

## ARTÍCULO

# Different methods for gas price forecasting

Hamid Abrishami,\* Vida Varahrami

<sup>a</sup>Faculty of Economics, University of Tehran, Tehran, Iran

Recibido el 10 de septiembre de 2011; aceptado el 21 de diciembre de 2011

### JEL CLASSIFICATION

C45;  
Q3;  
Q4

### KEYWORDS

Gas price forecasting;  
Group Method of Data  
Handling (GMDH)  
neural networks;  
Genetic Algorithm (GA);  
Hybrid Intelligent  
System;  
Rule-based Expert  
System (RES);  
MLF neural networks

### CÓDIGOS JEL

C45;  
Q3;  
Q4

### PALABRAS CLAVE

Previsión de los  
precios del gas;

**Abstract** The difficulty in gas price forecasting has attracted much attention of academic researchers and business practitioners. Various methods have been tried to solve the problem of forecasting gas prices however, all of the existing models of prediction cannot meet practical needs.

In this paper, a novel hybrid intelligent framework is developed by applying a systematic integration of GMDH neural networks with GA and Rule-based Expert System (RES) employs for gas price forecasting. In this paper we use a new method for extract the rules and compare different methods for gas price forecasting.

Our research reveals that during the recent financial crisis period by employing hybrid intelligent framework for gas price forecasting, we obtain better forecasting results compared to the GMDH neural networks and MLF neural networks and results will be so better when we employ hybrid intelligent system with for gas price volatility forecasting.

© 2011 Asociación Cuadernos de Economía. Published by Elsevier España, S.L. All rights reserved.

## Métodos para la previsión de los precios del gas

### Resumen

La dificultad de la previsión de los precios del gas ha atraído considerablemente la atención de los investigadores universitarios y los profesionales del sector. A pesar de que se ha intentado solucionar el problema de la previsión de los precios del gas con diferentes métodos, ninguno de los modelos de predicción existentes llegan a cumplir con las necesidades prácticas.

En este artículo, se ha desarrollado un novedoso sistema inteligente híbrido mediante la aplicación de la integración sistemática de redes neuronales de tipo *Group Method of Data Handling* (GMDH) con algoritmos genéticos (AG) y un sistema experto basado en reglas (SER) a la

\*Corresponding author.

E-mail: abrihami@ut.ac.ir (H. Abrishami).

Redes neuronales de tipo *Group Method of Data Handling* (GMDH);  
 Algoritmo genético (AG);  
 Sistema inteligente híbrido;  
 Sistema experto basado en reglas (SER);  
 Redes neuronales de tipo *Multi-Layer Feed-forward* (MLF)

previsión de los precios del gas. Igualmente, utilizamos un nuevo método para extraer las reglas y comparar los diferentes métodos para la previsión de los precios del gas.

Nuestra investigación revela que durante la reciente crisis económica se obtienen mejores resultados utilizando un sistema inteligente híbrido para la previsión de los precios del gas, en comparación con las redes neuronales de tipo GMDH y de tipo *Multi-Layer Feed-forward* (MLF), y que los resultados mejorarán si utilizamos un sistema inteligente híbrido en la previsión de la volatilidad de los precios del gas.

© 2011 Asociación Cuadernos de Economía. Publicado por Elsevier España, S.L. Todos los derechos reservados.

## 1. Introduction

Problems of complex objects modeling such as analysis and prediction of stock market, gas price and other such variables cannot be solved by deductive logical-mathematical methods with needed accuracy with a suitable number of hidden units. Neural networks get their intelligence from learning process, and then this intelligence makes them have the capability of auto-adaptability, association and memory to perform certain tasks. Gas price is primarily formed by supply and demand forces but is also influenced by factors such as gas products inventory levels, stock markets activities, foreign exchange rates, and political context.

In time series analysis, a review of the methodological linkage between statistical techniques and neural networks is given by Cheng and Titterton (1994). In comparison with statistical techniques, neural networks make less restrictive assumptions on the underlying distributions and provide a higher degree of robustness. Kuo and Reitsch (1995) showed that neural networks provide meaningful predictions when independent variables were correlated or missing. It is also known that neural networks tended to outperform the conventional regression analysis at the presence of ambiguity in independent variables. It is not surprising to learn that neural networks are superior to traditional approaches in terms of parsimony of parameterization. In addition, a network structure is trained by using part of the data and then tested by using the rest of the data. A well-trained network is therefore expected to provide robust predictions. A thorough literature review of neural network applications in finance and business are provided by Wang et al. (2004). Nasr et al. (2002) used artificial neural network (ANN) approach to gasoline consumption (GC) forecasting in Lebanon. Ambrishami et al. (2008) used GMDH neural network based on Genetic Algorithm to model and forecast the price of Gasoline by using two approaches; Deductive Method and Technical Analysis. The results of deductive method indicate that the accuracy of prediction could reach up to 96% and in technical analysis could reach up to 99%. Mehrara et al. (2008) used a GMDH neural network model with moving average crossover inputs to predict price in the crude oil futures market. The predictions of price are used to construct buy and sell signals for traders. Compared to those of benchmark models, cumulative returns, year-to-year returns, returns over a market cycle, and sharpe ratios all favor the GMDH

