Research Article

Modelling and Prediction of Particulate Matter, NO$_x$, and Performance of a Diesel Vehicle Engine under Rare Data Using Relevance Vector Machine

Ka In Wong,[1] Pak Kin Wong,[1] and Chun Shun Cheung[2]

[1] Department of Electromechanical Engineering, University of Macau, Taipa, Macau
[2] Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Correspondence should be addressed to Ka In Wong, imkain@gmail.com

Traditionally, the performance maps and emissions of a diesel engine are obtained empirically through many testes on the dynamometers because no exact mathematical engine model exists. In the current literature, many artificial-neural-network- (ANN-) based approaches have been developed for diesel engine modelling. However, the drawbacks of ANN would make itself difficult to be put into some practices including multiple local minima, user burden on selection of optimal network structure, large training data size, and overfitting risk. To overcome the drawbacks, this paper proposes to apply one emerging technique, relevance vector machine (RVM), to model the diesel engine, and to predict the emissions and engine performance. With RVM, only a few experimental data sets can train the model due to the property of global optimal solution. In this study, the engine speed, load, and coolant temperature are used as the input parameters, while the brake thermal efficiency, brake-specific fuel consumption, concentrations of nitrogen oxides, and particulate matter are used as the output parameters. Experimental results show the model accuracy is fairly good even the training data is scarce. Moreover, the model accuracy is compared with that using typical ANN. Evaluation results also show that RVM is superior to typical ANN approach.

1. Introduction

Air pollution is one of the most challenging problems today in many cities. The increased use of motor vehicles causes the amount of exhaust emissions to increase dramatically, which makes the problem more serious. Reducing the exhaust emissions from engines has then become an important concern of governments and motor vehicle manufacturers. Moreover, in view of the increasing oil price and the need to reduce emission of the global warming gas CO_2, there is a demand to reduce fuel consumption while maintaining the engine performance. Therefore, many researchers have focused on the relations between these two issues, namely, engine performance and emissions.

Diesel engines, though having the advantages of high fuel efficiency and high durability when compared to other engines, are the major source of nitrogen oxides (NO$_x$) and particulate matter (PM), which are harmful to human health and the environment. In particular, the fine and ultrafine particles ($\sim$10 micrometers or less) emitted by diesel engines can accumulate in the human respiratory system and cause various health problems [1] and influence global climate by absorbing solar radiation and reacting with other atmospheric constituents [2, 3]. Diesel engines are used extensively in buses and trucks; thus they are the major road-side emitters, posing significant threat to the health of the road users. In order to reduce these emissions, the combustion process of the engines has to be controlled. Additional hardware and instruments must be installed to monitor and control the engine operating parameters. Many experiments and tests must also be conducted to obtain a comprehensive understanding on the performance and emissions of the diesel engine. These are very complicated, time consuming, and expensive [4].

A way to solve these problems is to create a mathematical model for the diesel engine so that all the costly and immeasurable data can be predicted and virtual sensors can be used to replace the costly sensors. However, the combustion
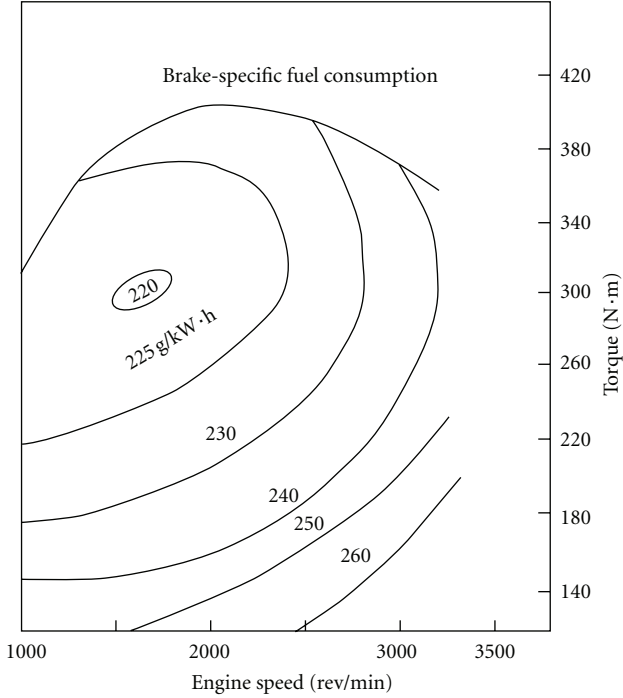
FIGURE 1: An example of diesel engine performance map [5].

process of a diesel engine is too complex that an exact mathematical model still does not exist today. Figure 1 shows one example of diesel engine performance map with only three variables in which the relationship among the engine load (torque), engine speed, and brake specific fuel consumption is already highly nonlinear. It could be imagined that if more variables are studied together, the model will be very complicated and very difficult to obtain. Moreover, the mathematical model varies for different engines.

