Ensemble of Adaptive Neuro-Fuzzy Inference System Using Particle Swarm Optimization for Prediction of Crude Oil Prices

Lubna A.Gabralla[1], Talaat M.Wahby[1], Varun Kumar Ojha[3] and Ajith Abraham[1,2,3]

[1]Faculty of Computer Science & Information Technology, Sudan University of Science and Technology, Khartoum, Sudan
lubnagabralla@gmail.com,Talaat.wahby@gmail.com
[2]Machine Intelligence Research Labs, WA, USA,
[3]IT4Innovations, VSB - Technical University of Ostrava, Czech Republic
ajith.abraham@ieee.org, varun.kumar.ojha@gmail.com

***Abstract*— Oil is the lifeblood of the global economy. Recently, oil prices have witnessed fluctuations and the prediction of oil prices has become a challenge for researchers. The aim of this research is to design a model that is able to predict the prices of crude oil with good accuracy. We used the daily data from 1999 to 2012 with 14 input factors to predict the price of West Texas Intermediate (WTI), which is a well-known benchmark. We propose an ensemble of Adaptive Neuro-Fuzzy Inference System using a Particle Swarm Optimization algorithm for oil price prediction and the empirical results illustrate high performance and accurate results.**

***Keyword* – Ensemble; Adaptive Neuro-Fuzzy Inference System; Particle Swarm Optimization; prediction; fluctuating crude oil prices.**

I. INTRODUCTION

Without any doubt the physical production of energy is the basis of the global economy. Development of the economy depends on the different resources of energy that even most of economic sectors such as commercial, industrial and transportation are impossible to operate without energy. Figure1 shows the proportion of energy use in a number of sectors. Among the different energy sources, oil plays an important role to become the most efficient and most important source of energy. It stands for the primary origin of many manufactures such as plastics, chemicals, and more than half of the world's oil are used as fuel for aircraft, engines and cars. Figure 2 shows the oil supply compared to the other sources of energy.

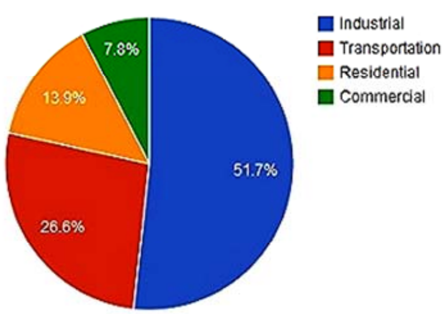

Figure 1: World energy consumption by sector, 2012 [1].

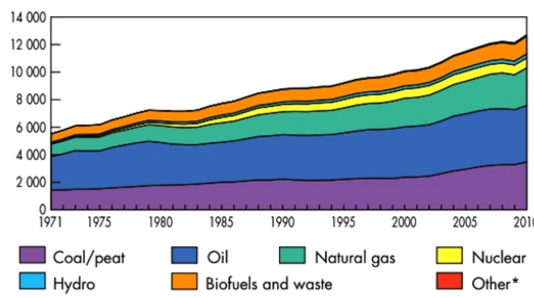

Figure 2: World total primary energy supply from 1971 to 2010 [2]

In recent years, oil prices suffer from the fluctuation because of the affects of political factors [3], economics [4], climate [5] and other several factors. These factors together have caused the oil prices to exhibit chaotic behavior and therefore, finding models to predict the oil prices has become a fertile area of research. It motivated many researchers to design the models that will be able to predict oil prices in the future in order to support the global economy, companies and institutions to hedge against surprise changes to make sound decisions and building a healthy and successful economy. Ensemble or hybrid technique is considered as one of the most promising scientific techniques to obtain satisfactory results. Hybrid system integrates two or more learning models by taking advantage of the strong points of each of the single techniques. Therefore, it has been applied to many areas and applications such as data mining [6] and medical applications [7]. This research uses an ensemble method for ANFIS, which is a fuzzy inference system fine tuned using neural network learning methods to improve the accuracy of predicting oil prices. The work has gone through several stages. The most prominent were pre-processing, dividing the data, constructing of base ANFIS models and finally the ensemble model. This paper is organized as follows: literature related work summarized in Section 2, methods are presented in Section 3. Section 4 describes the experiments and the analysis of results. Finally, conclusions for the work is presented in Section 5.