model by a large factor. The significant profitability of the GMDH model casts doubt on the efficiency of the oil futures market. Brito Buarque (2009) used methods of multiple linear regression and artificial neural networks for the prediction of gasoline properties from information of composition obtained by gas chromatography, as well as a methodology for prediction of properties using a hybrid method composed of neural networks and group contribution. Gencay (1996) use foreign exchange markets to pioneer the use of technical analysis rules as inputs for neural networks, which are flexible, nonlinear models with powerful pattern recognition properties. In a series of articles, Gencay et al. (1998) and Gencay and Ramazan (1998) and Gencay and Ramazan (1999) show that simple technical rules result in significant forecast improvements for current returns over a random walk model for both foreign exchange rates and stock indices.

In this paper, we employ moving average daily gas prices from 2006 to 2010 for forecasting the gas price and which are then modeled by developed a GMDH neural networks model and MLF neural network. In addition, the effects of irregular and infrequent events on gas price are explored by using RES techniques. Over all, we observed that the hybrid intelligent framework improve the forecasting results of gas price.

This paper is organized as follows. Section 2 provides modeling using neural network. In Section 3, general discussion of RES and GMDH neural networks modeling is presented. In Section 4, empirical results are presented and Section 5 offers concluding reviews.

## 2. Modeling using Neural Network

Artificial Neural Networks (ANN) is biologically inspired network based on the organization of neurons and decision making process in the human brain (Madala and Ivakhnenko, 1994). In other words, it is the mathematical analogue of the human nervous system. This can be used for prediction, pattern recognition and pattern classification purposes. It has been proved by several authors that ANN can be great used when the associated system is so complex that the underline processes or relationship are not completely understandable or display chaotic properties (Srinivasan et al., 2002). Development of ANN model for any system involves three important issues: (i) topology of the network, (ii) a proper training algorithm and (iii) transfer function.

Basically an ANN involves an input layer and an output layer connected through one or more hidden layers. The network learns by adjusting the inter connections between the layers. When the learning or training procedure is completed, a suitable output is produced at the output layer. The learning procedure may be supervised or unsupervised. In prediction problem supervised learning is adopted where a desired output is assigned to network before hand (Nariman-zadeh et al., 2002).

### 2.1. MLFF Neural Network

MLFF neural network is one the famous and it is used at more than 50 percent of researches that are doing in economy field recently. This class of networks consists of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. In many applications the units of these networks apply a sigmoid function as a transfer function ( $f(x) = \frac{1}{1 + e^{-x}}$ ).

It has a continuous derivative, which allows it be used in back-propagation. This function is also preferred because its derivative is easily calculated:  $y' = y(1 - y)$ .

Multi-layer networks use a variety of learning techniques; the most popular is back-propagation algorithm (BPA). The BPA is a supervised learning algorithm that aims at reducing overall system error to a minimum. This algorithm has made multilayer neural networks suitable for various prediction problems. In this learning procedure, an initial weight vectors  $w_0$  is updated according to:

$$w_i(k + 1) = w_i(k) + \mu(T_i - O_i) f'(w_i x_i) x_i \tag{1}$$

Where,  $w_i \Rightarrow$  The weight matrix associated with  $i^{th}$  neuron;  $x_i \Rightarrow$  Input of the  $i^{th}$  neuron;  $O_i \Rightarrow$  Actual output of the  $i^{th}$  neuron;  $T_i \Rightarrow$  Target output of the  $i^{th}$  neuron, and  $\mu$  is the learning rate parameter.

Here the output values ( $O_i$ ) are compared with the correct answer to compute the value of some predefined error-function. The neural network is learned with the weight update equation (1) to minimize the mean squared error given by:

$$E = \frac{1}{2} (T_i - O_i)^2 = \frac{1}{2} [t_i - f(w_i x_i)]^2$$

By various techniques the error is then fed back through the network. Using this information, the algorithm adjusts the weights of each connection in order to reduce the value of the error function by some small amount. After repeating this process for a sufficiently large number of training cycles the network will usually converge to some state where the error of the calculations is small. In this case one says that the network has learned a certain target function. To adjust weights properly one applies a general method for non-linear optimization that is called gradient descent. For this, the derivative of the error function with respect to the network weights is calculated and the weights are then changed such that the error decreases.

