International Journal of Natural and Engineering Sciences 7 (3): 07-15, 2013
ISSN: 1307-1149, E-ISSN: 2146-0086, www.nobel.gen.tr

Application of Genetic programming to predict an SI engine brake power and torque using ethanol- gasoline fuel blends

Mostafa Kiani DEH KIANI Barat GHOBADIAN* Fathollah OMMI Gholamhassan NAJAFI*
Tarbiat Modares University, Tehran- Iran P.O.Box:14115-111

*Corresponding author
ghobadib@modares.ac.ir

Received: September 12, 2013
Accepted: October 27, 2013

Abstract

The main objective of this study is to predict the performance of a commercial spark ignition (SI) engine using multigene genetic programming (GP). To acquire data for training and testing of the proposed GP, a four-cylinder, four-stroke test engine was fueled with ethanol-gasoline fuel blends. The fuels were blended with various percentages of ethanol (0, 5, 10, 15 and 20%), and the engine was operated at different engine speeds and loads. The experimental results showed that using ethanol–gasoline blend fuels increased the brake power and torque of the engine. Numerous runs were performed with model of GP and the performance of developed equations was evaluated. The optimum models were selected according to statistical criteria of root mean square error (*RMSE*) and coefficient of determination (R^2). The values of *RMSE* and R^2 for brake power were found to be 0.388 and 0.998. It was observed that the GP model can predict engine torque with correlation coefficient (R^2) in the range of 0.99–1 and *RMSE* was found to be 0.731. The simulation results demonstrated that GP model is a good tool to predict the engine brake power and torque under test.

Keywords: SI engine: ethanol-gasoline blends: multigene genetic programming.

Abbreviations

EC	Evolutionary computation	*GP*	Genetic programming
GA	Genetic Algorithm	*RMSE*	Root mean square error
t	Experimental value	o	Predicted value
n	Total number of data	R^2	Coefficient of determination
VE	Variation explained	P_b	Engine brake power
T	Engine torque	x_1	Engine speed
x_2	Engine load	x_3	Percentage of ethanol

INTRODUCTION

Evolutionary computation (EC) is drawing attentions for solving real engineering problems. This approach is to be robust in delivering global optimal solutions and coping with the restrictions encountered in traditional methods. EC harnesses the power of natural selection to turn computers into optimization tools [1-4]. This is very applicable to different problems in the manufacturing industry [5-8]. One of most important EC methods is genetic programming (GP). GP is a similar technique as genetic algorithm, an evolutionary computation method for imitating biological evolution of living organisms.

Genetic Algorithms (GAs) and genetic programming (GP) have been found to offer advantages dealing with system modeling and optimization, especially for complex and nonlinear systems. GP has been applied to a wide range of problems in artificial intelligence, engineering and science, chemical and biological processes and mechanical issues [9-13].

Pires, Alvim-Ferraz, Pereira and Martins [14] used GP method to predict the next day hourly average tropospheric ozone (O3) concentrations. The results showed very good agreement between predicted and measured data. Prediction of compressive and tensile strength of limestone was carried out via genetic programming as reported by Baykasoglu, Gullu, Canak and Ozbakir [15].

Another interesting Genetic Programming application was conducted by Cevik and Cabalar [16] for prediction of peak ground acceleration (PGA) using strong-ground-motion data. In this research, they demonstrated a high correlation between PGA and predictions. Multigene genetic programming is a recently developed approach for improving accuracy of GP that was developed by Hinchliffe, Willis, Hiden, Tham, McKay and Barton [17] and Hiden [18] and have been utilized in some recent research works [19-20].

The aim of this paper is to use a multigene GP algorithm based mathematical model for predicting an SI engine brake power and torque in relation to input variables including engine speed, engine load and ethanol-gasoline fuel blends.

MATERIALS AND METHODS

Data collection procedure

A KIA 1.3 SOHC, four-cylinders, four-stroke, spark ignition (SI) engine was used in this study. Technical specifications of the test engine are shown in Table 1. To measure engine brake power and torque, a 190 kW SCHENCK-WT190 eddy-current dynamometer was used to perform the experiments. Five separate fuel tanks were arranged containing gasoline and the ethanol-gasoline blends. The test setup for performing experiments is demonstrated in Fig. 1. The brake power and torque of the test engine were measured with different blends of ethanol-gasoline (E0, E5, E10, E15 and E20). Each test of fuel blends was performed under varying engine speed (1000-5000 rpm with 500 rpm interval) and load conditions (25, 50, 70 and 100%). Properties of gasoline and ethanol used in this study have been presented in Table 2.

Table 1. Technical specifications of the test engine.

Engine type	SOHC, fuel injected
Number of cylinder	4
Compression ratio	9.7
Bore (mm)	71
Stroke (mm)	83.6
Displacement volume (cc)	1323
Max. power (kW)	64 at 5200 rpm
Cooling system	Water-cooled

Table 2. Properties of ethanol and gasoline.

