KKU Engineering Journal

https://www.tci-thaijo.org/index.php/kkuenj/index

The factors affecting on consumption of some petroleum product in Thailand using general linear model and generalized linear model

Ampika Konkaew, Sujitta Suraphee and Rojanee Homchalee*

Research Unit on Applied Statistics, Department of Mathematics, Faculty of Science, Mahasarakham University, Maha Sarakham 44150, Thailand.

Received April 2016
Accepted June 2016

Abstract

The objective of this research was to investigate forecasting consumption of some petroleum products in Thailand. Two methods of statistical modeling, general linear models (GLM) and generalized linear model (GLMs) were applied. The GLM compared with GLMs in which response variables are the consumption of gasohol 95, gasohol 91, gasohol E20, and gasohol E85. The explanatory variables were gasoline 95 price, gasohol 91 price, gasohol 95 price, LPG price, NGV price, ethanol price, exchange rates, and world crude oil prices. The secondary data from many resources in the time period Jan 2010 to Dec 2015 were collected for analysis. The factor analysis was applied to reduce some-covariate variables that are high relationship. The study found that the GLMs with gamma distribution was more appropriate than GLM when response variable is not normality. In case of the response variable is normally distributed both of them give well the same result.

Keywords: Alternative energy products, Consumption of petroleum product, General linear models, Generalized linear models, Gamma distribution, Factor analysis

1. Introduction

Currently, petroleum energy has important role to the global economy motivating for the development to move forward on all fronts. The realization of the extinction of fuel processing from natural energy (crude oil) is very important. Therefore, Thai government policy about renewable energy development plans (REDP) for 15 year (2008-2022) has been set [1].

The government is supporting the production and use of renewable energy to reduce dependence on energy import. The renewable energy that blends ethanol in gasoline is environmental friendly. Thus, an ethanol product which is an alternative fuel that has been identified in the development plan of renewable energy is very important. The study about renewable energy sources such as the production of ethanol from the sugarcane and cassava helping agricultural products has been increased.

Since the data of consumption of all response variable is right tail (positively skewed distribution). It is not normality. The GLMs formula with Gamma distribution was applied. Therefore, in this paper the factors affecting consumption of gasohol 95, gasohol 91, gasohol E20, and gasohol E85 using two statistical methods, classical general linear models (GLM) and generalized linear model (GLMs) are investigated. The results from both techniques will be compared to choose more appropriate model.

2. Materials and methods

To investigate the factor affecting consumption of gasohol 95, gasohol 91, gasohol E20, and gasohol E85, we propose two statistical modeling, GLM and GLMs [2], as the following:

2.1 General linear models (GLM)

General linear model or linear model conceptualizes *Y* as the sum of its mean, μ, and a random variable, ε:

$$Y = \mu + \varepsilon \quad ; \mu = E(Y) \tag{1}$$

It is assumed that the expected value of *Y, E(Y)* can be written as a linear combination of the covariates, *X*, that is $\mu = \alpha + \beta_1 x_1 + \ldots + \beta_k x_k$. The error term, ε, is normally distributed with mean zero and constant variance. If error term is normally distributed, response variable is also normality. Linear models can be easily to solve with well-known linear algebra approaches. However, the required assumptions are violated in many applications.

2.2 Generalized linear models (GLMs)

Generalized linear models extend the linear modelling framework to response variables that are not normally

*Corresponding author. Tel.: +66 4375 4244, +6681 941 4720
Email address: rojaneeh@hotmail.com
doi: 10.14456/kkuenj.2016.80

distributed and not constant variance. The variance is permitted to vary with mean of the distribution. That is, the variance can be a function of the mean. This situation should be applied by GLMs to construct suitable statistical modeling. GLMs consist of a wide range of models that include linear models as a special case. The response variable is assumed to be a member of the exponential family of distributions. Finally, the effect of the covariates on the response variable is assumed to be additive on a transformed scale.