The gradient descent back-propagation learning algorithm is based on minimizing the mean square error. An alternate

approach to gradient descent is the exponentiated gradient descent algorithm which minimizes the relative entropy.

### 2.2. GMDH neural networks

GMDH neural networks are based on the concept of pattern recognition, and in that sense such networks are a refinement of traditional methods of technical analysis. They are highly flexible, semi parametric models, and have been applied in many scientific fields, including biology, medicine and engineering.

For economists, neural networks represent an alternative to standard regression techniques and are particularly useful for dealing with non-linear unvaried or multivariate relationships.

By applying GMDH algorithm a model can be represented as set of neurons in which different pairs of them in each layer are connected through a quadratic polynomial and thus produce new neurons in the next layer. Such representation can be used in modeling to map inputs to outputs. The formal definition of the identification problem is to find a function  $\hat{f}$  that can be approximately used instead of actual one,  $f$ , in order to predict output  $\hat{y}$  for a given input vector  $X=(x_1, x_2, x_3, \dots, x_n)$  as close as possible to its actual output  $y$ . Therefore, given  $M$  observations of multi-input-single-output data pairs so that:

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad i = 1, 2, \dots, M \tag{1}$$

It is now possible to train a GMDH-type neural network to predict the output values  $y^i$  for any given input vector  $X=(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})$ , that is:

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad i = 1, 2, \dots, M \tag{2}$$

The problem is now to determine a GMDH-type neural network so that the square of difference between the actual output and the predicted one is minimized, in the form of:

$$\sum_{i=1}^M [\hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) - y_i]^2 \rightarrow \min \tag{3}$$

General connection between inputs and output variables can be expressed by a complicated discrete form of the Volterra functional series that is:

$$\hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \tag{4}$$

$$n = 1, 2, \dots, N$$

This is known as the Kolmogorov-Gabor (Farlow, 1984; Iba et al., 1996; Ivakhnenko, 1971; Nariman-zadeh et al., 2002; Sanchez et al., 1997). The full form of mathematical description can be represented by a system of partial quadratic polynomials consisting of only two variables (neurons) in the form of:

$$\hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \tag{5}$$

$$i = 1, \dots, M, j = 1, 2, \dots, N$$

In this way, such partial quadratic description is recursively used in a network of connected neurons to build

the general mathematical relation of inputs and output variables given in Eq. (4). The coefficients  $a_i$  in Eq. (5) are calculated using regression techniques (Farlow, 1984; Nariman-zadeh et al., 2003) so that the difference between actual output,  $y$ , and the calculated one,  $\hat{y}$ , for each pair of  $x_i, x_j$  as input variables is minimized. Indeed, it can be seen that a tree of polynomials is constructed using the quadratic form given in Eq. (5) whose coefficients are obtained in a least-squares sense. In this way, the coefficients of each quadratic function  $G_i$  are obtained to optimally fit the output in the whole set of input-output data pairs, that is:

$$E = \frac{\sum_{i=1}^M (y_i - G_i)^2}{M} \rightarrow \min \quad (6)$$

In the basic form of the GMDH algorithm, all the possibilities of two independent variables out of total  $n$  input variables are taken in order to construct the regression polynomial in the form of Eq. (5) that best fits the dependent observations  $(y_i, i=1, 2, \dots, M)$  in a least-squares sense. Consequently,  $\binom{n}{2} = \frac{n(n-1)}{2}$  neurons will be built up in the first hidden layer of the feed forward network from the observations  $\{(y_i, x_{ip}, x_{iq}); (i=1, 2, \dots, M)\}$  for different  $p, q \in \{1, 2, \dots, n\}$ . In other words, it is now possible to construct  $M$  data triples  $\{(y_i, x_{ip}, x_{iq}); (i=1, 2, \dots, M)\}$  from observation using such  $p, q \in \{1, 2, \dots, n\}$  in the form:

$$\begin{bmatrix} x_{1p} & x_{1q} & | & y_1 \\ x_{2p} & x_{2q} & | & y_2 \\ \dots & \dots & | & \dots \\ x_{Mp} & x_{Mq} & | & y_M \end{bmatrix} \quad (7)$$

Using the quadratic sub-expression in the form of Eq. (5) for each row of  $M$  data triples, the following matrix equation can be readily obtained as:

$$Aa = Y \quad (8)$$

Where  $a$  is the vector of unknown coefficients of the quadratic polynomial in Eq. (5).

$$a = \{a_0, a_1, a_2, a_3, a_4, a_5\} \quad (9)$$

And  $Y = \{y_1, y_2, y_3, \dots, y_M\}^T$  is the vector of output's value from observation. It can be seen that:

$$\begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^2 & x_{1q}^2 \\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & x_{2p}^2 & x_{2q}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}x_{Mq} & x_{Mp}^2 & x_{Mq}^2 \end{bmatrix} \quad (10)$$