In general, black-box identification is one of the commonly used modelling techniques suitable for engines because it can manage complex and uncertain information. Many recent researches in black-box identification have described the use of artificial neural network (ANN) for modelling of diesel engine performance [5–9] and emissions [7, 8, 10–13] based on experimental data sets. The ANN has in general, however, three main drawbacks for its learning process [14].

(1) The architecture, including the number of hidden neurons, has to be determined a priori or modified while training by heuristic, which results in a suboptimal network structure.

(2) The training process (i.e., the minimization of the residual squared error cost function) in ANN can easily become stuck in local minima. Various ways of preventing local minima, like early stopping, weight decay, have been employed. However, these methods greatly affect the generalization of the estimated function (i.e., the capacity of handling new input cases).

(3) The amount of training data is usually large. Normally at least 200~400 sets of training data is re-

quired to build an accurate ANN engine model [15]. However, the collection of diesel engine emission and performance data is usually time consuming and costly, so the data set is usually lower than 50, resulting in that ANN may not be a good solution for diesel engine modelling.

To overcome the disadvantages of ANN, an algorithm entitled relevance vector machine (RVM) was proposed by Tipping [16]. This approach is an emerging machine learning technique that is able to utilize more flexible candidate models, which are typically much sparser, offer probabilistic prediction, and avoid the need to set additional hyperparameters. The other advantage is that the training algorithm of RVM can ensure a global optimal solution whereas the learning process of ANN may cause a local optimal solution, so ANN requires more training data to minimize the risk [14]. With this good property, RVM is likely not to require too much sample data to build an accurate model. However, one deficiency of this approach is that the training time is approximately in the cube of the sample numbers. Thankfully, a fast training algorithm [17] is developed for RVM which initializes with an "empty" model, and sequentially "add" samples to increase the marginal likelihood, and also modify their weights. Within the same principal framework, the objective function can also be increased by deleting the samples which subsequently become redundant.

Recently, RVM has been applied to system modelling and predictive control [18–20]. These researches show that RVM is generally superior to the ANN. Moreover, the application of RVM to modelling of diesel engines under rare data is very few. For these reasons, in the present paper, RVM is employed to model the performance and emission characteristics of NO_x and PM of the diesel engine. Experiments are still required to provide sample data for RVM training. To demonstrate the effectiveness of this approach, a neural-network-based diesel engine model is also constructed and compared with the RVM model.

2. Relevance Vector Machine

The procedure of the RVM modelling is introduced here. Consider a training data set **D** of N input vectors $\{\mathbf{X}_n\}_{n=1}^N$, along with N corresponding scalar-valued output $\{y_n\}_{n=1}^N$. The output y_n is assumed to contain zero-mean Gaussian noise with variance σ^2. Hence, the probability of prediction error ε_n for y_n is a Gaussian distribution of zero mean and variance σ^2: $p(\varepsilon_n \mid \sigma^2) = N(0, \sigma^2)$, with

$$y_n = f(\mathbf{X}_n, \mathbf{w}) + \varepsilon_n. \tag{1}$$

That is,

$$p(y_n \mid \mathbf{X}_n, \mathbf{w}, \sigma^2) = N(f(\mathbf{X}_n, \mathbf{w}), \sigma^2), \tag{2}$$

where $f(\mathbf{X}_n, \mathbf{w})$ in (1) is the prediction model for the model output, y_n, with the input $\mathbf{X}_n$ and $\mathbf{w} = [w_1, \ldots, w_N]$ is the weight vector for the RVM model.

