A New Interpretation of the Logistic Model in Estimating Seasonal and Yearly Natural Gas Consumption

Ahmet Goncu[1,*]

[1]Dept. of Mathematical Sciences, Xian Jiaotong Liverpool University, Suzhou, China

*Corresponding author: Dept. of Mathematical Sciences, Xian Jiaotong Liverpool University, 215123 Jiangsu, Suzhou, China. Tel: 86-151-900-15627. E-mail: Ahmet.Goncu@xjtlu.edu.cn

Abstract

The logistic type model, which requires the use of an optimization routine for the estimation of model parameters, is one of a number of widely used methods in modeling natural gas consumption. In this study, we derive a regression model by considering the Euler discretization of the logistic type model and show that this regression based approach offers a simpler estimation procedure, in addition to better modeling natural gas consumption data. For comparison, the regression model derived from the logistic model is fitted to the same dataset published in Forouzanfar et *al.* (2010). Furthermore, in the regression approach, confidence intervals for point forecasts can be obtained, whereas the logistic function model is a deterministic function of time that does not provide confidence intervals.

Keywords: natural gas consumption; logistic function; non-linear programming; genetic algorithm; ordinary least squares

1. Introduction

Natural gas is an important energy source and occupies an increasing share of the world's total energy mix. With the increasing supply and demand of natural gas, it has become more important to model and estimate natural gas consumption accurately. Better models for natural gas consumption can help the natural gas industry in the management of energy sources. The modeling of natural gas consumption has drawn significant attention in literature. An excellent survey of the literature on natural gas consumption modeling is given by Soldo (2012). A variety of studies employ statistical or econometric models for the estimation of natural gas consumption, including examples by Liu and Lin (1991), Lee and Singh (1994), Aras and Aras (2004), Sanchez-Ubeda and Berzosa (2007), Vondracek et al. (2008), Yoo et al. (2009), and Erdogdu (2010).

An alternative approach in the literature on modeling natural gas consumption is the use of the logistic type model. Forouzanfar et al. (2010) proposed a logistic function based approach to model residential and commercial natural gas consumption. In this approach, model parameters are estimated via non-linear programming (NLP) and genetic algorithm (GA). Among others, logistic type models are employed in studies by Siemek et al. (2003), Gracias et al. (2012), and Melikoglu (2013).

For our study, we considered the Euler discretization of the logistic type model and derived a regression model that can be estimated via the least squares method. The regression model we derived from the logistic function approach has three advantages. First, the estimation of model parameters can be done via the simple least squares method. Second, confidence intervals for point forecasts can be obtained. Third, the regression model proposed provides a better fit for the considered dataset, which was originally published by Forouzanfar et al. (2010).

The next section describes the logistic function method and derives an alternative interpretation. In Section Three we provide the numerical results and our conclusions follow in Section Four.

2. Regression Interpretation of the Logistic Function Method

We started by considering the logistic function model, which was proposed by Forouzanfar et al. (2010) as

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

where t is time, c(t) is the natural gas consumption for any time t, and r and K are positive constants defined as the growth rate and the carrying capacity, respectively. The solution of the differential equation in Equation (1) is given by

$$c(t) = \frac{K}{1 + \exp(-r(t - t_0))},\qquad(2)$$

where t_0 is the time of the inflection point of $c(t)$. Equation (2) implies that as $t \to \infty$, consumption approaches the carrying capacity. The carrying capacity is assumed to be constant by Forouzanfar et al. (2010); however, this assumption may not be very realistic, as infrastructure investments can change the carrying capacity.

The parameters of the model in Equation (1) are estimated by minimizing the squared differences between the actual consumption and the theoretical consumption values. The objective function is given by

$$E = \sum_{i=1}^{n} (c(t_i) - c_i)^2,\qquad(3)$$

where c_i is the actual yearly (or seasonal) natural gas consumption for year i and $(c(t_i) - c_i)^2$ is the corresponding squared error term. The above minimization problem has the following constraints; $K, t_0, r > 0$, $t_1 < t_0 < t_n$, and $K > \max\{c_i\}$, for $i = 1,2,....,n$.