The least-squares technique from multiple-regression analysis leads to the solution of the normal equations as shown in Eq. (11):

$$a = (A^T A)^{-1} A^T Y \quad (11)$$

This determines the vector of the best coefficients of the quadratic Eq. (5) for the whole set of  $M$  data triples. It should be noted that this procedure is repeated for

each neuron of the next hidden layer according to the connectivity topology of the network. However, such a solution directly from normal equations is rather susceptible to round off errors and, more importantly, to the singularity of these equations. Recently, genetic algorithms have been used in a feed forward GMDH-type neural network for each neuron searching its optimal set of connection with the preceding layer (Nariman-zadeh et al., 2002). Jamali et al. (2006) have proposed a hybrid use of genetic algorithm for a simplified structure GMDH-type neural network in which the connections of neurons are restricted to adjacent layers. In this paper using GA for finding GMDH-type neural networks for modeling the Pareto optimized data.

### 3. The Hybrid Intelligent System for forecasting

A superior approach is employed to develop a hybrid intelligent system that can implement gas price forecasting in the volatile gas market. The hybrid intelligent system for gas price forecasting consists of GMDH based time series forecasting module, RES module with rules extract from regression.

In Section 3.1 RES is reviewed. Section 2.2 covers GMDH neural network.

#### 3.1. Rule-based Expert System (RES)

A rule-based expert system has five components: the knowledge base, the database, the inference engine, the explanation facilities, and the user interface.

The knowledge base contains the domain knowledge useful for problem solving. In a rule-based expert system, the knowledge is represented as a set of rules. Each rule specifies a relation, recommendation, directive, strategy or heuristic and has the IF (condition) THEN (action) structure. When the condition part of a rule is satisfied, the rule is said to fire and the action part is executed. The database includes a set of facts used to match against the IF (condition) parts of rules stored in the knowledge base. The inference engine carries out the reasoning whereby the expert system reaches a solution. It links the rules given in the knowledge base with the facts provided in the database. The explanation facilities enable the user to ask the expert system how a particular conclusion is reached and why a specific fact is needed. An expert system must be able to explain its reasoning and justify its advice, analysis or conclusion. The user interface is the means of communication between a user seeking a solution to the problem and an expert system. The communication should be as meaningful and friendly as possible. These five components are essential for any rule-based expert system (Negnevitsky, 2005).

The key to an expert system is the construction of its knowledge base (KB). In this study, KB is represented by all types of rules from knowledge engineers who collect and summarize related knowledge and information as well as from history and from domain experts. The main work of an RES module is to collect and extract the rules or knowledge category from the KB. Our expert system module is required to extract some rules to judge abnormal variability in the gas price by summarizing and concluding relationships

between gas price fluctuation and irregular key factors affecting gas price volatility. To formulate a useful price volatility mechanism to predict gasoline price movements one has to first observe historical price patterns that occur frequently in the gas market (Yu et al., 2003).

In this paper, the terms “patterns”, “factors” or “events” will be used interchangeably. The relationships between the gas price variability and the factors affecting gas price are examined.

Finally, if there are strong connections between price influencing factors and price movements, then the factors are elicited from the historical price patterns examined and a KB for predicting gas price variability can be constructed. As previously mentioned, world events such as wars can have an immediate impact on the gas price. Furthermore, these factors can exert either an individual or composite effect. In order to represent the irregular patterns in a more organized and systematic way, the price patterns are classified into individual patterns and combination patterns. Individual patterns that have relatively simple conditions and attributes are used in defining combination patterns. In this study, the pattern itself can be considered to be the representation of a rule because the conditions of a pattern can be seen as conditions of a rule in the rule representation. Figures 1 and 2 show how individual patterns and combination patterns are defined and constructed. The syntax of an individual pattern uses reserved words such as PATTERN, IF, AND, OR and EXPLANATION, as illustrated in figure 1. If certain important events are matched with the IF condition of a particular pattern, then the pattern is identified by the conditions, and the EXPLANATION part gives the information about what the pattern really means. The individual pattern itself has its own meaning and can be an important clue in predicting gas price volatility. Likewise, the combination patterns integrate several conditions or patterns to explain a certain sophisticated phenomenon, as illustrated in figure 2 (Wang et al., 200

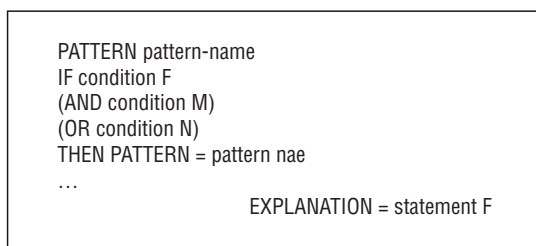


Figure 1 The syntax of individual pattern.

