively adjusted throughout the training process while inputs are fed to the network. Based on the pattern of connections between neurons in adjacent layers, the method of determining weights on the connections, and the node activation function (hidden and output layer nodes), the network is designed in a way that it can capture causal relationships between inputs and outputs in a dataset [16].

III. Uncertainty Events

Uncertainty events occur without prior knowledge of when they will occur, such as war, revolution, financial crises, terrorist attacks, political conflicts, false news, natural disasters, earthquakes, and extreme weather conditions. These types of uncertainties, when related to crude oil, have significant effects on the price of crude oil and will contribute to oil price volatility. In a study conducted by [17], prices of crude oil before, during, and after the 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 (WTI) and Brent crude oil prices, dating from March 30th, 1990 to May 31st, 1991. The aim was to assess the impact of unexpected events on crude oil prices. Empirical evidence from the study indicated that crude oil prices normally spike during unexpected events and return to normal prices after the event ends. In a related study, [18] analyzed daily crude oil prices for a period of 25 years with information extracted from the US Department of Energy, and they noted that unexpected events affect crude oil prices.

IV. Proposed Conceptual Framework

Our study focuses on crude oil price prediction under uncertainty events. Crude oil markets are complex and dynamic, which are characterized by high volatility, nonlinearity, high levels of uncertainty, nonstationary behavior, and hidden relationships, and are affected by both demand/supply and uncertainty events. This study adopts emerging GA and NN computational techniques to perform this challenging task. Using GAs and NNs, we can construct an intelligent model capable of handling the effects of uncertainty events and project crude oil prices on the basis of the flow chart illustrated in Fig 1. The conceptual framework process consists of three major stages: data collection, cleaning, and preprocessing; intelligent model construction; and model evaluation. The conceptual framework shown in Fig 1 is meant to address the problems pointed out in the introductory section. The three major stages of our framework are described in the following section.

A. Stage 1: Data Collection, Cleaning and Preprocessing

Historical data of Brent crude oil prices and regular factors affecting crude oil markets were collected from the Energy Information Administration website of the US Department of Energy from 1987 through 2011. These regular factors are OPEC crude oil production, World crude oil production, US domestic crude oil supply, non-OPEC crude oil production, Organization for Economic Cooperation and Development (OECD) crude oil consumption, OECD ending stocks, US crude oil production, US crude oil imports, US ending stocks of gasoline, and US crude oil stocks at refineries.

In a large real-world database, there is a very high probability of having inconsistency, outliers, missing values, typing mistakes, and differences in measurement units, among others. These anomalies prompt the need for data cleaning and preprocessing, which constitute important aspects of an intelligent model construction process, so that any anomaly found in the data can be corrected or removed. The data collected for this research is real-world data. Therefore, cleaning and preprocessing must be performed on the collected data in order to improve the quality of the data, and thus, improve the performance accuracy of the intelligent model. Despite the data being collected from credible source, it was thoroughly inspected for possible missing values, outliers, and other inconsistencies, but the majority of the data was found to be complete. Only data for US domestic crude oil supply and OPEC crude oil production were found to have missing values but an imputation technique was used to complete the missing values. The data was normalized in order to prevent the possible saturation of neurons, reduce computational complexity, and improve accuracy.

Feature subset selection plays a significant role in successful pattern recognition and data mining task in which optimum results can be obtained in a relatively small computational time. There are several features that determine the price of crude oil which are very difficult to accommodate in our intelligent model construction. In real-world data, only a small portion of the available subset features are relevant for a particular task (e.g., prediction). Some of the inputs can be eliminated because of redundancy and irrelevance. However, great caution and application of a systematic procedure are necessary for selecting the optimal feature subsets that adequately represent all features and enhance performance accuracy of the intelligent model. In our proposed system, GA is applied for optimal feature subset selection. The feature subsets will be subjected to factor analysis to determine any relationship among independent variables.