To solve the constrained minimization issue in Equation (3), Forouzanfar et al. (2010) use two alternative estimation procedures based on non-linear programming (NLP) and a genetic algorithm (GA).

Alternatively, we considered the Euler discretization of the differential equation given in Equation (1) and obtained

$$c(t_{i+1}) - c(t_i) = r\Delta t c(t_i) - \frac{r}{K}\Delta t c^2(t_i),\qquad(4)$$

where $t_{i+1} - t_i = \Delta t$ is on the uniform grid and thus Equation (4) is a first order approximation with an order of convergence $O(\Delta t)$.

Next, by adding a noise term to Equation (4), we reduced the original problem to a statistical curve fitting problem.

Suppose that the noise in the consumption process is independently and identically distributed with a mean of zero and constant variance σ^2. As such, we can re-write the discretized model in Equation (4) as

$$c(t_{i+1}) = \phi_1 c(t_i) + \phi_2 c^2(t_i) + \varepsilon(t_i) ,$$

$$(5)$$

where $\phi_1 = r\Delta t$, $\phi_2 = -\dfrac{r}{K}\Delta t$, and $\varepsilon(t_i) \sim (0, \sigma^2)$ are used for $i = 1, 2, ..., n-1$.

Equation (5) specifies a regression model in which today's realized consumption values are explained by yesterday's consumption values, $c(t_i)$ and $c^2(t_i)$. Therefore, the minimization problem in Equation (3) reduces the least squares (LS) minimization of the sum of the squared residuals

$$E = \sum_{i=1}^{n-1} \varepsilon^2(t_i) ,$$

$$(6)$$

where the above summation includes n-1 residuals because we need to use the first lag of natural gas consumption. Minimization of Equation (6) via the LS method does not require a distributional assumption of the noise term ε.

By including the possibility of a constant and a trend term, we propose explaining natural gas consumption in terms of the lagged consumption values with the following regression model

$$c_{t+1} = \phi_0 + \phi_1 t + \phi_2 c_t + \phi_3 c_t^2 + \varepsilon_t ,$$

$$(7)$$

where $\varepsilon_t \sim (0, \sigma^2)$. The tables and figures in Forouzanfar et al. (2010) are updated, including the results of the regression model proposed in Equation (7).

3. Numerical Results

We have used the dataset provided by Forouzanfar et al. (2010) for comparison between the logistic and regression approaches. The dataset contains 13 years of annual and seasonal natural gas consumption in the residential and commercial sectors of Iran.

Table 1: Comparison of the sum of squared errors under different models for residential and commercial natural gas consumption

Time Interval	Residential			Commercial		
	NLP	GA	LS	NLP	GA	LS
Yearly	1.445×10^4	1.446×10^4	1.109×10^4	99.38	99.45	62.05
Spring	2.631×10^3	2.403×10^3	1.632×10^3	5.66	6.00	3.59
Summer	1.256×10^3	1.256×10^3	1.281×10^3	3.94	3.94	3.09
Autumn	2.149×10^3	2.042×10^3	1.376×10^3	5.37	5.70	4.40
Winter	8.564×10^3	8.197×10^3	6.192×10^3	6.86	5.73	1.41

