Energy 55 (2013) 519–528

Contents lists available at SciVerse ScienceDirect

Energy

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

Modeling and optimization of biodiesel engine performance using advanced machine learning methods

Ka In Wong [a,*], Pak Kin Wong [a], Chun Shun Cheung [b], Chi Man Vong [c]

[a] Department of Electromechanical Engineering, University of Macau, Macao
[b] Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong
[c] Department of Computer and Information Science, University of Macau, Macao

ARTICLE INFO

Article history:
Received 24 October 2012
Received in revised form
14 March 2013
Accepted 24 March 2013
Available online 26 April 2013

Keywords:
Biodiesel
Engine performance
Engine modeling
Engine optimization
Machine learning

ABSTRACT

This study aims to determine optimal biodiesel ratio that can achieve the goals of fewer emissions, reasonable fuel economy and wide engine operating range. Different advanced machine learning techniques, namely ELM (extreme learning machine), LS-SVM (least-squares support vector machine) and RBFNN (radial-basis function neural network), are used to create engine models based on experimental data. Logarithmic transformation of dependent variables is used to alleviate the problems of data scarcity and data exponentiality simultaneously. Based on the engine models, two optimization methods, namely SA (simulated annealing) and PSO (particle swarm optimization), are employed and a flexible objective function is designed to determine the optimal biodiesel ratio subject to various user-defined constraints. A case study is presented to verify the modeling and optimization framework. Moreover, two comparisons are conducted, where one is among the modeling techniques and the other is among the optimization techniques. Experimental results show that, in terms of the model accuracy and training time, ELM with the logarithmic transformation is better than LS-SVM and RBFNN with/without the logarithmic transformation. The results also show that PSO outperforms SA in terms of fitness and standard deviation, with an acceptable computational time.

1. Introduction

Diesel engines are increasingly used on many commercial vehicles due to their superior fuel efficiency and high durability performance. However, the amount of fuel consumed and pollutants emitted during their combustion process bring along serious environmental problems to the world. Some recent studies have already shown that particulates, being emitted from diesel engines, are associated with premature death [1,2]; it can accumulate in the human respiratory system [3] and cause serious health problems such as lung cancer according to World Health Organization. Moreover, the International Energy Agency forecasted that the world oil consumption would increase from 82.1 million barrels per day in 2004 to 115.4 million barrels per day in 2030 [4]. To deal with these environmental and energy issues, biodiesel becomes a popular viable alternative energy source for diesel engines nowadays because it can significantly reduce the fuel consumption and exhaust emissions. In addition, it is usually mixed with diesel fuel in some specific ratios so that the blends can be fueled into the engine without any significant engine modification [5,6].

Some recent studies [7–13] have investigated the performance and emission characteristics of biodiesel engines at different engine speeds, loads and biodiesel ratios. According to their results, with an increasing amount of biodiesel in fuel, both the BSFC (brake-specific fuel consumption) and the NO_x (nitrogen oxides) emission increase, whereas the CO (carbon monoxide) emission, the HC (hydrocarbon) emission and the PM (particulate mass) concentration reduce, particularly at high engine loads. These results imply

Abbreviations: AFR, air-fuel ratio; ANN, artificial neural network; BSFC, brake-specific fuel consumption; CO, carbon monoxide; CO_2, carbon dioxide; ELM, extreme learning machine; GA, genetic algorithm; HC, hydrocarbon; HCLA, heated chemiluminescent analyzer; LOOCV, leave-one-out cross-validation; LS-SVM, least-squares support vector machine; MAPE, mean absolute percentage error; $MAPE_{test}$, mean absolute percentage error for test data sets; $MAPE_{train}$, mean absolute percentage error for training data sets; NDIR, non-dispersive infra-red analyzer; NO_x, nitrogen oxides; NSGA-II, non-dominated sorting genetic algorithm-II; PM, particulate mass; PSO, particle swarm optimization; QNM, quasi-Newton method; RBFNN, radial-basis function neural network; RMSPE, root mean square percentage error; $RMSPE_{test}$, root mean square percentage error for test data sets; $RMSPE_{train}$, root mean square percentage error for training data sets; rpm, revolution per minute; SA, simulated annealing; SLFN, single-hidden-layer feedforward neural network; SPEA-2, strength Pareto evolutionary algorithm 2; SVM, support vector machine; TEOM, tapered element oscillating microbalance.

* Corresponding author. Tel./fax: +853 83974292.
E-mail address: imkain@gmail.com (K.I. Wong).

0360-5442/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.energy.2013.03.057

that the engine performance is sensitively affected by the ratio of biodiesel in the fuel. Besides, biodiesel is currently more expensive than diesel. As a result, an optimal biodiesel ratio that can result in fewer emissions at all the engine loads while keeping a low fuel cost is desired.

Nevertheless, there is no shortcut in determining this optimal biodiesel ratio because the factors (e.g., fuel cost and amount of exhaust emissions) are opposing each other. A very straightforward method for determining the optimal biodiesel ratio is by conducting numerous experiments on a dynamometer to obtain the optimal ratio subject to the user's requirements. The disadvantage of this method is that it requires a lot of time and money, and the procedure is not flexible since the constraints, like fuel cost, might vary.