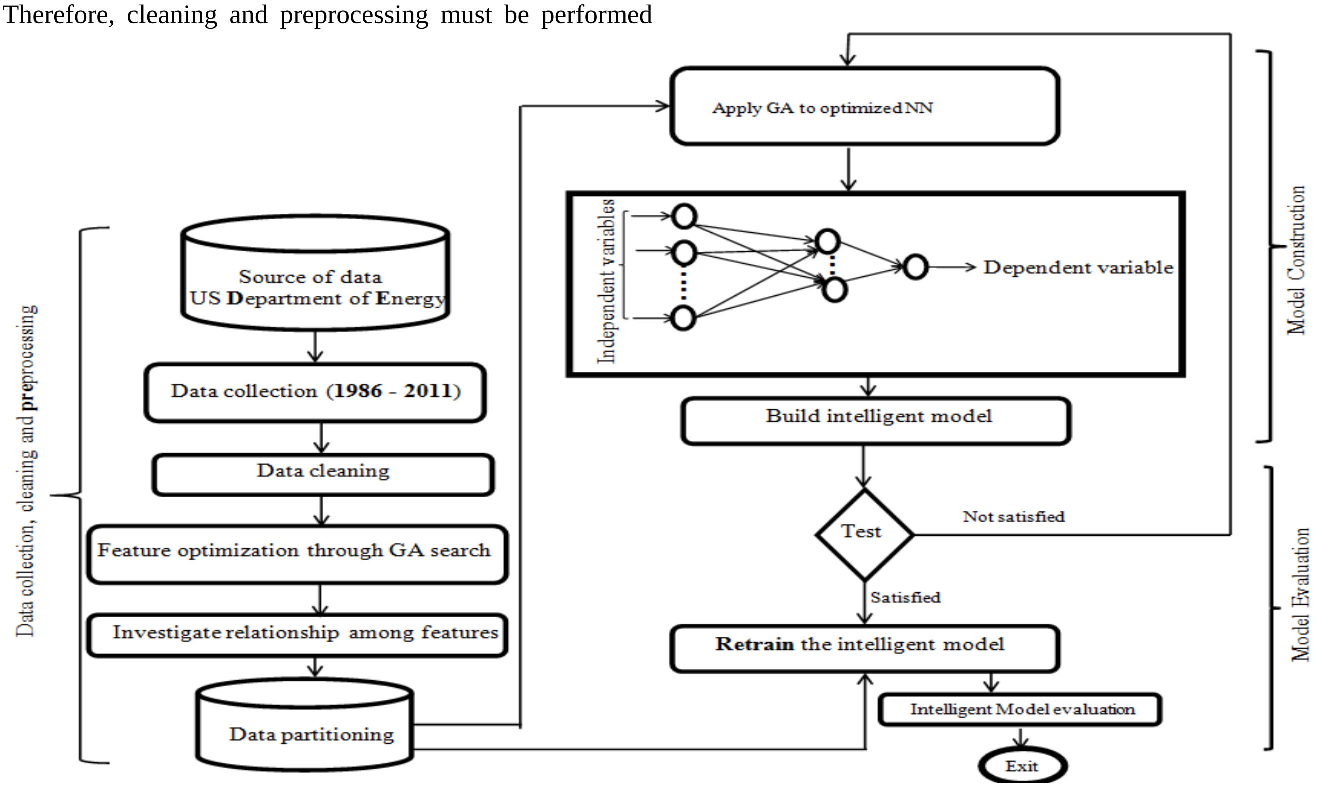

Fig 1 Conceptual framework for the proposed intelligent system implementation

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Hamilton [3] revealed that crude oil prices were significantly affected during the first Gulf War, Venezuelan unrest, and the second Gulf War. In [18], evidence indicated that oil markets were distorted during the Asian financial crisis, the US twin tower attack, and the world financial recession of 2007. Therefore, these uncertainty events are selected for our case study.

A. Stage 2: Intelligent model construction

For the first phase of the intelligent model construction, GA is proposed to be implemented with three different combinations of population size, crossover, and mutation probabilities (adopted from [19 – 21] in three separate trials so as to optimize the architectural configurations of the NN. The best NN architecture with superior performance will be adopted and used for constructing the intelligent model using data that is available prior to the unexpected events.

New sets of data distorted during the first Gulf War of 1991 will be added into the training data and the intelligent model will be retrained, leaving out oldest dataset (window training), so that the proposed model learns and captures new patterns in the data. This process will be periodically repeated with data distorted during the Asian financial crisis of 1997, Venezuelan unrest of 2002, the second Gulf War of 2003, the US twin tower attack of 2001, and the global financial recession of 2007. In each of the experiments, the intelligent model is expected to capture the effects of uncertainty events on crude oil prices, learn new patterns, and subsequently, project crude oil prices.

B. Stage 3: Model evaluation

Box – Jenkins model is the widely accepted benchmark for evaluating the effectiveness of NN [22 – 23]. For evaluation purpose, the Box–Jenkins model will be applied to predict crude oil prices and compare the results obtained with our model. Any significant statistical difference between performance of the Box–Jenkins and the intelligent models in terms of prediction accuracy will be measured.

V. CONCLUSIONS

This novel approach of crude oil price projection is expected to produce more accurate results than the intelligent models discussed in literature, and subsequently, provide more realistic projected prices of crude oil for proper planning by governments and private businesses. This may in turn help avert the negative effects of crude oil price volatility in times of crisis. This conceptual framework will possibly inspire researchers to focus attention on exploring other computational intelligent techniques capable of handling effects of uncertainties in the search for an optimal solution to oil price volatility, considering the critical role crude oil plays in the global economy.