Fuel property	Ethanol	Gasoline
Formula	C2H5OH	C4 to C12
Molecular weight	46.07	100-105
Density, g/cm3 at 20 ºC	0.79	0.74
Lower heating value, MJ/kg	25.12	45.26
Stoichiometric air- fuel ratio, Wight	9	14.7
Specific heat, kJ/kg K	2.38	1.99
Latent heat of vaporize, kJ/kg	839	305

Genetic programming concept

Genetic programming (GP) is a sub-branch of evolutionary algorithms (EAs) emulating the natural evolution of species. Koza [21] was one of the scientists who first suggested the use of GP to find a symbolic regression tree matching to the mathematical formula which can best fit the data according to a fitness criterion. Fitting such a model was performed in an optimization frame work in which the error (e.g., rmse) of the created symbolic trees versus sample data is minimized via regression. Thus; it is intrinsically suitable for modeling of complex industrial problems. In order to emulate the evolutionary process in the design of GP, certain components should be defined. These include n-ary arithmetic functions, problem decision variables and evolutionary operators such as reproduction, crossover, and mutation to symbolic expressions. The symbolic expressions, called individuals or solutions, are generated to create the initial population. A population in evolutionary algorithms is a set of a defined number of solutions at an iteration of the

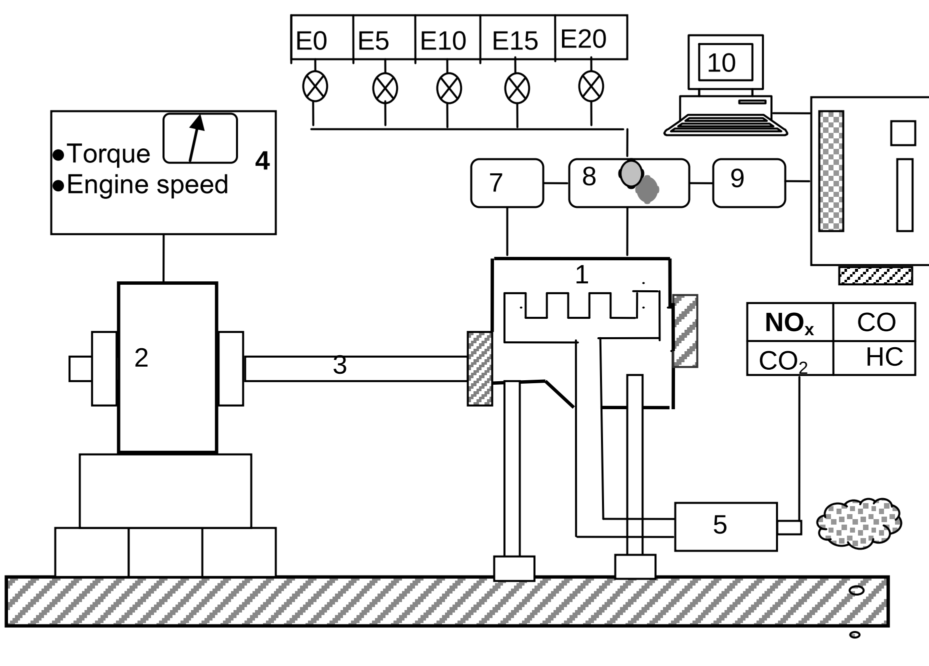

Fig.1. Schematic diagram of experimental setup. (1. Engine; 2.Dynamometer; 3.Drive shaft; 4.Dynamometer control unit, load & speed indicator; 5.Exhaust; 6.Gas analyzer; 7. Air flow meter; 8.Fuel measurement system; 9.Measuring boom; 10.Computer)

algorithm. The initial expressions are produced with tree-based encoding. These expressions are constituted of elements from two distinctive parameter groups: (i) a functional set and (ii) a terminal set. The functional set is generally arithmetic function, e.g. f= {*, +, -, sin, cos, log, power …}. The arguments for these functions are supplied from the terminal set that includes the decision variables and constants. The initial solutions are restricted in terms of tree depth or length of expression to fill the first population in the algorithm with the potential building blocks of individuals to be created at the next step of the algorithm [21.

At each generation a new population is created through selecting individuals based on their fitness and using the genetic operators (reproduction, crossover, and mutation). In reproduction operation, part of population (the fittest individuals) is preserved so that new generation is the result of genetic operations on the individuals of the actual population. In crossover operation two individuals (parents) are selected, their tree structures are broken at a randomly selected crossover point, and the produced sub-trees are recombined to form two new individuals (offspring) [22]. The existing population will then be substituted with the new population. The procedure is iterated until a termination criterion (achievement of the maximum number of generations or a determined error defined) is satisfied.

A newly developed method for improving the precision of GP is ‘‘multigene genetic programming’’. The main difference between the traditional and the multigene GP is the number of trees which can be used. In the traditional GP, a single tree represents the model however, in the multigene GP several trees may express the model. All of these genes possess specific optimal weights and sum of weighted genes plus a bias term would form the final formula as the best resulted numerical model. Multigene GP can be shown as the follow:

$$Y = a_0 + a_1 \times gene_1 + a_2 \times gene_2 + ... + a_n \times gene_n \quad (1)$$

In which, a0 is the bias term and ai is weight of the ith gene. Indeed, multigene GP is a linear combination of nonlinear terms, and this feature allows identifying the model of engineering problems in a highly precise manner.

Design of multigene GP

The GP was used in this study to perform a multigene genetic programming for precise prediction of engine brake power and torque. The Genetic Programming & Symbolic Regression is a new code written on the basis of multigene GP for use with MATLAB [23].