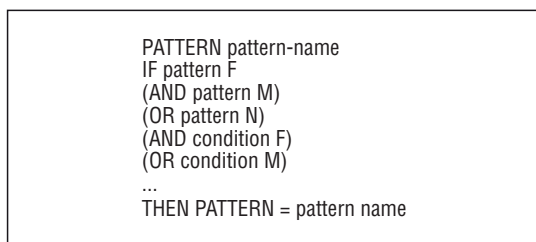


Figure 2 The syntax of a combination pattern.

## 4. Empirical results

In this Section, we first describe the data used in this research in Section 4.1 and then define some evaluation criteria for prediction purposes. Afterwards, the empirical results and explanations are presented in Section 4.2.

### 4.1. Data description

We employ daily Henry Hub Gulf Coast Natural gas spot price covering the period from January 1, 2006 through to December 31, 2010, based on gas contracts obtained from EIA. We use daily data of Natural Gas Futures Contract 1 (Dollars per Million BTU), WTI Spot Price FOB (Dollars per Barrel) and Crude Oil Future Contract 1 (Dollars per Barrel) for regression. Figure 3 shows trend of gas price.

For tractability, we utilize neural networks with two hidden layers and a direct connection between the lagged moving average crossovers and prices.

2 lags of the  $5[MA_5, MA_5(-1), MA_5(-2)]$ ,  $50[MA_{50}, MA_{50}(-1), MA_{50}(-2)]$ , day moving average crossover<sup>1</sup>, as input variables to the neural networks. The gas price data used in this study are daily Henry Hub Gulf Coast Natural gas spot price obtained from EIA (Energy Information Administration). We use the daily data from January 2006 to July 2009 as the in sample data sets for training and validation purposes and the remainder as the out of sample data sets for testing purposes.

In order to evaluate the prediction performance, it is necessary to introduce a forecasting evaluation criterion. In this study, two main evaluation criteria, root mean square error (RMSE) and direction statistics (Dstat) are introduced. The RMSE is calculated as: (Casella and Lehmann, 1999).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \tag{12}$$

Where  $e_i$  denotes the difference between forecasted and realized values and  $n$  is the number of evaluation periods. In the gas price forecasting, a change in trend is more important than precision level of goodness of fit from the viewpoint of practical applications. As a result, we introduce directional change statistics, Dstat. Its computational equation can be expressed as:

$$Dstat = \frac{1}{n} \sum_{i=1}^n |e_i| \tag{13}$$

Where  $e_i=1$  if  $(y_{t+1} - y_t) (\hat{y}_{t+1} - y_t) \geq 0$ , and  $e_i=0$  otherwise. (Wang et al., 2004).

In addition, as the effects on gas price of irregular events can be measured in the rational range, then the interval forecasting results can be obtained.

Subsequently, irregular events and their effects are examined and explored. We find the irregular events and

1. Such models are all based on rules using moving averages of recent prices. A typical moving average is simply the sum of the closing prices for the last  $n$  number of days divided by  $n$ , where  $n$  may be from 1 to 200 days the rules for using these tools are very similar and usually involve making a decision when a short-term average crosses over a long-term average. For example, the rule may be to buy when the 5-day moving average exceeds the 50-day moving average and to sell when the 5-day average is below the 50-day average (Gencay et al., 1996).