II. LITERATURE REVIEW

Hybrid techniques have been applied to predict crude oil prices and achieved acceptable results. This section introduces some of these studies with a focus on Neuro-fuzzy systems, which is used in our work. Liu et al. [8] used individual forecasting models, namely radial basis function neural networks, a Markov chain-based semi-parametric model and wavelet analysis as input to fuzzy neural network. They used Brent data for a period from 20 May 1987 to 30 August 2006, and the actual crude oil prices represented the output. The results indicated that the model was suitable to address forecasting of oil prices by achieving a high degree of prediction accuracy and reinforcement learning. Panella et al. [9] collected data from Europe (Brent crude oil) and the US (West Texas Intermediate crude oil) from 2001 to 2010 to forecast crude oil, natural gas, electricity, and coal prices using three different models radial basis function neural networks, adaptive neuro-fuzzy inference system networks and least-square approximation, The experimental results showed the superiority of adaptive neuro-fuzzy inference system. Zimberg [10] presented feed-forward neural network, Elman neural network, finally followed by Adaptive Neuro-Fuzzy Inference System (ANFIS), using West Texas Intermediate and Brent crude oil prices in the period from1991 to 2003. ANFIS results are higher in the term of accuracy when compared with the prediction based on an econometric model. Ghaffari and Zare [11] applied an adaptive network-based fuzzy inference system for forecasting WTI crude oil spot price. Using daily data from 5 January 2004, to 30 April 2007, 68.18% prediction accuracy was achieved.

III. RESEARCH METHODOLOGY

A. *Fuzzy inference system (FIS) [12]*

Mathematical and statistical methods are not well suitable for expression of human experiences such as perception, logic and uncertain concepts. A fuzzy inference system employing fuzzy if-then rules can provide a framework to model human knowledge. Takagi, Sugeno and Kang (TSK) [13] proposed a fuzzy inference method in which the conclusion of a fuzzy rule is constituted by a weighted linear combination of the crisp inputs rather than a fuzzy set.

B. *Adaptive Neuro Fuzzy Inference System (ANFIS) [14]*

There is no systematic way to transform experiences of knowledge of human experts to the knowledge base of a fuzzy inference system (FIS). On the other hand, Artificial Neural Network (ANN) learning mechanism hard to extract structured knowledge from either the weights or the configuration of the ANN. To overcome these drawbacks and to take advantages of these two approaches integrated system was built by combining the concepts of (FIS) and (ANN) modeling this system called Adaptive

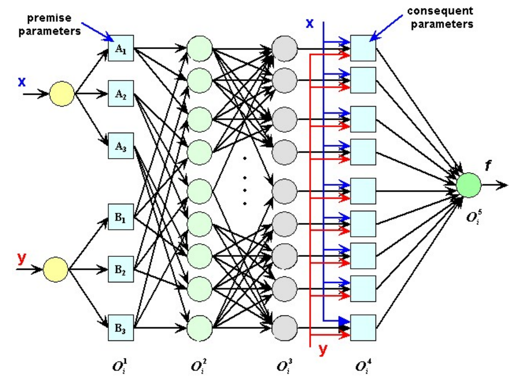

Figure 3: The architecture of the ANFIS [14]

Neuro Fuzzy Inference System (ANFIS). ANFIS implements a Takagi Sugeno Kang (TSK) fuzzy inference system. For a first order TSK model, a common rule set with two fuzzy *if-then* rules is represented as follows:

Rule 1: *If x is A1 and y is B1, then f1 = p1x + q1y + r1*
Rule 2: *If x is A2 and y is B2, then f2 = p2x + q2y + r2*

Where x and y are linguistic variables and A1, A2, B1, B2 are corresponding fuzzy sets and p1, q1, r1 and p2, q2, r2 are linear parameters. ANFIS makes use of a mixture of back propagation to learn the premise parameters and least mean square estimation to determine the consequent parameters. Figure 3 illustrates the ANFIS structure.