The GP has the possibility of setting some limitations to avoid bloating. Bloating is defined as the unnecessary growth of the model without any significant improvement in the fitness. In order to avoid bloating, some restrictions were imposed on initial parameters such as maximum number of genes, maximum depth of genes and trees, and maximum number of nodes per tree. In addition, lexicographic tournament selection that is an efficient method for restraining the model bloating was used in GP. It is noteworthy that the present investigation has considered root mean square error (RMSE) as the fitness function of the analysis. The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(t_i - o_i)^2} \quad (2)$$

Where t is the experimental value, o is the predicted value and n is the total number of data. Table 3 indicates the range of the initial parameters used in the GP runs. Other initial parameters were adjusted to their default values in GPTIPS, according to Searson [24].

Table 3. Multigene GP ranges of initially parameters.

parameter	Range
Number of generations	100-300
Population size	100-400
Function set	{+, -, ×,√ ÷, sin, cos, exp, log, tanh}
Maximum number of genes	3-5
Maximum depth of tree	4-8
Probability of crossover	0.70-0.95
Probability of mutation	0.04-0.2
Probability of reproduction	0.01-0.1

Three statistical evaluation criteria were applied to assess the model performance. (i) The coefficient of determination (R2) defined as:

$$R^2 = \left(1 - \frac{\sum_{i=1}^{n}(t_i - o_i)^2}{\sum_{i=1}^{n}(t_i - \bar{t}_i)^2}\right) \quad (3)$$

(ii) Variation explained (VE) defined as:

$$VE = 100R^2 \quad (4)$$

(iii) The root mean square error (RMSE) defined as Eq. (2).

For developing the GP model, the data set was divided into two subsets including the training and testing with a ratio of 0.75 and 0.25, respectively (i.e., 135 samples for training and 48 samples for testing).

RESULTS AND DISCUSSIONS

Experiment results

The effects of the ethanol addition to gasoline on the engine brake power and torque between 1000 and 5000 rpm engine speed at different loads (25, 50, 75% and full load) are shown in Fig. 2 and Fig.3 respectively. In general, increase of ethanol content increases the brake power and torque of the engine. Although the addition of ethanol to the gasoline decreases its heating value, the increase in torque and power was observed. Useful effect of ethanol as an oxygenated fuel is a possible reason for more complete combustion, and thus increasing the brake power and torque. Another possible reason is that the latent heat of evaporation of blended fuels is

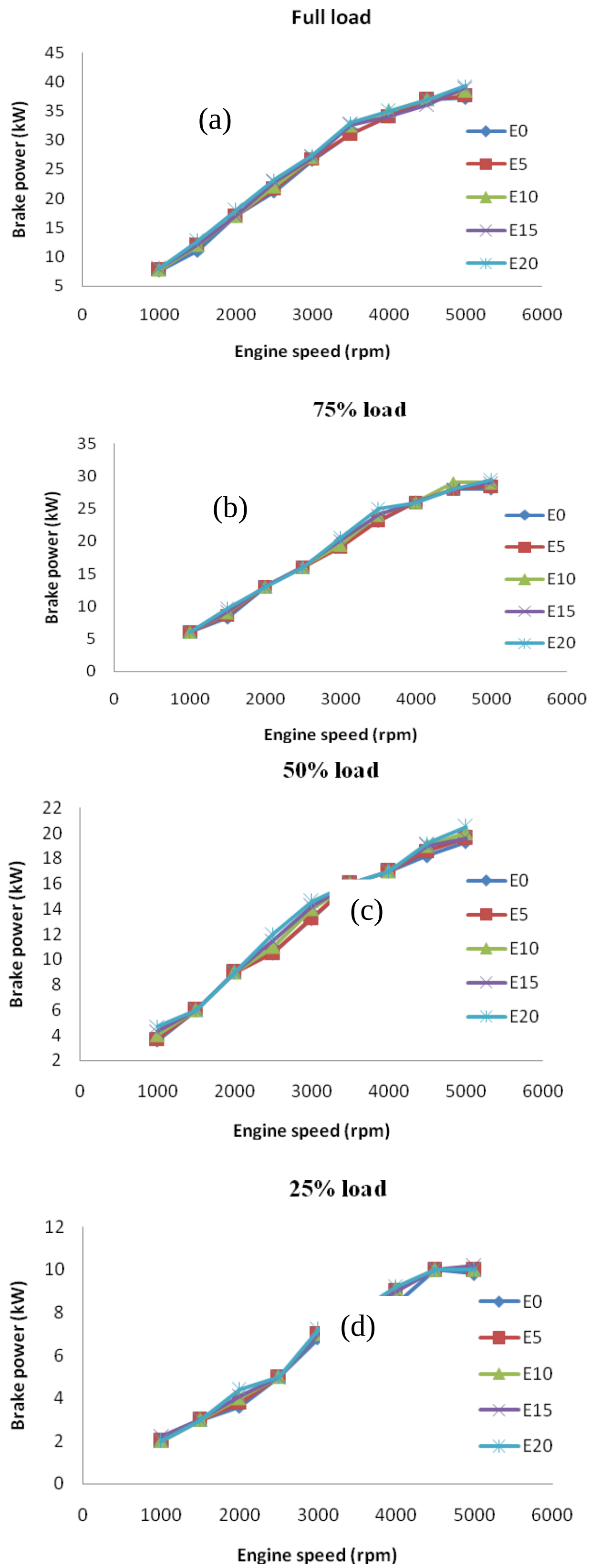

Fig.2. Experimental results of brake power at different ethanol-gasoline blends and engine speeds for: (a) full load, (b) 75% load, (c) 50% load, (d) 25% load.