2.2.1 Components of a generalized linear model

- *Random component:* each component of Y is independent and is from one of the exponential family of distributions [3].
- *Systematic component:* the p covariates are combined to give the linear predictor.
- *Link function:* the relationship between the random and systematic components is specified via a canonical link function, g, which is differentiable and monotonic. The model formula states that $g(\mu) = \alpha + \beta_1 x_1 + \ldots + \beta_k x_k$

2.2.2 The GLMs formula with Gamma distribution

Since the response variables in this study are consumption of gasohol 91, gasohol 95, gasohol E20, and gasohol E85, which are non-negative values and not normality, therefore, we apply GLMs with Gamma distribution to construct the appropriated model. The canonical link function is $g(\mu) = \mu^{-1}$, which is the inverse link:

$$\mu^{-1} = \alpha + \beta_1 x_1 + \ldots + \beta_k x_k$$

2.3 Factor analysis

Factor analysis is a method of variable reduction. In this study, it was applied to reduce covariate variables that high relationship before take them to forecast the response variables. The eigenvalue is a measure to explain the total variance of all observed variables. The relationship of each variable to the underlying factor is expressed by the factor loading [4].

2.4 Criteria for assessing the goodness of fit of different models

The main criteria used to assess the goodness of fit of models is mean absolute percent error (MAPE) , as shown in Table 1. A smaller MAPE model is preferred to the other models. The mathematical form of MAPE is given by

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{Y}_i - Y_i}{Y_i}\right| \times 100 ,$$

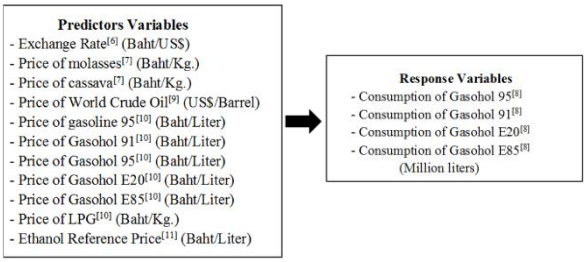
where Y_i and $\hat{Y}_i$ are the actual and predicted value of response variable, respectively.

Table 1 Typical MAPE Values for Model Evaluation

MAPE (%)	Evaluation
MAPE ≤ 10%	High Accuracy Forecasting
10% < MAPE ≤ 20%	Good Forecasting
20% < MAPE ≤ 50%	Reasonable Forecasting
MAPE > 50%	Inaccurate Forecasting

Source: Lewis, C. D. (1982) [5]. International and Business Forecasting Methods. London: Butterworths.

2.5 Response and explanatory variables



3. Results

For reducing the explanatory variables by factor analysis technique, the factor that affect to consumption of gasohol 95 and gasohol 91 can be grouped the three dimensions, the group of energy price (factor1), the price of raw material of ethanol product and price of LPG (factor2), and price of ethanol (factor3). All of the components account for a substantial percentage 85.669% of the total variance in the transformed variables. The forecasting equation from both general linear model and generalized linear model showed in the Table 2. Both models obtained MAPE between 10% and 20% that were good forecasting models.

There were two factors affecting to consumption of gasohol E20, energy price (factor1) and the price of raw material of ethanol product (factor2). There were three factors affecting consumption of gasohol E85, the price of raw material of ethanol product (factor1), economic factor (factor2), and energy price (factor3). Explaining about the eigenvalue which is a measure of how much of the variance of the observed variables a factor explains can be explain same as the consumption of gasohol 95 and gasohol 91.

Moreover, forecasting equation by the general linear model and generalized linear model showed in the Table 3 and 4. However, MAPE of both models was greater than 50%, which were inaccurate forecasting models.

4. Discussion and conclusions

From the results, generalized linear models give the better result than general linear model. Since the consumption of petroleum products are the right tailed distribution. We found that the distribution of consumption of gasohol is Gamma distribution. Therefore, classical general linear model is not suitable for applying for this situation. For the forecasting models of consumption of gasohol E20 and gasohol E85 give unsatisfying result both GLM and GLMs. It may the predictor variables which are chosen to predict the consumption of both of them is not appropriated. They may depend on others factors.

Table 2 The factors affecting consumption of Gasohol 95 and Gasohol 91 and forecasting equations.