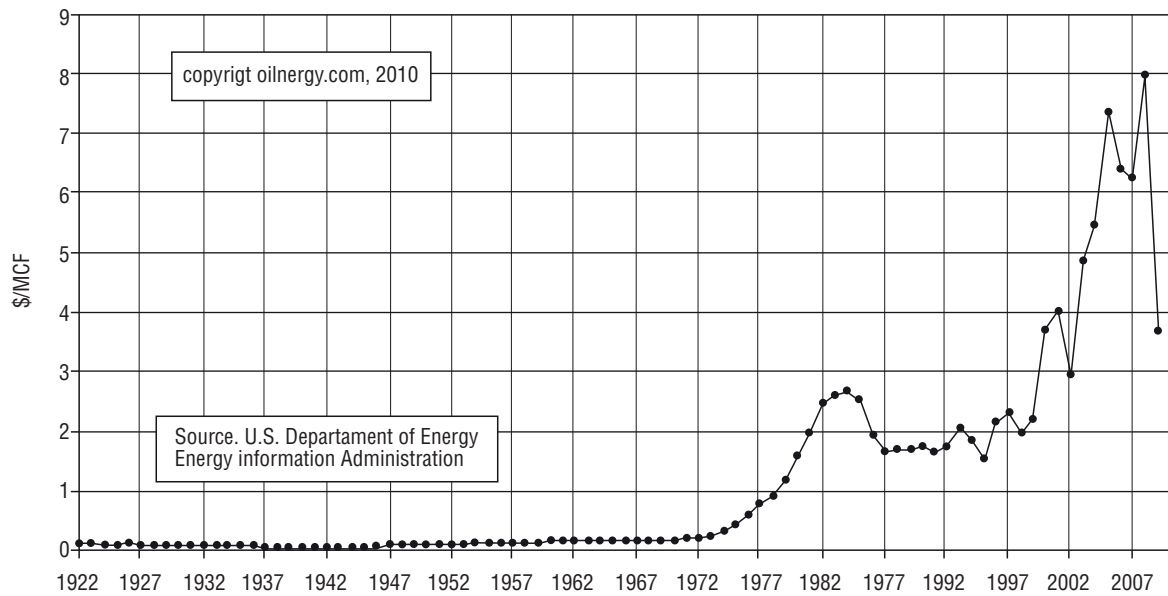


Figure 3 Evolution of gas price in \$/MCF (1922-2007).

RES is utilized to measure the degree of impact of these irregular events.

We find some irregular events that affect the gas price from the Internet<sup>2</sup>. All of these factors are created from theories, therefore we survey all papers and theories which are about factors that affect on gas price and then we choose our variables. In this paper, we consider Henry Hub Gulf Coast Natural gas price and we consider environment policy in USA and Europe. All of forecasting in this paper is not out of sample but we can use from this method for out of sample forecasting. Some main factors are concluded by analyzing past events, as shown in table 1.

Then for extract rules, we regress gas price on factors that affect it. There for I regress gas price as a dependent variable on oil price, crude oil futures price, natural gas futures price, global demand for gas, global demand for crude oil and dummy variable for environmental policy, taxes placed on gas price and OPEC cut production.

We regress these variables with OLS and shown results in table 2.

Table 2 presents in details the main judgmental or forecasting rules in this study according to the extraction of historical events affecting the gas price and contracts obtained from EIA (Energy Information Administration). In this table we are shown every events that affect gas price, these events are collected from sites. Therefore from table 2 we can build rules, for example (as shown in table 3) when crude oil price increases 1 percentage then gas price increases 25 percentages.

With the help of this information, one can judge the effect of irregular future events on the gas price by using the RES module. The rules should be adjusted with time and events in order to keep the expert system robust. We are shown

rules in blow table and enter these rules to GMDH neural network and build an intelligent system for forecasting.

#### 4.2. A simulation study

We employ a simulation experiment for proposed the hybrid intelligent system for gas price forecasting. In the simulation study, we reveal that forecasting rules from expert system and moving average gas price are modeled by using GMDH neural networks. We used the Multi-Objective Optimization Program (Atashkari et al., 2007) and Pareto based multi-objective optimization (Amanifard et al., 2008) that was designed with this target: reducing error in modeling and forecasting that simultaneously increase the exactitude of forecasting and the stability of process for measurement the scale of variables effects in different patterns. Accordingly, the evaluation criteria are the root mean square error (RMSE) and direction change statistics (Dstat). For a comparison, the full evaluation period is divided into five sub-periods in terms of chronology. In addition, the individual GMDH forecasting method is used as a benchmark model in this research and we use MLF neural network for gas price forecasting and compare its results with other methods. The corresponding results are summarized in table 4.

It observed that the hybrid intelligent results is better than MLF and GMDH and GMDH results is better than MLF in terms of either RMSE or Dstat. Notably, the values of Dstat of our hybrid intelligent forecasting method for each evaluation period exceed 70%, indicating that the proposed hybrid intelligent forecasting approach has good performance for the gas price forecasting considering the complexity of the gas market.

Focusing on the RMSE indicator, in the case of individual GMDH method, the second sub-period 2007 performs the best, followed by 2006, 2008, 2009 and 2010. While in the case of the hybrid intelligent method, the results of 2008 outperform those of the other evaluation period.

2. [www.bloomberg.com](http://www.bloomberg.com); [www.wtrg.com/prices.htm](http://www.wtrg.com/prices.htm); [www.washingtontimes.com](http://www.washingtontimes.com); [www.engdahl.oilgeopolitics.net](http://www.engdahl.oilgeopolitics.net); [www.iirenergy.com](http://www.iirenergy.com)