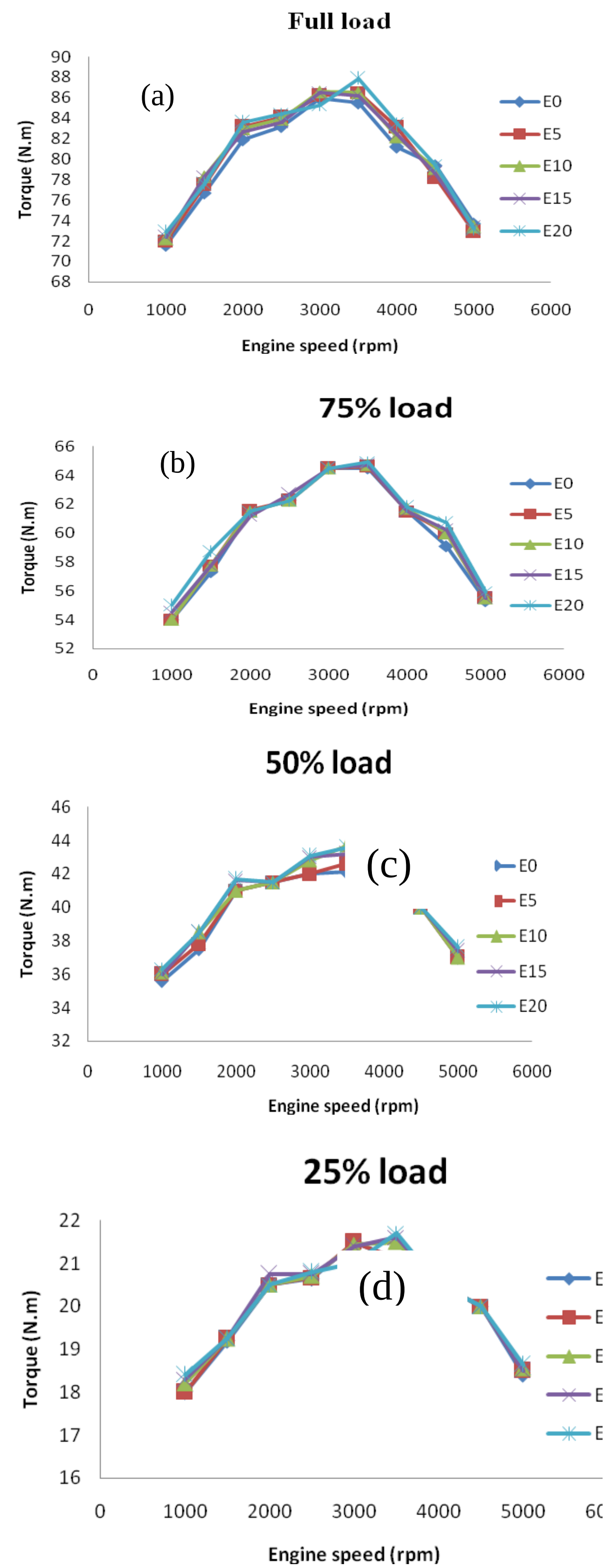

Fig.3. Experimental results of torque at different ethanol-gasoline blends and engine speeds for: (a) full load, (b) 75% load, (c) 50% load, (d) 25% load.

higher than that of gasoline; this provides lower temperature intake manifold and increases volumetric efficiency (Table 2). So, the brake power and torque increase with increasing percentage of ethanol in fuel blends.

Several runs were carried out with different initial parameters of multigene GP and the performance of developing equations was assessed for each run. Finally the best models were selected according to the performance evaluation criteria (Sec. 2.3) for prediction of engine brake power and torque.

Multigene GP model for brake power

The following equation was selected as the best model for brake power.

$$P_b = 0.3351x_1 + 0.4875x_2 + 0.3925\cos(x_2 - x_1) + 0.3925\cos(\cos(x_2)) + 0.2031\tanh(x_1 + x_2 - 0.8742) + 0.3925\cos(\cos(x_1 - x_2 - x_3)) - 0.2742\tanh(x_1 - x_2 - 0.6396) + 0.4773\sin(x_1) - 0.7268 \quad (5)$$

Where Pb is the engine brake power (kW), and x1, x2 and x3 are the engine speed (rpm), engine load (%) and percentage of ethanol in fuel blends respectively.

Accuracy of the equation is studied by plotting the measured against predicted values for training and testing sets (Fig.4). The values of R2 and RMSE are equal to 0.9985 and 0.3382, respectively, for training sets (fig.4a) and 0.9955 and 0.6596, respectively, for testing sets (fig.4b). There is a good correlation between the predictions from multigene GP and the measured data.

A comparison of the error during training and testing by using multigene GP and experimental results are illustrated in Fig.5. As can be observed, the training sets include the results of 135 samples and the testing sets include the results of 45 samples. It was seen that the multigene GP model can predict engine brake power with a high variation explained (99.85% for training and 99.55% for testing) and low root mean square errors (RMSE).