Consequently, creating a mathematical model for biodiesel engines may be the best solution to the above limitation because the optimal biodiesel ratio can then be determined by applying some computer-aided optimization methods to the engine model. However, due to the complex nature of the relationship between the parameters of biodiesel engines, an exact mathematical model still remains unknown so far. In the current literature, there is an increasing interest in developing engine models from experimental data using ANN (artificial neural network) [14–19]. For instance, Canakci et al. [14] investigated the use of ANN for the prediction of the performance and exhaust emissions of a biodiesel diesel engine. Oguz et al. [15] developed ANN models to estimate the performance of the engine fueled with blends of diesel, biodiesel, and bioethanol. Yusaf et al. [16] also used ANN to predict the performance of a compression ignition engine using blends of crude palm oil and diesel as fuel. More recently, Ozgur et al. [17], Shivakumar et al. [18] and Mohamed Ismail et al. [19] still used ANN with backpropagation algorithm to predict the performance, emission characteristics and other engine responses of diesel engines fueled with different biodiesel blends. However, according to Wong et al. [20] and Haykin [21], traditional ANN has actually many drawbacks for its learning process, such as multiple local minima, user burden on selection of optimal network structure, slow learning speed, large training data size and over-fitting risk.

To compensate for the aforementioned limitations, several advanced machine learning techniques, such as LS-SVM (least-squares support vector machine) [22] and ELM (extreme learning machine) [23], have recently been introduced. They exhibit the major advantages of global optimum and higher generalization capability, which means that the prediction results should be more accurate than ANNs. Some recent researches successfully presented the application of LS-SVM [20,24–26] and ELM [27–29] to different engineering applications, and showed that they are superior to traditional ANNs. Nevertheless, the application of these two advanced machine learning techniques to biodiesel engine modeling is still very rare. As a result, this study aims to apply these two methods to model and predict the performance and emission characteristics of a biodiesel engine.

These aforesaid modeling techniques are data-driven, so experimental data are required for model training and verification. In order to collect the sample data within a reasonable time and cost, together with the narrow range of operating speeds of diesel engines as compared with gasoline engines, only a small set, say 60, of experimental data can be acquired, resulting in data scarcity. Moreover, some performance outputs of the sample data may suffer from the problem of exponentiality (i.e., the output $\boldsymbol{y}$ increases or decreases exponentially along input $\boldsymbol{x}$) [7,30] that seriously deteriorate the prediction accuracy. To deal with these two problems simultaneously, logarithmic transformation of dependent variables is applied to improve the prediction accuracy of the model. The purpose of the logarithmic transformation is to scale down the range and variation of all the dependent variables (i.e., the performance outputs) in the sample dataset.

After the engine models are created from LS-SVM and ELM respectively, a comparison among these methods with traditional ANNs is carried out accordingly. RBFNN (Radial-basis function neural network), being a well-known type of ANN that has a faster computation time and less uncertainty during the design procedure, was recently used for the prediction of biodiesel engine emissions [31]. It is therefore chosen to represent ANN to make a fair comparison.

From the comparison, the most accurate model is then selected for optimization. As this engine model could be very complicated, discontinuous and out of gradient information, traditional optimization methods are not applicable. In Wong et al. [25], QNM (quasi-Newton method), GA (genetic algorithm) and PSO (particle swarm optimization) were successfully applied to gasoline engine setup optimization based on LS-SVM engine model. Their results already proved that PSO has performed better than QNM and GA. Moreover, SA (simulated annealing), being a popular optimization method, has not been explored for engine optimization problems yet. Hence, it is an original and interesting research to compare PSO with SA for engine optimization problems. As a result, these two well-known optimization algorithms are selected to perform the optimization, and the optimization results are further verified through the experimental data.

2. Experimental setup for sample data collection

The test engine employed for the experiment in this study was a naturally aspirated, water-cooled, 4-cylinder, direct-injection diesel engine, and its specifications are shown in Table 1.

The engine was connected to an eddy-current dynamometer with a control system used for adjusting its speed and torque. Six fuels, including pure diesel, pure biodiesel, and four blended fuels (20%, 40%, 60%, 80% of biodiesel by volume), were used for data sampling. The experimental setup is illustrated in Fig. 1.

In the setup, HCLA (heated chemiluminescent analyzer) and NDIR (non-dispersive infra-red analyzer) were adopted to measure the gaseous species in the engine exhaust including CO, CO_2 (carbon dioxide) and NO_x on a continuous basis. The gas analyzers were calibrated with standard and zero gases before each experiment. PM concentration was measured with a TEOM (tapered element oscillating microbalance). The exhaust gas from the engine was diluted before passing through the TEOM with a Dekati mini-diluter.

The experiments were conducted at engine speeds of 1800 rpm (revolution per minute) and 2400 rpm, with five different engine torque of 28, 70, 130, 190 and 230 Nm for each engine speed and fuel blend, resulting in 60 sets of sample data. To ensure the repeatability and comparability of the measurements, the cooling water temperature was automatically controlled by a temperature controller to 80 °C, and held to within ±2 °C, while the lubricating

Table 1
Engine specifications.

Model	Isuzu 4HF1
Type	In-line four-cylinder
Maximum power	88 kW/3200 rpm
Maximum torque	285 Nm/1800 rpm
Bore × stroke	112 mm × 110 mm
Displacement	4334 cc
Compression ratio	19.0:1
Fuel injection timing (BTDC)	8°
Injection pump type	Bosch in-line type
Injection nozzle	Hole type (with five orifices)