C. *The generalised ensemble method [15]*

The ensemble method depends on the behavior that a collection of predictor such as machine learning algorithms (neural network, support vector machine, decision trees and so on) can do better than the individual approaches. Predictors are combined through some weighted average or weighted combination. The generalized ensemble method find weights for each output that minimizes the Mean Absolute Error (MAE) of the ensemble. The general ensemble model (GEM) is defined by:

$$F_{GEM} = \sum_{i=1}^{n} \alpha F_i(X) \quad (1)$$

where
$$\sum_{i=1}^{m} \alpha = 1 \quad (2)$$

Where α F_i (x) are chosen to minimize the MAE between the outputs and the desired values. Finding the optimal values of α is not an easy task. We used a Particle swarm optimization method to determine the optimal weights. Particle swarm optimization [16] is a technique for simulating the social and cooperative behavior of different types such as birds, fish, bees and human beings. The PSO composed of a population (swarm) of possible solutions called particles. These particles move through the search domain with a specified velocity in search of optimal solutions. Each particle maintains a memory, which helps it in keeping the track of its previous best position.

IV. Empirical Results

All ANFIS experiments were implemented using the MATLAB R2013a software package. The steps of the proposed model are shown in Figure 4:

Step1: The dataset consists of 14 variables as listed below to predict the West Taxes Intermediate (Output) [1, 17] .

- Date (DT): The daily data from 1999 to 2012
- West Texas Intermediate (WTI): crude oil price benchmark plays an important role as a reference point to determine the price.

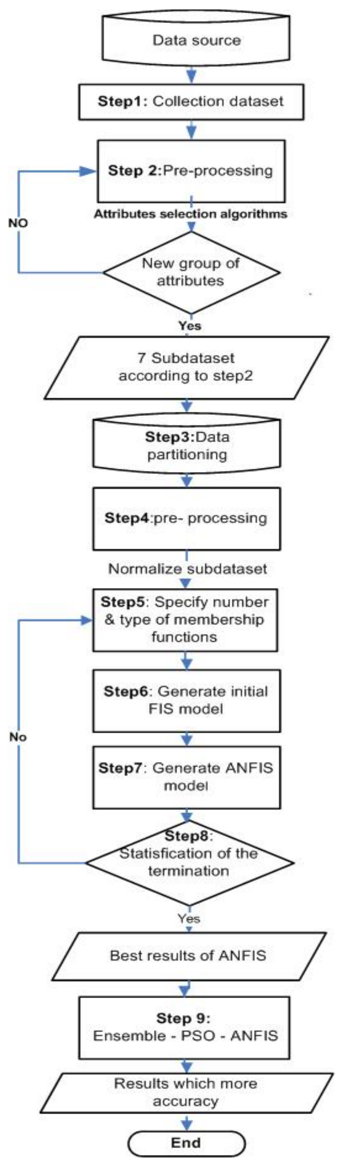

Figure 4: Flowchart for proposed model

- Federal Fund rate (FFR): one of the most influential interest rates in the U.S. economy, because it effects on monetary and financial conditions, which in turn have an impact on fundamental aspects of the broad economy including employment, growth and inflation [18].
- Volatility Implied Equity Index (VIX): measures the contribution of the instability of the market.
- The regional Standard & Poor's equity index (SPX): represent the market performance.
- Gasoline prices New York Harbor and US Gulf Coast (GPNY; GPUS): as example to assesses oil products.
- Heating oil spot prices (HP): as indication of seasonality in the energy market.
- Future contracts 1, 2, 3, 4 (FC1; FC2; FC3; FC4): for WTI to maturity traded on NYMEX.
- Exchange rate (ER): the price of oil and exchange rates of other currencies against the U.S. Dollar price.
- Gold prices (GP): gold is that less volatile than crude oil and could reflect the real trend in the commodity market rather than the noise and gold used as the results of investors hedge against inflation caused by the oil price shock [19].

