Applied Energy 87 (2010) 268–274

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Modeling and estimation of the natural gas consumption for residential and commercial sectors in Iran

Mehdi Forouzanfar, Ali Doustmohammadi *, M. Bagher Menhaj, Samira Hasanzadeh

424, Hafez Ave., P.O. Box 15875-4413, Tehran, Iran

ARTICLE INFO

Article history:
Received 20 February 2009
Received in revised form 11 June 2009
Accepted 6 July 2009
Available online 6 August 2009

Keywords:
Natural gas consumption prediction
Natural gas consumption forecast
Logistic equation
Nonlinear programming
Genetic algorithm

ABSTRACT

In this paper, a logistic based approach is used to forecast the natural gas consumption for residential as well as commercial sectors in Iran. This approach is relatively simple compared with other forecasting approaches. To make this approach even simpler, two different methods are proposed to estimate the logistic parameters. The first method is based on the concept of the nonlinear programming (NLP) and the second one is based on genetic algorithm (GA). The forecast implemented in this paper is based on yearly and seasonal consumptions. In some unusual situations, such as abnormal temperature changes, the forecasting error is as high as 8.76%. Although this error might seem high, one does not need to be deeply concerned about the overall results since these unusual situations could be filtered out to yield more reliable predictions. In general, the overall results obtained using NLP and GA approaches are as well as or even in some cases better than the results obtained using some older approaches such as Cavallini's. These two approaches along with the gas consumption data in Iran for the previous 10 years are used to predict the consumption for the 11th, 12th, and 13th years. It is shown that the logistic approach with the use of NLP and GA yields very promising results.

1. Introduction

Natural gas is one of the most important energy sources in the world. The natural gas consumption growth has been the fastest of all the fossil fuels in recent years. While the share of oil in the world's total energy produced declined to 36.7% in 2000 from 45% in 1970, the share of natural gas has gone up to 22.8% from 17.2% for the same period. In the last 20 years, global production of natural gas has increased about 1.7 times and the US Energy Information Administration predicts its use to double by 2020 [1]. The percentage share of natural gas in marginal consumption of energy in Iran has dramatically increased in the two recent decades. Iran, a member of OPEC since 1961, is ranked amongst the world's top holders of proven oil (ranked fifth owning 8.6% of oil resources in the world) and natural gas reserves (ranked second owning 15% of natural gas resources in the world).

Because of large gas reserves in Iran and advantages of gas versus oil, the energy policy in this country is based on increase of gas/decrease of oil usage in commercial as well as residential sectors. At the first of 2007, National Iranian Gas Company (NIGC) with 440 million cubic meter refining capacity per day, had 25,000 km power transition lines and 131,320 km gas supplying network under its coverage and was providing natural gas to more than 11.6 million homes comprising of approximately 49.9 million people [2].

It should be noted that efficient use of energy resources require accurate prediction of future energy demand. Numerous researchers and practitioners have analyzed various energy issues and focused on developing appropriate energy demand models to reduce forecasting errors. Herbert [3] has analyzed monthly natural gas sales to the residential consumers in America [3]. Liu and Lin [4] have used time-series models to forecast residential natural gas consumption in Taiwan [4]. Eltony [5] has forecasted natural gas demand in Kuwait by means of econometric models [5]. Siemek [6] have estimated natural gas consumption in Poland based on the logistic curve interpretation [6]. Sarak and Satman [7] have described a deterministic model to forecast natural gas consumption for residential heating in certain areas of Turkey [7] based on previous studies performed by Durmayaz et al. [8]. Kaboudan and Liu [9] have forecasted quarterly US demand for natural gas in short term using combination of genetic programming with a two-stage least squares (2SLS) regression system of equations [9]. Aras and Aras [10] have forecasted natural gas demand for residential sector in Turkey using auto-regression approach [10].

In this paper, a logistic based approach is used to forecast the natural gas consumption for residential as well as commercial sectors in Iran. The forecast is for the 11th, 12th, and 13th year based

* Corresponding author. Tel.: +98 2164543329.
E-mail addresses: mehdi_f_s@yahoo.com (M. Forouzanfar), adoust20@gmail.com, doustm@aut.ac.ir (A. Doustmohammadi), tmenhaj@ieee.org (M.B. Menhaj), samira_hasanzadeh@modares.ac.ir (S. Hasanzadeh).

0306-2619/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apenergy.2009.07.008

on the existing data for the previous 10 years. The remainder of this paper is organized as follows. Historical reviews about sigmoid curves are presented in Section 2. Description of the logistic model is given in Section 3. In Section 4, the model is used to forecast the gas consumption in Iran. Conclusions and future research suggestions are given in Section 5.