The predicted output $\hat{y}$ at an input $\mathbf{X}$ in the kernel model can be represented by

$$
\begin{aligned}
\hat{y} &= f(\mathbf{X}, \mathbf{w}) \\
&= \sum_{i=0}^{N} w_i K(\mathbf{X}, \mathbf{X}_i) \\
&= \mathbf{\Phi} \mathbf{w},
\end{aligned}
\tag{3}
$$

where $K(\mathbf{X}, \mathbf{X}_i)$ is a basis function and $\mathbf{\Phi}$ is the $N \times (N + 1)$ design matrix with $\mathbf{\Phi} = [\phi(\mathbf{X}_1) \cdots \phi(\mathbf{X}_N)]^T$, wherein $\phi(\mathbf{X}_N) = [1 \, K(\mathbf{X}_n, \mathbf{X}_1) \cdots K(\mathbf{X}_n, \mathbf{X}_N)]^T$. In this research, radial basis function (RBF) is chosen as the basis function K because it is commonly used for modelling problems. The approach for estimating $\hat{y}$ is to maximize the likelihood in

$$
p(\mathbf{y} \mid \mathbf{w}, \sigma^2) = (2\pi)^{-N/2} \sigma^{-N} \exp\left\{ -\frac{\|\mathbf{y} - \mathbf{\Phi}\mathbf{w}\|^2}{2\sigma^2} \right\}.
\tag{4}
$$

The likelihood function in (4) is complemented by a prior over the weights $\mathbf{w} = \{w_i\}$, $i = 0$ to N, to control the complexity of the model function and avoid overfitting. The prior is a zero-mean Gaussian probability distribution and is defined over every weight w_i as follows:

$$
p(\mathbf{w} \mid \boldsymbol{\alpha}) = (2\pi)^{-N/2} \prod_{i=0}^{N} \alpha_i^{1/2} \exp\left(-\frac{\alpha_i w_i^2}{2} \right).
\tag{5}
$$

The hyperparameters vector, $\boldsymbol{\alpha} = [\alpha_0, \ldots, \alpha_N]^T$, controls how far for each weight, w_i, is allowed to deviate from zero. Consequently, using Bayes' rule, the posterior over $\mathbf{w}$ is given as follows:

$$
p(\mathbf{w} \mid \mathbf{y}, \boldsymbol{\alpha}, \sigma^2) = \frac{p(\mathbf{y} \mid \mathbf{w}, \sigma^2) \, p(\mathbf{w} \mid \boldsymbol{\alpha})}{p(\mathbf{y} \mid \boldsymbol{\alpha}, \sigma^2)},
\tag{6}
$$

where $p(\mathbf{y} \mid \boldsymbol{\alpha}, \sigma^2)$ is the normalizing factor. $p(\mathbf{y} \mid \mathbf{w}, \sigma^2)$ and $p(\mathbf{w} \mid \boldsymbol{\alpha})$ are both Gaussian priors. The posterior mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ are as follows [17]:

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \left(\mathbf{A} + \sigma^{-2} \mathbf{\Phi}^T \mathbf{\Phi} \right)^{-1}, \\
\boldsymbol{\mu} &= \sigma^{-2} \boldsymbol{\Sigma} \mathbf{\Phi}^T \mathbf{y},
\end{aligned}
\tag{7}
$$

where $\mathbf{A}$ defines as $\mathrm{diag}(\alpha_0, \ldots, \alpha_N)$. In fact, the $\mathbf{w}$ in (3) can be set to the fixed $\boldsymbol{\mu}$ for the purpose of point prediction.

Rather than extending the model to include Bayesian inference over those hyperparameters (which is analytically intractable), a most-probable point estimate, $\boldsymbol{\alpha}_{\mathrm{MP}}$, may be found via a type II maximum likelihood procedure. That is called sparse Bayesian learning which is formulated as the local maximization with respect to $\boldsymbol{\alpha}$ of the marginal likelihood, or equivalently, its logarithm $L(\boldsymbol{\alpha})$:

$$
\begin{aligned}
L(\boldsymbol{\alpha}) &= \log p(\mathbf{y} \mid \boldsymbol{\alpha}, \sigma^2) = \log \int_{-\infty}^{\infty} p(\mathbf{y} \mid \mathbf{w}, \sigma^2) p(\mathbf{w} \mid \boldsymbol{\alpha}) d\mathbf{w} \\
&= -\frac{1}{2} \left[N \log 2\pi + \log|\mathbf{C}| + \mathbf{y}^T \mathbf{C}^{-1} \mathbf{y} \right],
\end{aligned}
\tag{8}
$$

Where

$$
\mathbf{C} = \sigma^2 \mathbf{I} + \mathbf{\Phi} \mathbf{A} \mathbf{\Phi}^T.
\tag{9}
$$