***Step 2: Pre-processing* and *attribute selection*:** We used attribute selection algorithms as follows:

- Attributes ranking principal: ranked list of attributes based on evaluated individually each attribute [20].
- Wrapper attributes Selection: It depends on an induction algorithm to estimate the merit of feature subsets [21].
- ReliefF for regression: Evaluates quality of attributes according to the value in the given attribute for the near instance to each other and different predicted (class) value [22].
- Correlation based Feature Selection (CFS): assesses the value of a group of attributes by looking at the individual predictive ability of each feature as well with the possibility of repetition among the features [20].

Selecting a subset of the original attributes to reduce the dimensionality of the data and then constructing a model from these reduced number of features in some cases could improve the prediction accuracy and performance, and a simpler model that is easier to interpret [1] [20]. Table 1 summarizes the results of attribute selection and the 7 sub-data sets (SBDS1-SBDS7) obtained.

***Step 3 data partitioning*:** We also examined the effect of training and testing data by randomly splitting them as follows:

TABLE 1. ATTRIBUTE SELECTION METHODS AND THEIR FEATURES

Sub Dataset	Method	Features
SBDS 1	Correlation based Feature Selection subset evaluator	WTI; SPX; FG1
SBDS 2	Correlation based Feature Selection subset evaluator	DT; VIX; WTI; SPX; GPNY; GPUS; HP; ER; FC1; FC2; FC3; FC4
SBDS 3	Correlation based Feature Selection subset evaluator	VIX; WTI; GPNY; ER; FC1
SBDS 4	Correlation based Feature Selection subset evaluator	WTI; GPNY; FC1
SBDS 5	Correlation based Feature Selection subset evaluator	VIX; WTI; GPNY; FC1
SBDS 6	Wrapper subset evaluator	WTI; FC1
SBDS 7	Wrapper subset evaluator	WTI; GPUS

- 90% - 10% (A)
- 80% - 20% (B)
- 70% - 30% (C)
- 60% - 40% (D)

Step 4: *Pre-processing normalization:* We normalized the dataset within the range of [-1,1].
Step 5: *Specify number and type of membership functions*. We used two trapezoidal membership function for each input variable.
Step 6: *Generate initial TSK based FIS model* (grid partitioning).
Step 7: *Configure ANFIS system* and adjust the membership functions and consequent parameters using the hybrid learning method for 100 epochs.

The membership function, rule base and ANFIS structure are displayed in Figures 5, 6 and 7 respectively for SBDS7-B (two inputs). Also the four *if-then* rules will appear as follows:

If (input1 is in1mf1) and (input2 is in2mf1) then (output is out1mf1)
If (input1 is in1mf1) and (input2 is in2mf2) then (output is out1mf2)
If (input1 is in1mf2) and (input2 is in2mf1) then (output is out1mf3)
If (input1 is in1mf2) and (input2 is in2mf2) then (output is out1mf4)

Step 8: The training process is terminated after 100 epochs.
Step 9: *Ensemble approach:* Table 2 shows that the best results for each sub-dataset based on groups A and B. Then we used the Ensemble ANFIS for each group separately and generated dynamic weights by using PSO to optimize the MAE. The best results were obtained using group B as shown in Table 3. To ensure the effectiveness of the proposed model, we compared the results with previous works, the proposed model achieved the clear superiority. Table 3 shows this comparison.