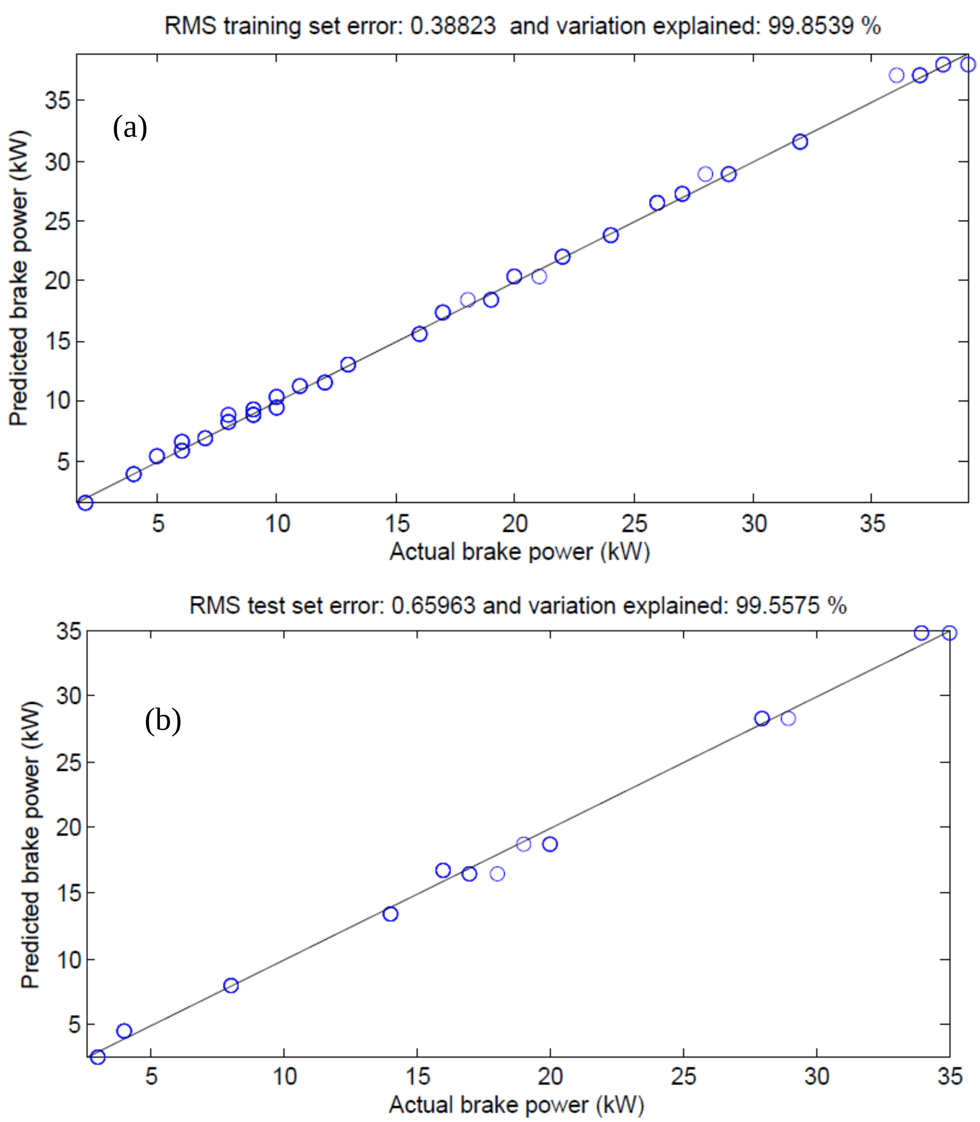

Fig.4. Measured versuse prdicted values of brake power for, (a) training set data and (b) testining data set.