Two estimation procedures, the NLP and GA algorithms, are compared with the LS regression model derived in Equation (6) in terms of the sum of squared errors obtained from $n-1$ observations. In Table 1, we observe that the NLP and GA methods give similar results, whereas the LS regression gives a smaller sum of squared errors.
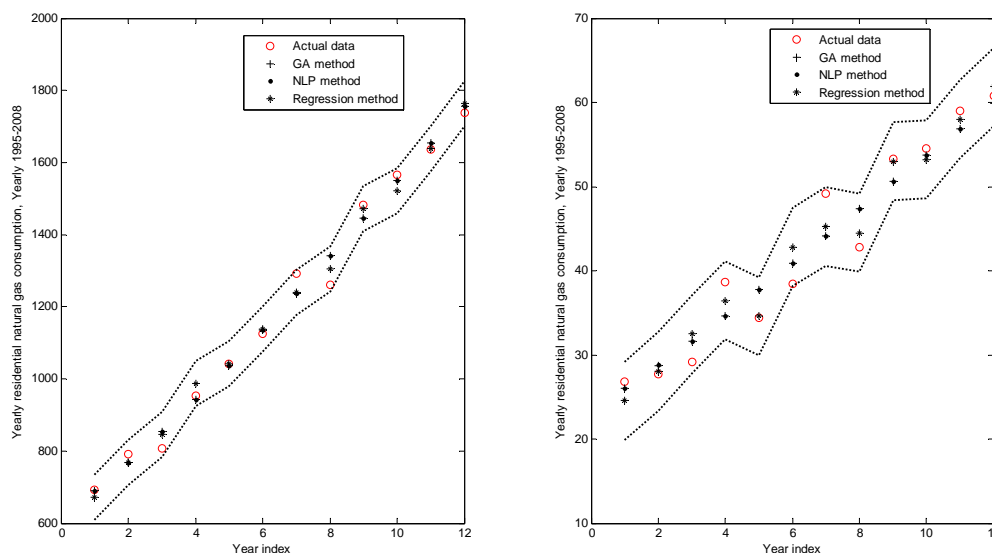

Figure 1: Yearly natural gas consumption versus year index from 3/21/1995 to 3/20/2008
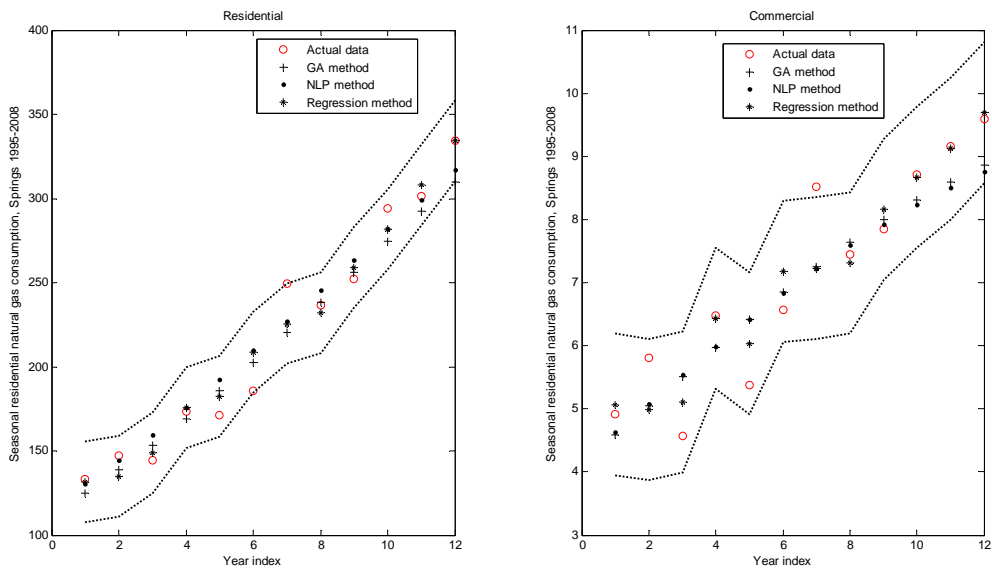
Figure 2: Spring natural gas consumption versus year index from 3/21/1995 to 3/20/2008
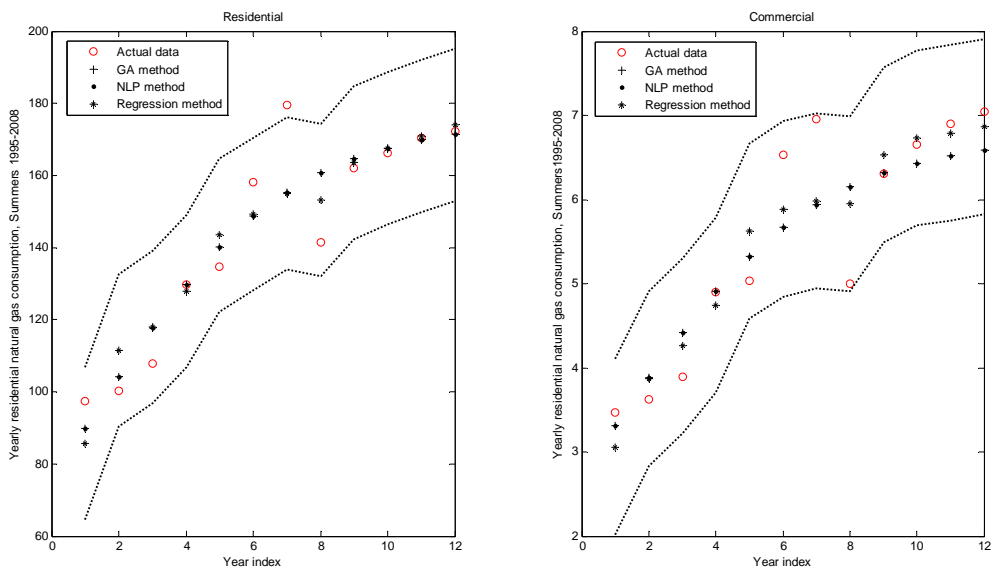