**Table 1** The factor classification

Year	Month	Day	Important events affecting gas price
2006	11	1-25	Environmental policy
2006	12	8-28	
2007	1	3-7	Taxes placed on gas price
		10-14	
		18-21	
		24-28	
2007	1	17-25	Increase natural gas futures prices
2007	6	2	Environmental policy
2007	5	10-13	OPEC cut production affected gas price
		17-21	
		24-28	
2007	12	18	Taxes placed on gas price
2008	1	2-5	Global demand of gas increased and affected gas price
		8-12	
		16-19	
2008		22-26	
2008	1	22-29	OPEC cut production affected gas price
2008	2	5-9	Global demand of gas increased and affected gas price
		12-16	
		26-28	
2008	2	12-16	OPEC cut production affected gas price
2008		20-23	
2008	6	5	OPEC cut production affected gas price
2008	7	10-25	Increase natural gas futures prices
2008	10	15	Environmental policy
2008	12	3-12	Increase natural gas futures prices
2009	3	4-6	Increase natural gas futures prices
		10-12	
		26-30	
2009	7	1-3	Increasing global gas demand
2009	11	10-14	Crud oil price increased and affected gas price
		18-21	
		25-28	
2010	4	16	Environmental policy
2010	10	10-13	Increasing global gas demand
		17-20	
2010	12	3-18	OPEC cut production affected gas price

**Table 2** OLS results

Variables	Coefficient
Crude oil price	0.25
Crude oil futures price	0.12
Natural gas futures price	0.23
Environmental policy	0.06
Taxes placed on gas price	0.26
OPEC cut production	0.17
Global demand for crude oil	0.09
Global demand for gas	0.16

**Table 3** The typical rules

Rule NO	Condition	Direction movements	The movements (%)
1	Increasing crude oil price	Increase	25
2	Taxes placed on gas price	Increase	26
3	OPEC design to cut production	Increase	17
4	Increasing global gas demand	Increase	16
5	Increasing global crude oil demand	Increase	9
6	Environmental policy	Increase	6
7	Increasing crude oil futures price (increasing demand for crude oil futures contracts increase)	Increase	12
8	Natural gas futures price (increasing Demand for gas futures contracts)	Increase	23

The main reason is that many important events affecting gas price volatility happened. The information of those important events could be obtained.

From a practitioners' point of view, the Dstat indicator is more important than the RMSE. This is because the former can reflect the movement trend of gas price and can help traders to make good trading decisions. For the test case of our hybrid intelligent approach and from the view of Dstat, the performance of 2008 is much better than 2006, 2007, 2009 and 2010, as shown in table 4.

From table 4, we observe that a smaller RMSE does not necessarily mean a higher Dstat value. For example, for the test case of the individual GMDH method, the RMSE for 2007 is slightly smaller than full-period period of 2006-2010, while the Dstat for 2006-2010 is larger than that for 2007. However, the overall prediction performance of the proposed hybrid intelligent approach is satisfactory because the RMSE for each evaluation period is smaller than 3.00 and the Dstat for each evaluation period exceeds 70%. Thus, the forecasting results of hybrid intelligent method is better than GMDH & MLF neural networks methods and this indicates that there are some profitable opportunities if traders use the proposed approach to forecast gas price.

### 5. Conclusions

In this paper, we find some irregular events that affect the gas price and reveal rules according to events affecting gas price and a hybrid intelligent framework integrating RES with GMDH neural networks is employed for gas price forecasting.

We observed that during the crisis period, when we investigate the effects of irregular and infrequent events on gas price by regression and RES, we obtain better forecasting results compared to the GMDH neural networks and MLF neural networks.

Overall, the obtained results reveals in 2008, when different important events took place, GMDH neural

**Table 4** The forecasting results of gas price for period of Jan. 2006 - Dec. 2010

Evaluation Method	Full period (2006-2010)	Sub-period I 2006	Sub-period II 2007	Sub-period III 2008	Sub-period IV 2009	Sub-period V 2010
<b>GMDH:</b>						
RMSE	3.482	3.516	3.118	3.219	3.311	3.433
<i>Dstat</i> (%)	59.28	54.28	50.21	68.34	55.71	58.24
<b>MLF:</b>						
RMSE	3.524	3.521	3.179	3.24	3.373	3.517
<i>Dstat</i> (%)	58.79	48.87	52.23	63.58	45.64	57.33
<b>Hybrid intelligent:</b>						
RMSE	2.495	2.942	2.866	1.830	2.171	2.234
<i>Dstat</i> (%)	81.37	73.49	75.81	92.00	75.44	82.26

networks and MLF neural networks can not reveal effects of these events on gas price forecasting and forecast's results of this methodology are not so well.

Hence, the novel hybrid intelligent forecasting model can be employed as an effective tool for gas price forecasting and can improve forecasting accuracy.