-Consumption of Gasohol 95	Factor1	Factor2	Factor3	
	- exchange rate	- price of molasses	- price of ethanol	
	- price of world crude oil	- price of cassava		
-Consumption of Gasohol 91	- price of gasohol 95	- price of LPG		
	- price of gasohol 91			
	- price of gasoline 95			
Eigenvalue	4.555	2.108	1.048	
% of variance	50.611	23.417	11.641	
Forecasting equation				MAPE
-GLM for Gasohol95	$consumption\ of\ gasohol\ 95 = 219.309 + 26.597(F3) - 18.951(F1).$			10.76%
-GLMs for Gasohol95	$consumption\ of\ gasohol\ 95 = \frac{1}{0.005\ +\ 0.000(F1) - 0.001(F3)}.$			11.32%
-GLM for Gasohol91	$consumption\ of\ gasohol\ 91 = 228.94 + 52.440(F3) - 45.383(F2) - 22.144(F1).$			13.95%
-GLMs for Gasohol91	$consumption\ of\ gasohol\ 91 = \frac{1}{0.005\ +\ 0.000(F1) + 0.001(F2) - 0.001(F3)}.$			13.54%

Table 3 The factors affecting consumption of Gasohol E20 and forecasting equations.

-Consumption of Gasohol E20	Factor 1	Factor 2	
	- exchange rate	- price of molasses	
	- price of world crude oil	- price of cassava	
	- price of ethanol		
	- price of gasohol E20		
Eigenvalue	2.677	1.475	
% of variance	44.614	24.584	
Forecasting equation			MAPE
-GLM for Gasohol E20	$consumption\ of\ gasohol\ E20 = 63.100\ - 27.899(F1) - 19.872(F2).$		105.43%
-GLMs for Gasohol E20	$consumption\ of\ gasohol\ E20 = \frac{1}{0.022\ +\ 0.006(F1) + 0.010(F2)}.$		88.64%

Table 4 The factors affecting consumption of Gasohol E85 and forecasting equations.

-Consumption of Gasohol E85	Factor 1	Factor 2	Factor 3	
	- price of molasses	- exchange rate	- price of ethanol	
	- price of cassava	- price of world crude oil	- price of gasohol E85	
Eigenvalue	2.073	1.821	1.105	
% of variance	34.556	30.349	18.423	
Forecasting equation				MAPE
-GLM for Gasohol E85	$consumption\ of\ gasohol\ E85 = 11.653 + 7.646(F3) - 6.286(F2) - 4.577(F1).$			666.84%
-GLMs for Gasohol E85	$consumption\ of\ gasohol\ E85 = \frac{1}{0.185\ +\ 0.072(F1) + 0.046(F2) - 0.071(F3)}.$			480.38%

5. Acknowledgements

The authors would like to thank the associate editor and the referees.

6. References

[1] Energy Policy and Planning Office. Energy data [Internet]. 2016 [cited 2016 Mar 27]. Available from: http://www.eppo.go.th/ccep/energy_3-5.html
[2] Anderson D, Feldblum S, Modlin C, Schirmacher D, Schirmacher E, Thandi N. A Practitioner's Guide to Generalized Linear Models. 3th ed. Arlington: Towers Watson; 2007.
[3] Nelder JA, Wedderburn RWM. Generalized Linear Models. Reviewed work(s): Journal of the Royal Statistical Society, Series A (General) 1972;135(3) :370-384.
[4] Kanlaya V. Multivariate Analysis by SPSS for Windows: Factor Analysis. Bangkok: Chulalongkorn University Printing House; 2001.
[5] Lewis CD. International and Business Forecasting Methods. London: Butterworths; 1982.
[6] Bot.or.th [Internet]. Thailand: Bank of Thailand. Available from: https://www.bot.or.th/Thai/Statistics /FinancialMarkets/ExchangeRate/Pages/StatExchange Rate.aspx
[7] Customs.go.th [Internet]. Thailand: The Customs Department. Available from: http://internet1.customs. go.th/ext/ Statistic/StatisticIndex2550.jsp
[8] Doeb.go.th [Internet]. Thailand: Department of Energy Business ministry of energy. Available from: http://www.doeb.go.th/v5/info_sum_import_export_f uel.php
[9] Eia.gov [Internet]. United Stated: U.S. Energy Information Administration. Available from: http://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm
[10] Eppo.go.th [Internet]. Thailand: Energy Policy and Planning Office. Available from: http://www. eppo.go.th/info/8prices_stat.htm
[11] Thaiethanol.com [Internet]. Thailand: Thai Ethanol Manufacturing Association. Available from: http:// www.thaiethanol.com/th/statistical-data/price.html