The covariance, $\boldsymbol{\Sigma}_{\mathrm{MP}} = \boldsymbol{\Sigma}$, can be obtained by substituting $\boldsymbol{\alpha} = \boldsymbol{\alpha}_{\mathrm{MP}}$ into $\mathbf{A}$ in (7), so that the posterior mean weight, $\boldsymbol{\mu}_{\mathrm{MP}}$, is obtained by evaluating (7) again with $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{MP}}$, giving a final (posterior mean) approximator:

$$
\begin{aligned}
Y &= f(\mathbf{X}^*, \boldsymbol{\mu}_{\mathrm{MP}}) \\
&= \sum_{i=0}^{N} \mu_{\mathrm{MP}_i} K(\mathbf{X}^*, \mathbf{X}_i) \\
&= \sum_{i=0}^{N} \mu_{\mathrm{MP}_i} \exp\left(-\frac{\|\mathbf{X}^* - \mathbf{X}_i\|}{\sigma^2} \right),
\end{aligned}
\tag{10}
$$

where Y is the prediction of the model output with the unseen input data $\mathbf{X}^*$. One crucial observation is that typically the optimal values of many hyperparameters are infinite [16]. With (7), this leads to a parameter posterior infinitely peaked at zero for many weights w_i with the consequence that $\boldsymbol{\mu}_{\mathrm{MP}}$ correspondingly comprises very few nonzero elements.

A recent analysis has showed that $L(\boldsymbol{\alpha})$ has a unique maximum with respect to α_i [16]:

$$
\alpha_i = \frac{s_i^2}{q_i^2 - s_i} \quad \text{if } q_i^2 > s_i
\tag{11}
$$

$$
\alpha_i = \infty \quad \text{if } q_i^2 \leq s_i,
\tag{12}
$$

and from these, it simply follows:

$$
s_m = \frac{\alpha_i S_i}{\alpha_i - S_i} \qquad q_i = \frac{\alpha_i Q_i}{\alpha_i - S_i}.
\tag{13}
$$

Note that when $\alpha_i = \infty$, $s_i = S_i$ and $q_i = Q_i$, then, it is convenient to utilize the Woodbury identity to obtain the quantities of interest:

$$
S_i = \phi_i^T (\sigma^{-2}\mathbf{I}) \phi_i - \phi_i^T (\sigma^{-2}\mathbf{I}) \mathbf{\Phi} \boldsymbol{\Sigma} \mathbf{\Phi}^T (\sigma^{-2}\mathbf{I}) \phi_i,
\tag{14}
$$

$$
Q_i = \phi_i^T (\sigma^{-2}\mathbf{I}) \mathbf{y} - \phi_i^T (\sigma^{-2}\mathbf{I}) \mathbf{\Phi} \boldsymbol{\Sigma} \mathbf{\Phi}^T (\sigma^{-2}\mathbf{I}).
\tag{15}
$$

The results of (14) and (15) imply that

(1) if ϕ_i is included in the model (i.e. $\alpha_i < \infty$) yet $q_i^2 \leq s_i$, then ϕ_i can be deleted (i.e., set α_i to ∞);

(2) if ϕ_i is excluded from the model ($\alpha_i = \infty$) and $q_i^2 > s_i$, ϕ_i can be added (i.e., set α_i to some optimal finite values).

To train and update the RVM model dynamically, a sequential learning algorithm is required. The algorithm starts with an empty model, and sequentially adds basis functions to increase the marginal likelihood, and modify their weights. Within the same principal framework, the likelihood can also be increased by deleting those basis functions which subsequently become redundant. Since this algorithm sequentially adds or deletes the basis function to or from the model, the likelihood can be continually increased by adding and deleting basis function and this mechanism make online

model update feasible. The steps of the sequential learning algorithm are shown below.

(1) Initialize σ^2 to some sensible values (e.g., var $[\mathbf{y}]\times 0.1$) and all other α_i are notionally set to infinity.

(2) Initialize S_n and Q_n with a single basis vector ϕ_i from (13) and (14) and compute new α_n from (11) which can be simplified as

$$\alpha_i = \frac{\left\|\phi_i\right\|^2}{\left\|\phi_i^{\mathrm{T}}\mathbf{y}\right\|^2/\left\|\phi_i\right\|^2 - \sigma^2}. \qquad (16)$$

(3) Explicitly compute $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$ (which are scalars initially), along with initial values of s_i and q_i for all N basis functions ϕ_i.

(4) Select a candidate basis vector ϕ_i from the set of all N basis functions.

(5) Compute $\theta_i = Q_i^2 - S_i$.

(6) If $\theta_i > 0$ and $\alpha_i < \infty$ (i.e., ϕ_i is included in the model), then reestimate α_i.

(7) If $\theta_i = 0$ and $\alpha_i = \infty$, then add ϕ_i to the model with updated α_i.