2. Sigmoid curves and the logistic function

Sigmoid curve is a tilted *S*-shaped curve that resembles trends in the lifecycle of living things and phenomena. It is used in many different contexts such as demographics, T biology, economics, etc. In demographics and T biology contexts, sigmoid curves have been used to describe the evolution of population and in the economic context they have been used to model economic growth. The curves are used because of their ability to describe these processes and display their typical phases [11]. These curves can be considered as having a base phase, a growth phase (logarithmic growth), and a maturing phase. The maturing phase is basically the point of stabilization or zero growth rates and its value is usually referred to as the 'saturation value' (symbolized by *K*) or 'carrying capacity' of the environment under study. *K* represents the point at which the upward curve begins to level off. The *S*-shaped curve is usually summarized mathematically by the logistic equation (sigmoid curve is actually a special case of the logistic equation). A logistic equation is produced by a logistic function which is a mathematical function. The use of the logistic function to describe the population growth was first introduced by Belgian mathematician, Verhulst in the early 19th century [12]. Almost a century later, in 1920, Pearl and Reed rediscovered the logistic curve in the course of their biological study of the evolution of fly population [13]. In 1903, the French sociologist Gabriel Trade was perhaps the first to use a sigmoid curve in an economic field to analyze economic growth in relation to innovation. Other scholars have followed up Trade's ideas during the first half of the 20th century. Fisher and Pry [14] applied a logistic diffusion innovation model based on the analogy between epidemic spread and information circulation [14]. Similar research was also adopted by scholars such as Blackman [15]. Sigmoid curves have also been applied to model market demand analysis.

The logistic growth process is considered very appropriate for modeling economic growth curves. Mead and Islam [16] compare the forecasting performance of different growth curve models. They have applied 17 different models to forecast the development of telecommunication markets, represented by 25 time-series describing telephone access in 15 different countries. Their results indicate that the logistic model performs significantly better than any other growth model [16]. Sigmoid curves have been used in more recent research developments as well. Reati [17] uses both the logistic and the Gompertz curves to model the spread of technological revolutions [17]. Foster and Wild [18] show how the logistic equation can be used to model growth curves in presence of self organizational changes [18]. The logistic growth process has also been used to model economic relationships that had, for a long time, been modeled as linear forms. Honda and Suzuki [19] for instance, estimated an investment function by applying the logistic model. Their results strongly supported the use of a logistic model to estimate investment functions which appeared to follow nonlinear relationships rather than linear ones [19]. Aoki and Yoshikawa [20] have captured the role of demand as a limiting growth using logistic curves to describe the time utility of the new products [20]. Kejak et al. [21] has employed the logistic model to test the relationship between the education sector and human capital formation [21]. The logistic model has been indeed very effective in forecasting models of many different contexts however as a drawback, it has the characteristic of underestimating the forecasts in many situations [22]. NIGC has very useful information regarding daily natural gas consumption in Iran during different month of the year for different years [23]. In this article, using yearly and seasonal natural-gas consumption data in Iran for the period 1995–2005 obtained from NIGC main office in Tehran, we have proposed a logistic function that predicts or forecasts the yearly and seasonal gas consumption for the years 2006–2008.

3. The proposed procedure

The procedure is mainly based on logistic type function for estimation and a proper optimization technique like NLP or GA to minimize the error difference between the model and the actual data. The proposed function is of the following from:

$$\frac{dc(t)}{dt} = rc(t)\left(1 - \frac{c(t)}{K}\right) \tag{1}$$

where t is time, $c(t)$ is the natural gas consumption for any given time, and r and K are positive constants defined as the growth rate and the carrying capacity, respectively. We will later elaborate on why K is defined as carrying capacity. Using separation of variables it can be shown that the solution of the above logistic equation is given by

$$c(t) = \frac{K}{1 + Me^{-rt}} \tag{2}$$

where M is an arbitrary constant. To solve for M, let us denote the inflection point of $c(t)$ by t_0 and set the second derivative of $c(t)$ equal to zero at $t = t_0$. It can be shown that the second derivative of $c(t)$ with respect to t is given by:

$$c''(t) = \frac{MKr^2 e^{rt}(M - e^{rt})}{(M + e^{rt})^3} \tag{3}$$

Therefore, by setting $c''(t_0) = 0$, one will obtain

$$M = e^{rt_0} \tag{4}$$

Substitution of (4) into (2) leads to the following:

$$c(t) = \frac{K}{1 + e^{-r(t-t_0)}} \tag{5}$$

It should be noted from (5) that the consumption, $c(t)$, approaches K (saturation level) as $t \to \infty$. This is the reason why K is defined as carrying capacity. It should also be noted that by estimating the logistic parameters, t_0, r, and K, the consumption could be forecasted. Therefore, the objective is to estimate the above logistic parameters such that for a given set of n data points the error given by the following equation is minimized.

$$e = \sum_{i=1}^{n} (c(t_i) - c_i)^2 \tag{6}$$

In (6), the term c_i is the actual consumption at time t_i and the term $(c(t_i) - c_i)^2$ is the squared error made at the time instant t_i. The following constraints will be used to minimize (6):

$$\begin{aligned} K, t_0, r > 0 \quad &\text{(i)} \\ t_1 < t_0 < t_n \quad &\text{(ii)} \\ K > \max\{c_i\} \quad &\text{(iii)} \end{aligned} \tag{7}$$

The optimization technique used by Cavallini has the drawback that it may not yield a satisfactory result on the first try and consequently one may have to use different initial guesses to achieve an adequate result [24]. In this paper, we use two different procedures that will not have the above problem. The first procedure is based on NLP and the second one is based on GA. In both NLP and