Different methods for gas price forecasting

Hamid Abrishami,* Vida Varahrami

*Faculty of Economics, University of Tehran, Tehran, Iran

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

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 atraido 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...
1. Introducción

Problemas de complejos objetos modelado como análisis y predicción de la Bolsa, el precio del gas y otros tales variables no pueden ser resueltos por métodos lógico-matemáticos deductivos con precisión suficiente y con un número de unidades ocultas adecuado. Las redes neuronales tienen una inteligencia aprendida, y esta inteligencia hace que se adquiera la capacidad de auto-adaptabilidad, asociación y memoria para ejecutar ciertas tareas. El precio del gas es principalmente formado por la oferta y la demanda, pero también está influido por factores como el inventario de productos, el mercado de valores y los mercados de divisas extranjeras.

En el análisis de series temporales, un estudio de la metodología de la interrelación entre las técnicas estadísticas y las redes neuronales es realizado por Cheng y Titterington (1994). En comparación con las técnicas estadísticas, las redes neuronales hacen menos suposiciones restrictivas sobre las distribuciones subyacentes y proporcionan un mayor grado de robustez. Kuo y Reitsch (1995) mostraron que las redes neuronales proporcionan predicciones significativas cuando las variables independientes fueron correlacionadas o faltantes. Es también conocido que las redes neuronales tienden a superar a los métodos convencionales de regresión lineal en presencia de ambigüedad en variables independientes. Es no sorprendente aprender que las redes neuronales son superiores a los métodos tradicionales en términos de parsimonia de parameterización. En adición, un modelo de red se configura por usan parte de los datos y luego entrenado usando el resto de los datos. Un modelo bien entrenado es entonces esperado para proporcionar predicciones robustas. Un amplio estudio de la red neuronal aplicada en la Bolsa, el gas y otras variables se proporciona por Wang et al. (2004). Nasr et al. (2002) usaron un modelo de red neuronal a través de un algoritmo genético (GC) para predecir el consumo de gasolina en Líbano. Abrishami et al. (2008) usaron un modelo de red neuronal basado en el algoritmo genético (GC) para predecir el precio de los combustibles. Brito Buarque (2009) utilizó métodos de regresión lineal múltiple y redes neuronales para prever las propiedades de la gasolina en función de la composición obtenida por gasometría por chromatografía. Brito Buarque (2007) utilizó modelos de regresión lineal y redes neuronales para prever el precio del gas. Igualmente, utilizamos un nuevo modelo 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 mejoran 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.

2. Modelado usando la Red Neuronal

Los Redes Neuronales (ANN) son una red neuronal biológicamente inspirada basada en la organización de los nervios y el proceso de toma de decisiones del cerebro humano (Madala y Ivakhnenko, 1994). En otras palabras, es el modelo matemático del sistema nervioso. Este puede ser utilizado para la previsión, patrón de reconocimiento y patrón de clasificación. Ha sido comprobado por varios autores que una red neuronal puede ser muy útil cuando el sistema es tan complejo que los procesos subyacentes o la relación no son completamente comprensibles o pueden representar un comportamiento caótico (Srinivasan et al., 2002). El desarrollo de un modelo ANN para cualquier sistema involucra tres importantes aspectos: (i) la topología del modelo, (ii) un algoritmo de entrenamiento correcto y (iii) transferencia de función.
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 to 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 \( w_i \) is updated according to:

\[
 w_i (k + 1) = w_i (k) + \mu (t_i - O_i) f'(w_i x_i)
\]

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 \( f \) that can be approximately used instead of actual one, \( f \), in order to predict output \( \hat{y} \) for a given input vector \( T=(x_1, x_2, x_3, \ldots x_n) \) as close as possible to its actual output \( y \). Therefore, given \( M \) observations of multi-input-single-output data pairs so that:

\[
y_i = f(x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in}) \quad l = 1, 2, ..., M
\]

It is now possible to train a GMDH-type neural network to predict the output values \( y^i \) for any given input vector \( X=(x_1, x_2, x_3, \ldots x_n) \), that is:

\[
y^i = f(x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in}) = f(T_i) \quad l = 1, 2, ..., M
\]

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:

\[
\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \rightarrow \min
\]

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

\[
\hat{y} = G(x_i, x_j) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2 + a_4 x_1^2 + a_5 x_2^2
\]

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:

\[
\hat{y} = G(x_i, x_j) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2 + a_4 x_1^2 + a_5 x_2^2
\]

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:

$$E = \frac{1}{M} \sum_{i=1}^{M} (y_i - \hat{y}_i)^2 \rightarrow \min$$

(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 ($y_i, i=1,2,\ldots,M$) 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 {($y_i, x_ip, x_iq$); ($i=1,2,\ldots,M$)} for different $p, q \in \{1, 2, \ldots, n\}$. In other words, it is now possible to construct $M$ data triples {($y_i, x_ip, x_iq$); ($i=1,2,\ldots,M$)} from observation using such $p, q \in \{1, 2, \ldots, n\}$ in the form:

$$\begin{bmatrix}
x_{ip} & x_{iq} & y_i \\
x_{p2} & x_{q2} & y_2 \\
\vdots & \vdots & \vdots \\
x_{Mn} & x_{Mn} & y_M
\end{bmatrix}$$

(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:

$$Aa = Y$$

(8)

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

$$a = \{a_0, a_1, a_2, a_3, a_4, a_5\}$$

(9)

And $Y = \{y_1, y_2, y_3, \ldots, y_M\}$ is the vector of output’s value from observation. It can be seen that:

$$\begin{bmatrix}
1 & x_{ip} & x_{iq} & x_{ip}^2 & x_{iq}^2 \\
1 & x_{p2} & x_{q2} & x_{p2}^2 & x_{q2}^2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & x_{Mn} & x_{Mn} & x_{Mn}^2 & x_{Mn}^2
\end{bmatrix}$$

(10)

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

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

(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 which 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., 2004).

### 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, 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).

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

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:

$$\text{Dstat} = \frac{1}{n} \sum_{i=1}^{n} \frac{|e_i|}{\max|e_i|}$$

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).
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. 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 Muti-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 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.

---

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...
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


www.bloomberg.com
www.engdahl.oilgeopolitics.net
www.iirenergy.com
www.washingtontimes.com
www.wtrg.com/prices.htm