(8) If $\theta_i < 0$ and $\alpha_i < \infty$, then delete ϕ_i from the model and set $\alpha_i = \infty$.

(9) Estimate the noise level, update σ^2 as follows:

$$\sigma^2 = \frac{\left\|\mathbf{y} - \{f(\mathbf{X}_n,\mathbf{w})\}_{n=1}^N\right\|^2}{(N - M + \sum_i \alpha_i \boldsymbol{\Sigma}_{ii})}. \qquad (17)$$

(10) Recomputed or update $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$ and all S_i and Q_i using (7), (13) to (15).

(11) If converged then terminate, otherwise go to Step 4.

It has to be noticed that the RVM modelling algorithm is only a multi-input but single-output modelling method. Therefore, individual model corresponding to each output needed to be constructed. A multi-input/multioutput model is then easily be obtained by combining all the individual models.

3. Experimental Setup

Sample data sets are required for RVM training and are generally collected through experiments. In this study, the experiments were conducted on a naturally aspirated, water-cooled, 4-cylinder, direct-injection diesel engine. The specifications of the engine are shown in Table 1.

The engine was connected to an eddy-current dynamometer, and a control system was used for adjusting its speed and torque. Ultralow sulfur diesel fuel containing less than 10-ppm-wt sulfur was adopted in the test. The experimental setup is illustrated in Figure 2.

The experiments were carried out at engine speeds of 1200, 1400, 1600, 1800, and 2000 rpm and each at engine loads of 28, 70, 140, 210, and 252 Nm. For each test, the

TABLE 1: Engine specifications.

Model	Isuzu 4HF1
Type	In-line four cylinders
Maximum power	88 kW/3200 rev min^{-1}
Maximum torque	285 Nm/1800 rev min^{-1}
Bore $\times$ stroke	112 mm $\times$ 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)

volumetric flow rate of fuel was measured using a measuring cylinder and then converted into mass consumption rate, which is used to calculate the brake-specific fuel consumption (BSFC) and the brake thermal efficiency (BTE). The gaseous species in the engine exhaust including CO, CO_2, and NOx, were measured on a continuous basis using the Anapol EU5000 exhaust gas analyzer which was suitable for measuring diesel engine emissions. The Anapol EU5000 used infrasensors for measuring CO and CO_2 concentrations and used chemical cells for measuring NO and NO_2 to obtain the NO_x concentration. The gas analyzer was calibrated with standard and zero gases before each experiment. Particulate mass concentration was measured with a tapered element oscillating microbalance (TEOM, Series 1105, Rupprecht & Patashnick Co., Inc.). The exhaust gas from the engine was diluted before passing through the TEOM with a Dekati minidiluter. The dilution ratio (DR) was evaluated based on the following equation:

$$\mathrm{DR} = \frac{[CO_2]_{\mathrm{exhaust}} - [CO_2]_{\mathrm{background}}}{[CO_2]_{\mathrm{diluted}} - [CO_2]_{\mathrm{background}}}, \qquad (18)$$

where $[CO_2]_{\mathrm{exhaust}}$, $[CO_2]_{\mathrm{diluted}}$, and $[CO_2]_{\mathrm{background}}$ represent the undiluted, the diluted, and the background CO_2 concentrations, respectively. The dilution ratio was around 8 in the tests.

At each speed and load, data were recorded after the engine had reached the steady state, which was indicated by the lubricating oil temperature and the coolant temperature. For the purpose of reducing experimental uncertainties and ensuring repeatability of test data, the data were recorded continuously for 5 minutes to reduce experimental uncertainties, and each test was carried out three times. The average values were used in this research.

Based on the measured data, the following parameters are derived:

Brake thermal efficiency:

$$\mathrm{BTE} = \eta_b = \frac{P_b}{\dot{m}_f LHV}, \qquad (19)$$

Brake specific fuel consumption:

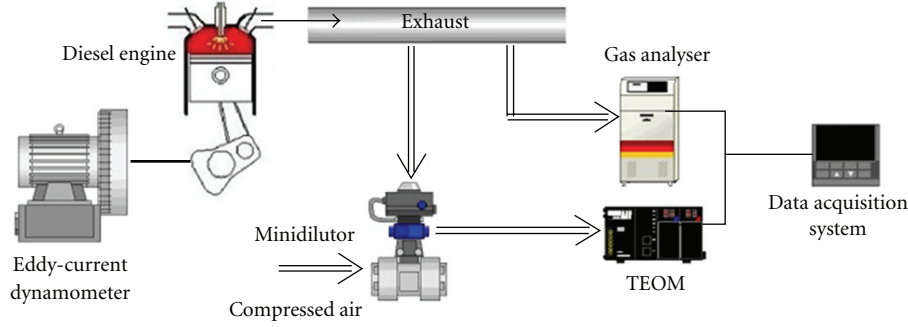$$\mathrm{BSFC} = \frac{\dot{m}_f}{P_b}, \qquad (20)$$

FIGURE 2: Schematic diagram of experimental setup.

where P_b is the brake power calculated from the measured torque and engine speed, $\dot{m}_f$ is the mass flow rate of the diesel fuel and LHV is the lower heating value of the diesel fuel.