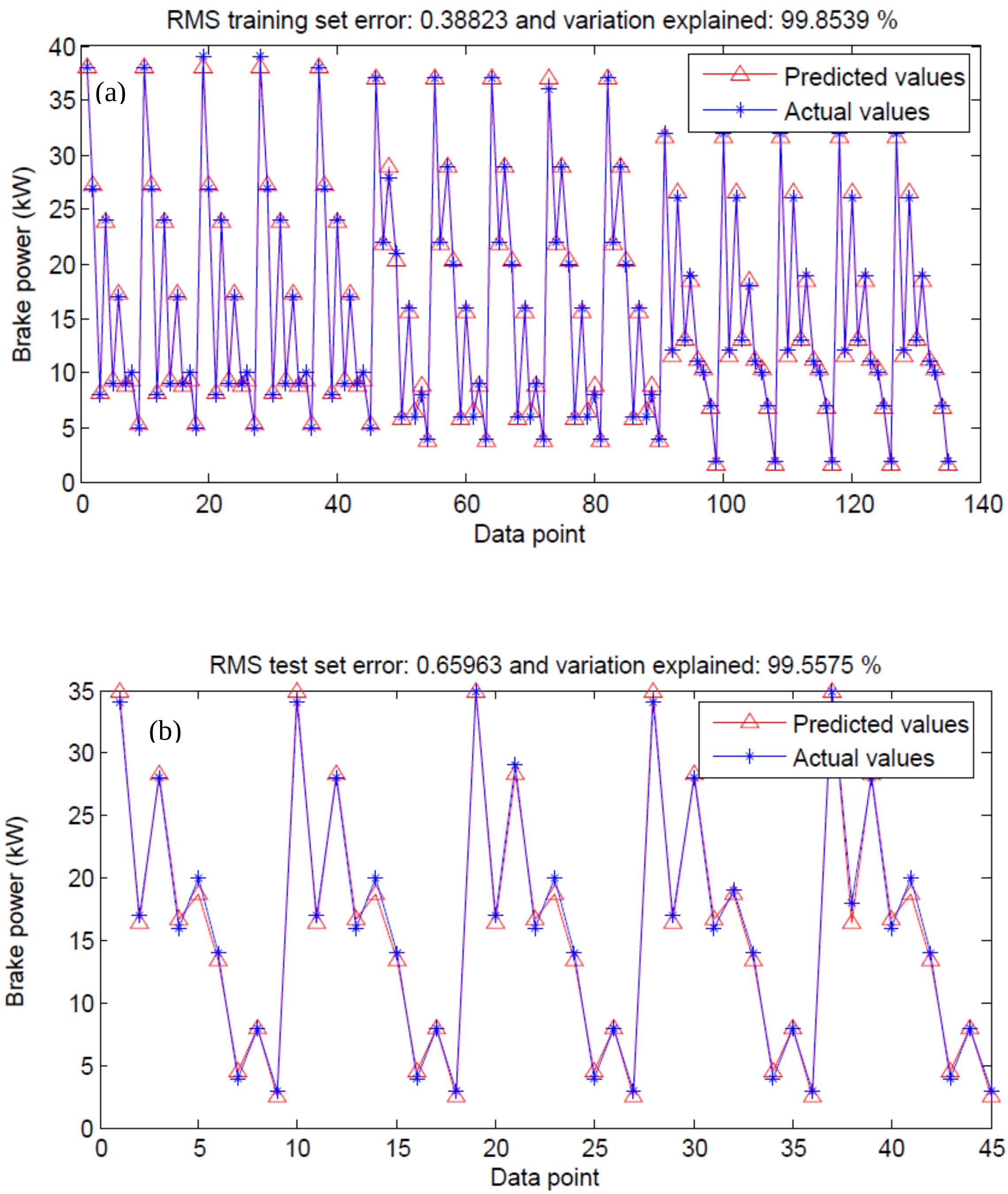

Fig.5. Comparisons of experimental results and the GP model predictions of brake power for, (a) training set data and (b) testining data set.

Multigene GP model for torque

The following equation was selected as the best model for prediction of engine torque.

$$T = 2.461x_2 + 0.2231\log(x_1 - x_2 + 11.64) - 3.353\log(x_2 + \cos(x_1) + 5.808)$$
$$+9.756\log(\cos(x_1 - x_3) - x_2 + 10.25) - 17.57 \quad (6)$$

Where T is the engine torque (N.m), and x1, x2 and x3 are engine speed (rpm), engine load (%) and percentage of ethanol in blend fuels respectively.

Precision of the developed equation is examined by plotting the measured against predicted values for training and testing sets (Fig.6). The values of R2 and RMSE are equal to 0.9994 and 0.5245, respectively, for training sets (fig.6a) and 0.9989 and 0.7313, respectively, for testing sets (fig.6b). There is a good correlation between the predictions from multigene GP and the measured data.

The error between predicted values by using multigene GP and experimental results of training and testing data are illustrated in Fig.7. As can be observed, the training sets include the results of 45 samples and testing sets include of 145 samples. It was seen that the multigene GP model can predict engine torque with variation explained 99.94% and 99.89% for training and testing data.