## References

- Abrishami, H., Mehrara, M., Ahrari, M., Varahrami, V., 2010. Gold price forecasting by hybrid intelligent systems with GARCH effects. *Far East Journal of Experimental and Theoretical Artificial Intelligence* 6, 111-128. Available from: <http://pphmj.com/journals/fejetai.htm>
- Abrishami, H., Moeini, A., Mehrara, M., Ahrari, M., Soleimanikia, F., 2008. Forecasting gasoline price using GMDH neural network based on GA. *Journal of Quarterly Iranian Economic Research*, Allameh Tabatabaiee University. 37-59.
- Amanifard, N., Nariman-zadeh, N., Borji, M., Khalkhali, A., Habibdoust, A., 2008. Modeling and Pareto optimization of heat transfer and flow coefficients in micro channels using GMDH type neural networks and genetic algorithms. *Energy Conversion and Management* 49, 311-325.
- Atashkari, K., Nariman-zadeh, N., Gölcü, M., Khalkhali, A., Jamali, A., 2007. Modeling and multi-objective optimization of a variable valve-timing spark-ignition engine using polynomial neural networks and evolutionary algorithms. *Energy Conversion and Management* 48, 1029-1041.
- Brito Buarque, H., 2009. Prediction of gasoline properties from composition data. Available from: [www.biblioteca.universia.net](http://www.biblioteca.universia.net)
- Casella, G., Lehmann, E.L., 1999. *Theory of point estimation*. Springer, New York.
- Cheng, B., Titterton, D., 1994. Neural networks: A review from a statistical perspective. *Statistical Science* 9, 2-30.
- Farlow, S.J., 1984. *Self-organizing method in modeling: GMDH type algorithm*. Marcel Dekker Inc, New York.
- Gencay, R., 1996. Non-linear prediction of security returns with moving average rules. *Journal of Forecasting* 15, 165-174.
- Gencay, R., 1998. The predictability of security returns with simple technical trading rules. *Journal of Empirical Finance* 5, 347-359.
- Gencay, R., 1999. Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. *Journal of International Economics* 47, 91-107.
- Gencay, R., Stengos, T., 1998. Moving average rules, volume and the predictability of security returns with feed forward networks. *Journal of Forecasting* 17, 401-414.
- Iba, H., deGaris, H., Sato, T., 1996. A numerical approach to genetic programming for system identifications. *Evolutionary Computation* 3, 417-452.
- Kuo, C.H., Reitsch, A., 1995. Neural network vs. conventional methods of forecasting. *The Journal of Business Forecasting Methods & Systems* 14, 17-25.
- Ivakhnenko, A.G., 1971. Polynomial theory of complex systems. *IEEE Trans. Syst, Man & Cybern*, SMC-1. pp. 364-378.
- Jamali, A., Nariman-zadeh, N., Atashkari, K., 2006. Inverse modeling of multi-objective thermodynamically optimized turbojet engines using GMDH and GA. 14th Annual (International) Mechanical Engineering Conference, Isfahan University of Technology, Isfahan, Iran.
- Madala, H.R., Ivakhnenko, A.G. *Inductive learning algorithms for complex systems modeling*. CRC Press Inc., Boca Raton, 1994, p. 384.
- Mehrara, M., Moeini, A., Ahrari, M., Erfanfard, A., 2008. Investigating the efficiency in oil futures market based on GMDH approach. *Expert Systems with Applications* 36, 7479-7483.
- Nariman-zadeh, N., Darvizeh, A., Darvizeh, M., Gharababaei, H., 2002. Modeling of explosive cutting process of plates using GMDH-type neural network and singular value decomposition. *Journal of Materials Processing Technology* 128, 80-87.
- Nasr, G.E., Badr, E.A., Joun, C., 2002. Cross entropy error function in neural networks; forecasting gasoline demand. *FLAIRS-02 Proceedings*. Available from: [www.aaai.org](http://www.aaai.org)
- Negnevitsky, M.A., 2005. *Artificial intelligence: A guide to intelligent systems* (2nd edition). Addison Wesley, Harlow, England; New York.
- Sanchez, E., Shibata, T., Zadeh, L.A., 1997. *Genetic algorithms and fuzzy logic systems*. World Scientific, River Edge, NJ.
- Srinivasan, N., Ravichandran, V., Chan, K.L., Vidhya, J.R., Ramakrishnan, S., Krishnan, S.M. Exponentiated backpropagation algorithm for multilayer feedforward neural networks. *Neural Information Processing*, 2002. *ICONIP apos; 02*. Proceedings of the 9th International Conference on Volume 1, 18-22 Nov. 2002. pp. 327-331.
- Wang, S., Yu, L., Lai, K.K., 2004. A novel hybrid AI system framework for crude oil price forecasting. Springer-Verlag Heidelberg, New York, pp. 233-242.
- Yu, L., Wang, S.Y., Lai, K.K., 2003. A hybrid AI system for forex forecasting and trading decision through integration of artificial neural network and rule-based expert system. *Expert System with Applications*. [www.bloomberg.com](http://www.bloomberg.com)  
[www.engdahl.oilgeopolitics.net](http://www.engdahl.oilgeopolitics.net)  
[www.iirenergy.com](http://www.iirenergy.com)  
[www.washingtontимаes.com](http://www.washingtontимаes.com)  
[www.wtrg.com/prices.htm](http://www.wtrg.com/prices.htm)