VI FURTHER RESEARCH

Remaining part of stage 1 (feature subset selection and correlation relationships of subsets), stage 2, and stage 3 of the proposed conceptual framework are currently ongoing research.

REFERENCES

[1] A. Khashman, I. N. Nwulu “Intelligent prediction of crude oil using support vector machines,” In: 9th IEEE Inter. Symposium Appl Machine Intell Informatics pp.165 – 169, 2011.

[2] M. Rast “Fuzzy neural networks for modeling commodity markets” In: Proc. IFSA WorldCongress and 20th NAFIPS Inter. Conf. vol. 2,pp. 952— 955, 2001.

[3] J. D. Hamilton “Historical oil schocks,” Prepared for the Handbook of Major Events in Economic History. Accessed 12 August, 2012, 2011.

[4] L. Kilian "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," American Economic Review, vol. 99, no.3, pp. 1053-69, 2009.

[5] A. Alizadeh, K. Mafinezhad “Monthly brent oil price forecasting using artificial neural networks and a crisis index,” In: Proc. Inter. Conf. Electronics Infor. Eng. Pp. 465 – 468, 2010.

[6] W. Shouyang, Y. Lean, K.K. LAI “Crude oil price forecasting with tei@i methodology,” J Syst Sci Complex vol. 18, no.2, pp. 145 – 166, 2005.

[7] Y. Bao, X. Zhang, L. Yu, K.L Keung, S. Wang “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. Accessed 12 June 2012, 2007.

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

[9] M.Z Mehdi “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. Accessed 20 May 2012, 2009.

[10] M. Sotoudeh, E. Farshad “Application of neural network for forecasting gas price in America,” J Math Comp Sci vol. 4, no.2, pp. 216 – 226 , 2012.

[11] L. Yu , S. Wang, K.L Keung “A generalized intelligent – agent – based fuzzy group forecasting model for oil price prediction,” In: Proc. IEEE Inter. Conf. Syst. man and cybernetics pp. 489 – 493, 2008.

[12] S. Oreski, D. Oreski, G. Oreski “Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment,” Expert Syst Appl vol. 39, pp.12605–12617, 2012.

[13] F.J. Hair, W.C. Black, J. B. Babin, R.E Anderson, “Multivariate data analysis,” Pearson Prentice Hall: New Jersey, 2010.

[14] J.H. Holland “Adaptation in natural and artificial systems,” MIT Press: Massachusetts. Reprinted in 1998, 1975.

[15] A.E. Amin “A novel classification model for cotton yarn quality based on trained neural network using genetic algorithm,” http://dx.doi.org/10.1016/j.knosys.2012.10.008 Knowl. Based Syst. 2012.

[16] D. Zhang and L. Zhou “Discovering Golden Nuggets: Data Mining in Financial Application,” IEEE Trans syst man cybernetics—part c: appl reviews vol. 34, no. 4, pp. 512-522, 2004.

[17] X. Zhang, L. Yu , S. Wang, K.L Keung “Estimating the impact of extreme events on crude oil prices: An EMD – based event analysis method,” Energ econ vol. 31, pp. 768 – 778, 2009.

[18] A. Ortiz – Cruz, E Rodriguez, C. Ibarra-Valdez, J. Alvarez-Ramirez Efficiency of crude oil markets: evidence from

informational entropy analysis. Energ Policy vol. 4, pp. 365 – 373, 2012.
[19] K.A De Jong, “An Analysis of the Behavior of a Class of Genetic Adaptive Systems” Ph.D. thesis,University of Michigan, Ann Arbor, 1975.
[20] J. J. Grefenstette “Optimization of control parameters for genetic algorithms,” IEEE Transa Syst., Man, and Cybernetics vol. 16, no. 1: 122–128, 1986.
[21] M.A Reza, E. G Ahmadi “A hybrid artificial intelligence approach to monthly forecasting of crude oil prices time series,” In: The Proc. 10th Inter. Conf. Eng Appl Neural Networks, CEUR- WS284, pp. 160 – 167, 2007.
[22] V. Fernandez “Forecasting crude oil and natural gas spot prices by classification methods,” IDEAS web. http://www.webmanager.cl/prontus_cea/cea_2006/site/asocfile/ASOCFILE120061128105820.pdf. Access 4 June 2012, 2006.
[23] G.P Zhang “An investigation of neural networks for linear time-series forecasting,” Comput. Oper. Res. Vol. 2,pp. 1183-1202, 2001.

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)