4. Application of RVM and Modelling Results

To evaluate the effectiveness of RVM, the prediction models were built based on the experimental data. As the collection of the experimental data is time consuming and costly, only 22 data sets corresponding to different load and speed settings were collected from the experiments, which are shown in Table 2. 18 sets of them were used as the training data for the model construction, and the rest 4 sets were used for model validation and testing. Actually, several weeks were required to collect the twenty-two data sets professionally. Table 3 illustrates the use of each of the data sets.

The measured parameters in each of the data sets can be basically separated into two categories, which are the input parameters and output parameters. Engine speed and engine load are the two most important independent parameters that affect the engine performance and emissions. They are included in the input parameters. The coolant controls the engine temperature so the coolant temperature is regarded as an important factor and is also treated as the input parameter. The brake-specific fuel consumption and the brake thermal efficiency represent the engine performance; thus, they are used as the output parameters. Moreover, the NO_x and particulate matter are two most serious exhaust emissions from diesel engine. Therefore, the output parameters also consist of the NO_x concentration and particle mass concentration.

The RVM modelling was implemented using MATLAB. There are three input parameters and four output parameters, indicating that four individual RVM models have to be built. Moreover, in order to have a more accurate modelling result and to prevent any input parameter from dominating the output value, the input data is conventionally normalized before training [21]. In this study, all the input values were normalized within the range $[-1, 1]$.

To verify the accuracy of the RVM model, the predicted output values is compared with the actual values from the test data sets and shown in Figures 3, 4, 5, and 6.
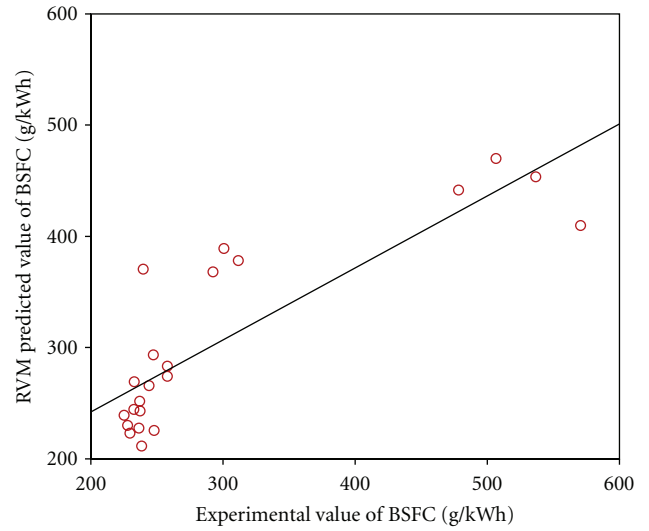

FIGURE 3: Comparison between RVM predicted values and the corresponding actual values for BSFC.

The corresponding prediction errors are presented by the mean absolute percentage error (*MAPE*); they were evaluated against the experimental data sets using (21). Moreover, the fraction of variance (*R*-squared value) is also calculated using (22) and (23). The smaller the MAPE, the better the modelling accuracy is. In addition, the best fitness of R^2 is 1

$$\text{MAPE} = \frac{1}{N_t} \sum_{k=1}^{N_t} \left| \frac{y_k - f(\mathbf{X_k})}{y_k} \right| \times 100\%, \qquad (21)$$

$$R^2 = 1 - \left(\frac{\sum_{k=1}^{N_t} (y_k - f(\mathbf{X_k}))^2}{\sum_{k=1}^{N_t} (y_k - \overline{y})^2} \right), \qquad (22)$$

$$\overline{y} = \frac{1}{N_t} \sum_{k=1}^{N_t} y_k, \qquad (23)$$

where $\mathbf{X_k}$ is the kth input vectors for the prediction, $f(\mathbf{X_k})$ is the prediction value corresponding to $\mathbf{X_k}$, y_k is the actual value corresponding to $\mathbf{X_k}$, $\overline{y}$ is the mean of the actual value, and N_t is the number of test data points.