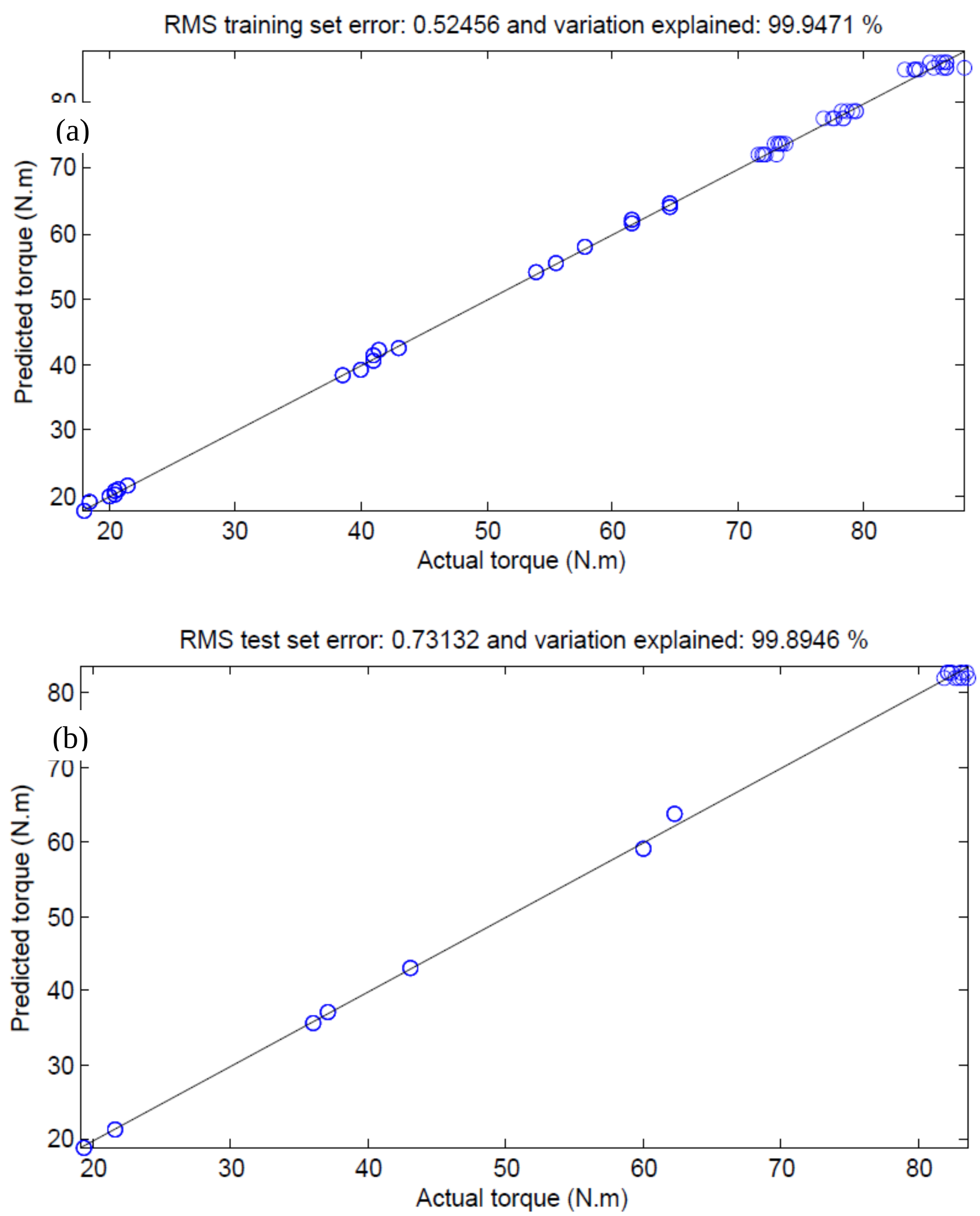

Fig. 6. Measured versuse prdicted values of torque for, (a) training set data and (b) testining data set.

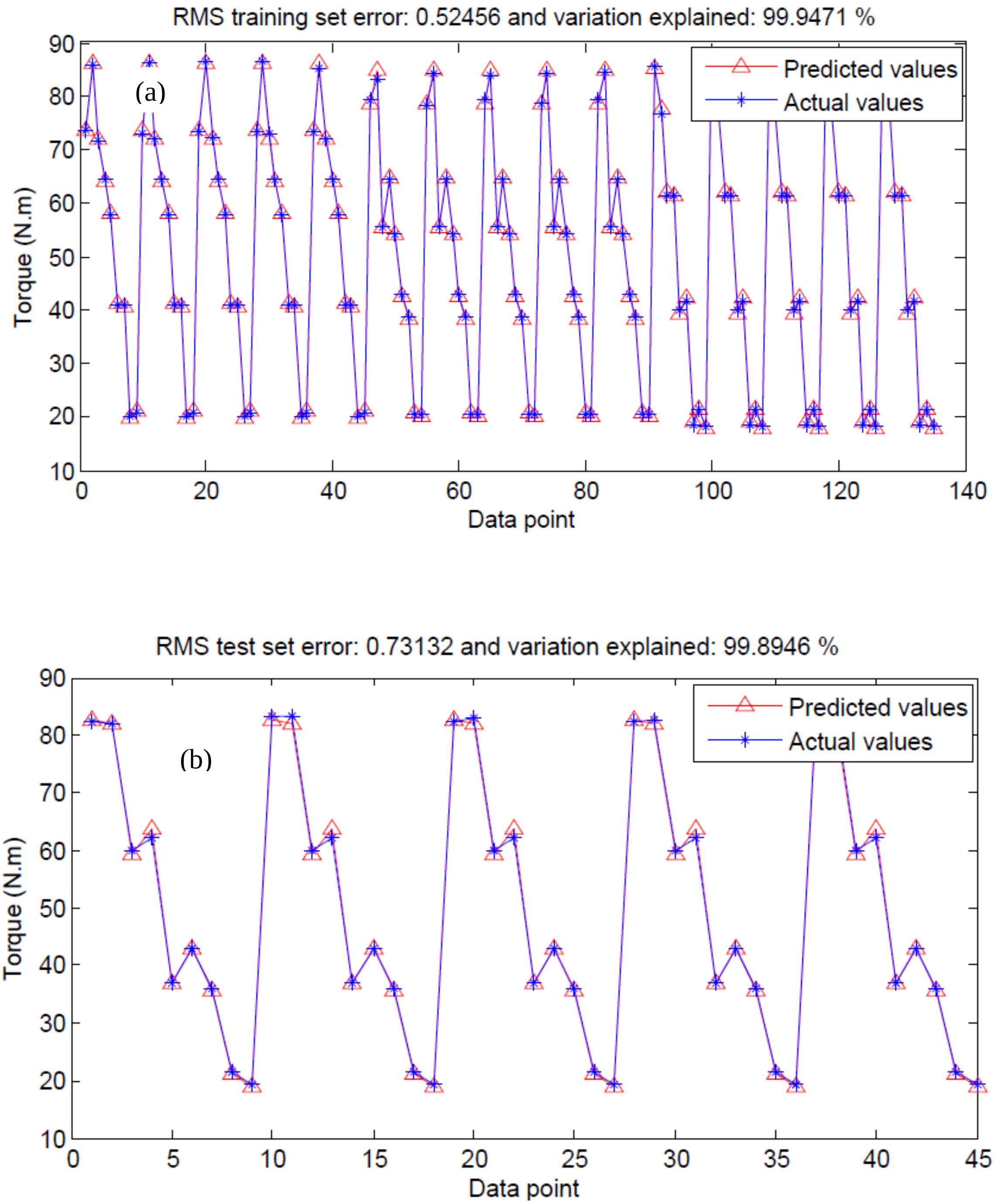

Fig. 7. Comparisons of experimental results and the GP model predictions of torque for, (a) training set data and (b) testining data set.