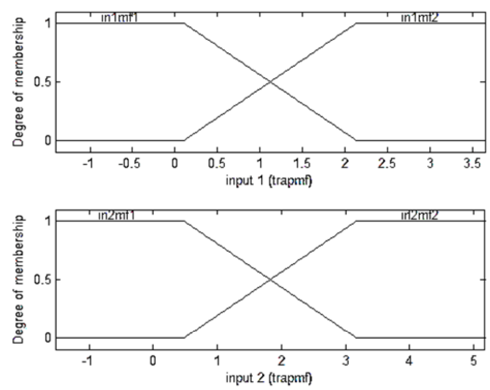

Figure 5: Trapezoidal-shaped membership function

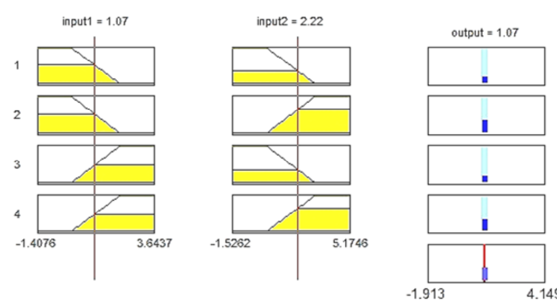

Figure 6: Developed TSK FIS using 2 inputs

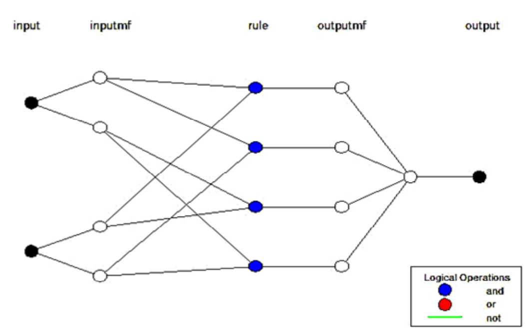

Figure 7: Developed ANFIS structure with 2 inputs

TABLE 2. ANFIS RESULTS FOR 7 SUB DATSETS

DATA	*Mean Absolute Error (MAE)*						
	SBDS 1	***SBDS 2***	***SBDS 3***	***SBDS 4***	***SBDS 5***	***SBDS 6***	***SBDS 7***
A	**1.23522E-05**	**8.94070E-06**	3.75200E-02	3.44406E-04	**9.96863E-04**	**2.39222E-06**	7.95274E-06
B	1.56636E-05	1.32849E-05	**2.89157E-02**	3.69172E-04	1.50347E-03	6.78925E-06	**4.70906E-07**
C	5.81167E-05	2.87358E-05	2.12495E-01	8.19114E-04	7.59600E-03	9.50917E-06	1.77291E-06
D	4.84699E-05	2.90189E-01	1.12921E-01	**1.90370E-04**	1.23274E-02	1.90200E-05	2.47023E-02

TABLE 3. ENSEMBLE ANFIS RESULTS

MAE-DATA (A)					
SBDS1	*SBDS2*	*SBDS5*	*SBDS6*	*Ensemble Average*	*Ensemble-PSO-ANFIS*
1.23522E-05	8.94070E-06	9.96863E-04	2.39222E-06	2.5011E-04	2.39222E-06
MAE-DATA (B)					
SBDS3		*SBDS7*		*Ensemble Average*	*Ensemble –PSO-ANFIS*
2.89157E-02		4.70906E-07		1.44578E-02	**4.62053E-07**
Previous results (MAE)					
Ensemble for IBL, KStar and SMOReg					0.2219 [23]
Ensemble of Random subspace and Multilayer Perceptron					0.0066 [24]

V. CONCLUSIONS

The problem of predicting oil prices is a complex issue because there are many factors that affect the oil prices directly and indirectly. In this paper, we used 14 different factors to find the price of WTI crude oil, to improve the results. We have implemented a number of steps such as attribute selection, normalization and data partitioning to 4 categories of training and testing. We finally developed an ensemble for ANFIS results with several sub datasets, and then generated dynamic weights by using PSO. Experimental results illustrate that the model performed well when compared to ourprevious results.