TABLE 2: Experimental dataset for model training and validation.

	Input parameters			Output parameters			
Data sets no.	Engine speed (rpm)	Engine load (Nm)	$T_{coolant}$ (°C)	BSFC (g/kWh)	BTE (%)	NO_x concentration (ppm)	PM mass concentration (mg/m³)
1	1200	140	80.7	233.16	36.23	665.25	3.363
2	1200	210	82.8	225.23	37.51	900.50	9.501
3	1200	252	83.2	232.83	36.28	957.19	38.403
4	1400	28	79.3	478.57	17.65	199.84	1.233
5	1400	70	80.2	292.57	28.88	325.21	2.049
6	1400	140	80.7	236.49	35.72	610.04	3.378
7	1400	210	83.2	227.94	37.06	818.10	11.484
8	1400	252	83.7	237.09	35.63	836.57	50.906
9	1600	28	80.3	506.77	16.67	175.13	2.203
10	1600	70	80.4	300.99	28.07	303.89	3.045
11	1600	140	81.5	238.77	35.38	566.38	4.880
12	1600	210	82.3	229.92	36.74	795.66	22.049
13	1600	252	84.3	244.17	34.60	814.56	72.437
14	1800	28	80.2	537.04	15.73	170.97	1.936
15	1800	70	80.7	311.98	27.08	283.75	2.836
16	1800	140	81.4	248.01	34.07	526.49	5.323
17	1800	210	83.1	237.34	3.60	748.64	22.093
18	1800	252	85.0	258.16	32.73	730.34	115.988
19	2000	28	80.7	570.84	14.80	171.73	1.550
20	2000	70	80.8	239.71	25.62	280.93	3.418
21	2000	140	82.8	257.99	32.75	507.44	4.192
22	2000	210	83.9	247.54	34.13	683.83	26.865

TABLE 3: Data set assignment. (T and X refer to training sets and test sets, resp.)

		Engine load/torque (Nm)				
		28	70	140	210	252
Engine speed (rpm)	1200	n/a	n/a	T	T	T
	1400	T	X	T	T	T
	1600	T	T	T	T	X
	1800	T	T	T	X	T
	2000	T	T	X	T	n/a

Table 4 summarizes the training MAPE, the MAPE over the test data sets, and the fraction of variance for each output parameter of the RVM model.

5. Comparison of RVM and ANN Modelling Results

To illustrate the advantages and superiority of the proposed RVM model, the prediction result was compared with a multilayer feed-forward neural network with backpropagation. Since multilayer feed-forward neural network is a well-known universal estimator [22] and many researches for
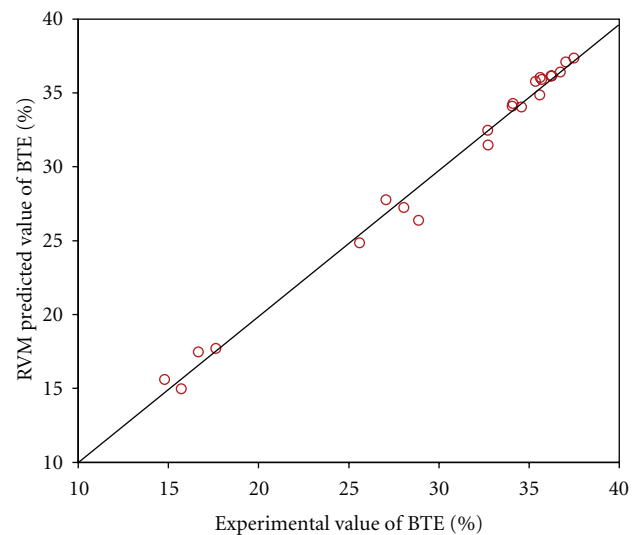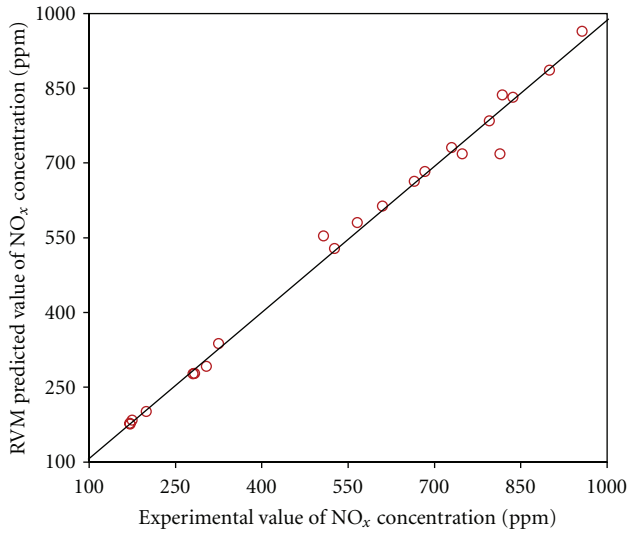

FIGURE 4: Comparison between RVM predicted values and the corresponding actual values for BTE.

diesel engine performance modelling [5–11, 13] were done based on this configuration, the results from it can be considered as a rather standard benchmark.