Figure 3: Summer natural gas consumption versus year index from 3/21/1995 to 3/20/2008
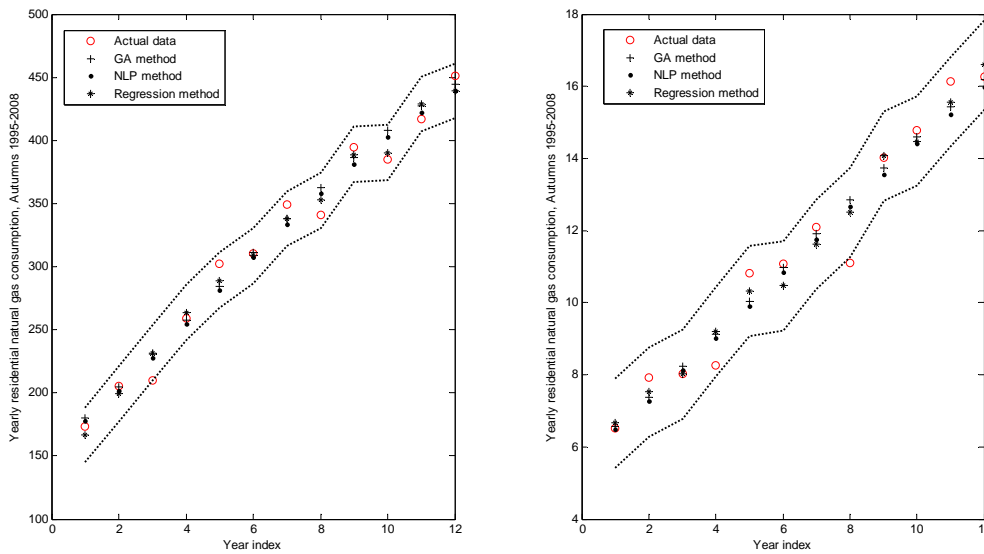
Figure 4: Autumn natural gas consumption versus year index from 3/21/1995 to 3/20/2008
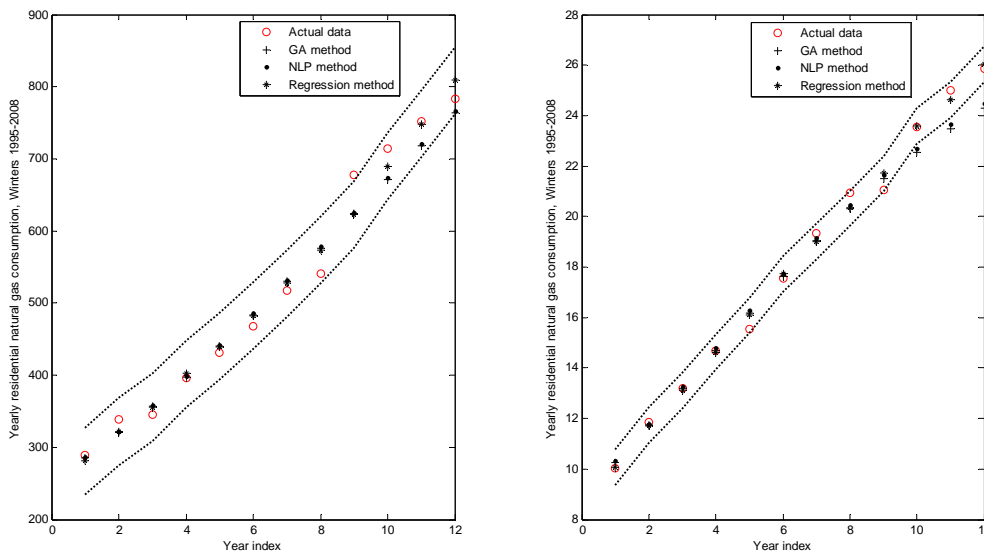