CONCLUSIONS

This research showed that a newly developed GP model can be used for predicting the brake power and torque of SI engine. The statistical parameters of R2 and RMSE demonstrated that the proposed multigene GP-based formulations results have best accuracy and can predict the engine brake power and torque close to experimental results. It is generally depicted that multigene GP is a powerful tool that has the capability to predict the engine performance and can be applied in the other industrial applications such as industrial processes and energy consumption which are even more sophisticated. The future work is using GP approach to predict other engine performance parameters and engine emissions using alternative fuels.

REFERENCES

[1] Kwok NM, Liu DK, Fang G. 2006. Evolutionary computing based mobile robot localization. Engineering Applications of Artificial Intelligence. 19(8):857-868.

[2] Izadifar M, Jahromi MZ. 2007. Application of genetic algorithm for optimization of vegetable oil hydrogenation process. Applied Thermal Engineering. 78:1–8.

[3] Laukkanen T, Tveit T, Ojalehto V, Miettinen K, Fogelholm C. Bilevel. 2012. Heat exchanger network synthesis with an interactive multi-objective optimization method. Applied Thermal Engineering. 48:301-316.

[4] Tsai J. 2012. An evolutionary approach for worst-case tolerance design. Engineering Applications of Artificial Intelligence. 25:917-925.

[5] Gusel L, Brezocnik M. 2011. Application of genetic programming for modelling of material characteristics. Expert Systems with Applications. 38:15014–15019.

[6] Chakraborty UK. 2009. Static and dynamic modeling of solid oxide fuel cell using genetic programming. Energy. 34:740–751.

[7] Chang HH. 2011. Genetic algorithms and non-intrusive energy management system based economic dispatch for cogeneration units. Energy. 36(1):181-190.

[8] Karakus M. 2011. Function identification for the intrinsic strength and elastic properties of granitic rocks via genetic programming (GP). Computers & Geosciences. 37(9):1818–1323.

[9] Pala M. 2008. Genetic programming-based formulation for distortional buckling stress of cold-formed steel members. Journal of Constructional Steel Research. 64:1495–1504.

[10] Asadi M, Eftekhari M, Bagheripour MH. 2011. Evaluating the strength of intact rocks through genetic programming. Applied Soft Computing. 11:1932–1937.

[11] Fonlupt C. 2001. Solving the ocean color problem using a genetic programming approach. Applied Soft Computing. 1:63–72.

[12] Can B, Heavey C. 2011. Comparison of experimental designs for simulation-based symbolic regression of manufacturing systems. Computers & Industrial Engineering. 61(3):447- 462.

[13] Tsai HC. 2011. Using weighted genetic programming to program squat wall strengths and tune associated formulas. Engineering Applications of Artificial Intelligence. 24:526-533.

[14] Pires JCM, Alvim-Ferraz MCM, Pereira MC, Martins FG. 2011. Prediction of tropospheric ozone concentrations: Application of a methodology based on the Darwin’s Theory of Evolution. Expert Systems with Applications. 38:1903–1908.

[15] Baykasoglu A, Gullu H, Canak H, Ozbakir L. 2008. Prediction of compressive and tensile strength of limestone via genetic programming. Expert Systems with Applications. 35:111-123.

[16] Cevik A, Cabalar AF. 2009. Modelling damping ratio and shear modulus of sand-mica mixtures using genetic programming. Expert Systems with Applications. 36(4):7749–7757.

[17] Hinchliffe MP, Willis MJ, Hiden HG, Tham MT, McKay B, Barton G. 1996. Modelling Chemical Process Systems Using a Multi-Gene Genetic Programming Algorithm. In: proceedings of the First Annual Conf on Genetic programming (Late Breaking Paper). pp. 56-65.

[18] Hiden HG. 1998. Data-Based Modelling using Genetic Programming. PhD Thesis. Dept. Chemical and Process Engineering, University of Newcastle. UK.

[19] Baziar MH, Jafarian Y, Shahnazari H, Movahed V, Tutunchian MA. 2011. Prediction of strain energy-based liquefaction resistance of sand–silt mixtures: An evolutionary approach. Computers & Geosciences. 37:1883–1893.

[20] Chen HM, Kao WK, Tsai HC. 2012. Genetic programming for predicting aseismic abilities of school buildings. Engineering Applications of Artificial Intelligence. 25:1103-1113.

[21] Koza JR. 1992. Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press. Cambridge.

[22] Can B, Heavey C. 2012. A comparison of genetic programming and artificial neural networks in metamodeling of discrete-event simulation models. Computers & Operations Research. 39:424–436.

[23] Searson DP. 2009a. GPTIPS:Genetic Programming & Symbolic Regressionfor MATLAB. Available at:http://gptips.sourceforge.net.

[24] Searson DP. 2009b. GPTIPS:Genetic Programming & Symbolic Regression for MATLAB, User Guide.