REFERENCES

[1] www.eia.gov.

[2] http://www.iea.org/publications/freepublications/publication/kwes.pdf.

[3] J. M. Griffin, "OPEC behavior: a test of alternative hypotheses," *The American Economic Review,* pp. 954-963, 1985.

[4] U. Soytas, R. Sari, S. Hammoudeh, and E. Hacihasanoglu, "World oil prices, precious metal prices and macroeconomy in Turkey," *Energy Policy,* vol. 37, pp. 5557-5566, 2009.

[5] R. S. Pindyck, "The dynamics of commodity spot and futures markets: a primer," *The Energy Journal,* pp. 1-29, 2001.

[6] Z. Zhao and Y. Zhang, "Design of ensemble neural network using entropy theory," *Advances in Engineering Software,* vol. 42, pp. 838-845, 2011.

[7] D. West, P. Mangiameli, R. Rampal, and V. West, "Ensemble strategies for a medical diagnostic decision support system: A breast cancer diagnosis application," *European Journal of Operational Research,* vol. 162, pp. 532-551, 2005.

[8] J. Liu, Y. Bai, and B. Li, "A new approach to forecast crude oil price based on fuzzy neural network," in *Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)*, 2007, pp. 273-277.

[9] M. Panella, F. Barcellona, and R. L. D'Ecclesia, "Forecasting energy commodity prices using neural networks," *Advances in Decision Sciences,* vol. 2012, 2012.

[10] B. Zimberg, "Crude oil price forecasting with ANFIS," *University of Vaasa web. Available online: http://lipas. uwasa. fi/~ phelo/ICIL2008TelAviv/24. pdf,* 2008.

[11] A. Ghaffari and S. Zare, "A novel algorithm for prediction of crude oil price variation based on soft computing," *Energy Economics,* vol. 31, pp. 531-536, 2009.

[12] L. A. Zadeh, "Roles of soft computing and fuzzy logic in the conception, design and deployment of information/intelligent systems," in *Computational intelligence: soft computing and fuzzy-neuro integration with applications*, ed: Springer, 1998, pp. 1-9.

[13] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *Systems, Man and Cybernetics, IEEE Transactions on,* pp. 116-132, 1985.

[14] J.-S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *Systems, Man and Cybernetics, IEEE Transactions on,* vol. 23, pp. 665-685, 1993.

[15] I. Maqsood, M. R. Khan, and A. Abraham, "An ensemble of neural networks for weather forecasting," *Neural Computing & Applications,* vol. 13, pp. 112-122, 2004.

[16] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the sixth international symposium on micro machine and human science*, 1995, pp. 39-43.

[17] www.usgold.com.

[18] http://www.investopedia.com/terms/f/federalfundsrate.asp.

[19] I. El-Sharif, D. Brown, B. Burton, B. Nixon, and A. Russell, "Evidence on the nature and extent of the relationship between oil prices and equity values in the UK," *Energy Economics,* vol. 27, pp. 819-830, 2005.

[20] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research,* vol. 3, pp. 1157-1182, 2003.

[21] M. A. Hall, "Correlation-based feature selection for machine learning," The University of Waikato, 1999.

[22] M. Robnik-Šikonja and I. Kononenko, "An adaptation of Relief for attribute estimation in regression," in *Machine Learning: Proceedings of the Fourteenth International Conference (ICML'97)*, 1997, pp. 296-304.

[23] L. A. Gabralla, R. Jammazi, and A. Abraham, "Oil price prediction using ensemble machine learning," in *Computing, Electrical and Electronics Engineering (ICCEEE), 2013 International Conference on*, 2013, pp. 674-679.

[24] L. A. Gabralla and A. Abraham, "Prediction of Oil Prices Using Bagging and Random Subspace," in *Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014*, 2014, pp. 343-354

.