Figure 5: Winter natural gas consumption versus year index from 3/21/1995 to 3/20/2008

Furthermore, we updated the figures given by Forouzanfar et al. (2010) in Figures 1-5. We plot the fitted models using the NLP, GA, and LS methods for yearly and seasonal natural gas consumption in the residential and commercial sectors. For the regression model, point forecasts are plotted in Figures 1-5, including dotted lines for the confidence intervals ($\mu \pm 1.96 \times \sigma$); the logistic function approach cannot produce confidence intervals for point forecasts. Fitted values and the confidence intervals plotted in Figures 1-5 show that the regression model in Equation (7) fits the data better and the actual consumption values are within the given confidence intervals.

Table 2: Comparison of percentage forecast errors for **residential** natural gas consumption: NLP and GA methods versus the LS regression method

Time Interval	NLP Method			GA Method			LS Method		
	2006	2007	2008	2006	2007	2008	2006	2007	2008
Yearly	-0.96	1.17	1.05	-0.96	1.16	1.05	-2.78	1.10	3.29
Spring	-4.19	-0.65	-5.24	-6.60	-2.95	-7.26	-3.68	5.44	3.35
Summer	0.77	-0.28	-0.4	0.77	-0.27	-0.39	3.39	3.81	5.40
Autumn	4.63	1.27	-2.64	-5.96	-2.55	-1.41	0.58	2.56	-3.64
Winter	-5.75	-4.24	-2.25	-6.05	-4.54	-2.55	4.57	4.89	6.54

Table 3: Comparison of percentage forecast errors for **commercial** natural gas consumption: NLP and GA methods versus the LS regression method

Time Interval	NLP Method			GA Method			LS Method		
	2006	2007	2008	2006	2007	2008	2006	2007	2008
Yearly	-1.42	-3.59	-1.49	-1.42	-3.58	-1.48	-4.31	-3.58	-0.77
Spring	-5.41	-7.11	-8.76	-4.54	-6.11	-7.66	0.11	0.79	2.62
Summer	-3.32	-5.44	-6.56	-3.32	-5.44	-6.59	-3.46	-8.04	-11.06
Autumn	-2.52	-5.68	-1.76	-1.22	-4.41	-0.44	-13.92	-12.26	-10.39
Winter	-3.59	-5.38	-5.35	-4.24	-6.03	-6.01	-0.77	-2.57	-0.82

In Table 2, residential percent forecast errors using the NLP, GA, and LS regression methods are given for the years 2006, 2007, and 2008, and Table 3 presents the results for the commercial sector. Tables 2 and 3 show that the LS regression method gives smaller percentage forecast errors for the majority of the cases. However, we should note that a larger dataset is needed for a more reliable comparison of the forecasting performances for different models.

4. Conclusion

In this study we provide a new interpretation of the logistic function method proposed by Forouzanfar et al. (2010). We have derived a regression model from the logistic function method by considering the Euler discretization of the logistic function model. The logistic method of Forouzanfar et al. (2010) is estimated via non-linear programming and genetic algorithm procedures. The regression model derived in this study provides a better fit for natural gas consumption and can be estimated with the least squares method.

Furthermore, confidence intervals for the point forecasts are obtained using the regression model. With the existence of a larger dataset, the fit of the regression model can be improved and distributional assumptions on the noise term can be made. This will enable us to use more efficient estimation procedures, such as the maximum likelihood estimation, and more reliable statistical results can be obtained.

Different from the original logistic type model, the regression model derived provides the statistical confidence intervals for the natural gas forecasts. This implies that natural gas

www.macrothink.org/rae

distributors can use the forecasts and the associated confidence intervals for supply management. By improving the estimation and forecasting methodology of our model, better planning and management opportunities of natural gas companies can be made.

We believe that this study contributes to the existing literature on modeling natural gas consumption by proposing a new interpretation of the logistic function method, which provides an efficient and simple estimation procedure that fits the considered dataset well.

Acknowledgement

References

Aras, H., & Aras, N. (2004). Forecasting residential natural gas demand. *Energy Sources*, *26*, 463-472. http://dx.doi.org/10.1080/00908310490429740

Erdogdu, E. (2010). Natural gas demand in Turkey. *Applied Energy*, *87*, 211-219. http://dx.doi.org/10.1016/j.apenergy.2009.07.006

Forouzanfar, M., Doustmohammadi, A., Menhaj, M.B., & Hasanzadeh, S. (2010). Modeling and estimation of the natural gas consumption for residential and commercial sectors in Iran. *Applied Energy*, *87*, 268-274. http://dx.doi.org/10.1016/j.apenergy.2009.07.008

Gracias, A.C., Lourenço, S.R., & Rafikov, M. (2012). Estimation of Natural Gas Production, Import and Consumption in Brazil Based on Three Mathematical Models. *Natural Resources*, *3*, 42-47. http://dx.doi.org/10.4236/nr.2012.32007

Lee, R.S., & Singh, N. (1994). Patterns in residential gas and electricity demand: an econometric analysis. *Journal of Business Economics and Statistics*, *12*, 233-241.

Liu, L.M., & Lin, M.W. (1991). Forecasting residential consumption of natural gas using monthly and quarterly time series. *International Journal of Forecasting*, *7*, 3-16.

Melikoglu, M. (2013). Vision 2023: Forecasting Turkey's natural gas demand between 2013 and 2030. *Renewable and Sustainable Energy Reviews*, *22*, 393-400.

Sanchez-Ubeda, E., & Berzosa, A. (2007). Modeling and forecasting industrial end-use natural gas consumption. *Energy Economics*, *29*, 710-742.

Siemek, J., Nagy, S., & Rychlicki, S. (2003). Estimation of natural-gas consumption in Poland based on the logistic-curve interpretation. *Applied Energy*, *75*, 1-7. http://dx.doi.org/10.1016/S0306-2619(03)00013-8

Soldo, B. (2012). Forecasting Natural Gas Consumption. *Applied Energy*, *92*, 26-37. http://dx.doi.org/10.1016/j.apenergy.2011.11.003

Vondracek, J., Pelikan, E., Konar, O., Cermakova, J., Eben, K., & Maly M. (2008). A statistical model for the estimation of natural gas consumption. *Applied Energy*, *85,* 362-370.

Yoo, S.H., Lim, H.J., & Kwak, S.J. (2009). Estimating the residential demand function for natural gas in Seoul with correction for sample selection bias. *Applied Energy*, *86*, 460-465. http://dx.doi.org/10.1016/j.apenergy.2008.08.023

Copyright Disclaimer