TABLE 4: Results of the RVM models.

Output parameters	MAPE over the training data sets (%)	MAPE over the test data sets (%)	R^2
BSFC	13.83	11.55	0.70
BTE	1.65	4.12	0.99
NO$_x$ emission	1.61	7.17	0.99
Mass concentration	11.52	63.14	0.97
Overall average	7.15	21.50	

FIGURE 5: Comparison between RVM predicted values and the corresponding actual values for NO$_x$ concentration.

FIGURE 6: Comparison between RVM predicted values and the corresponding actual values for PM mass concentration.

A neural network with one hidden layer was built based on the same training data sets used for RVM modelling. The neural network consists of 3 input neurons, 20 hidden neurons, and 4 output neurons. In fact, the number of hidden nodes was determined by a trial and error analysis, varying the number of hidden neurons between 3 and 30, this burden demonstrates the ineffectiveness of the ANN approach.

The activation function used inside the hidden layer was the Tan-Sigmoid transfer function, while a pure linear filter was employed for the output layer. Levenberg-Marquardt algorithm was used as the training algorithm. The learning rate of the weight update was set to be 0.05. Figure 7 depicts the architecture of the neural network.

The same test sets were also chosen so that the RVM and ANN model can be compared reasonably. The prediction accuracy of each output in the ANN model is illustrated in Figures 8, 9, 10, and 11 and Table 5.

Tables 4 and 5 show that the RVM outperforms the ANN by about 36.45% in terms of average MAPE under the same test sets. The relatively high training MAPE of the ANN shows that the data sets is not sufficient for building such a highly nonlinear model. Furthermore, only one initial
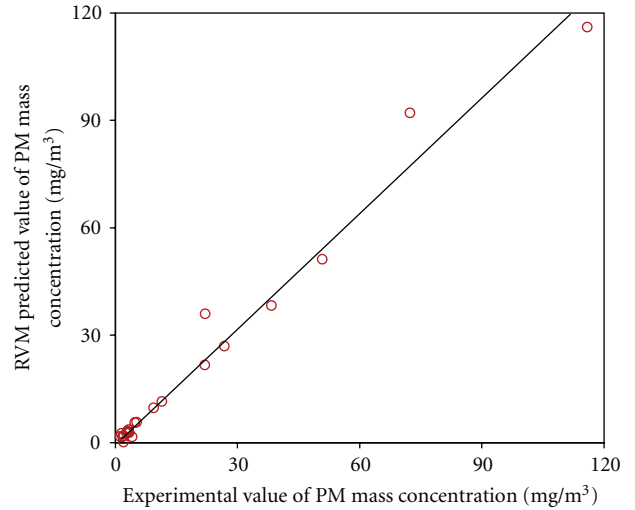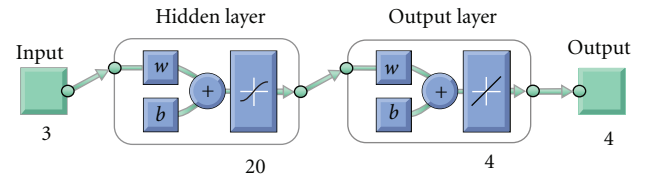

FIGURE 7: Architecture of the neural network.

value σ of the basis width is required by RVM, while the learning rate, number of hidden layers, and number of hidden neurons are required in ANN, which means a grid of guessed values for these parameters have to be prepared.

The MAPEs of both RVM and ANN for predicting the mass concentration of particulates are relatively large as compared to the other output parameter. This is because the variation of the mass concentration ranges from 0 to $12 \times 10^4 \, \mu g/m^3$. Actually, the RVM model tries to fit a function for the whole range rather than focusing on the low end of range, which is seen by the R-squared value of 0.97. In contrast, the ANN model tends to concentrate at the low end of the value. As a result, the R-squared value for the ANN model is only 0.02, which is unacceptable.
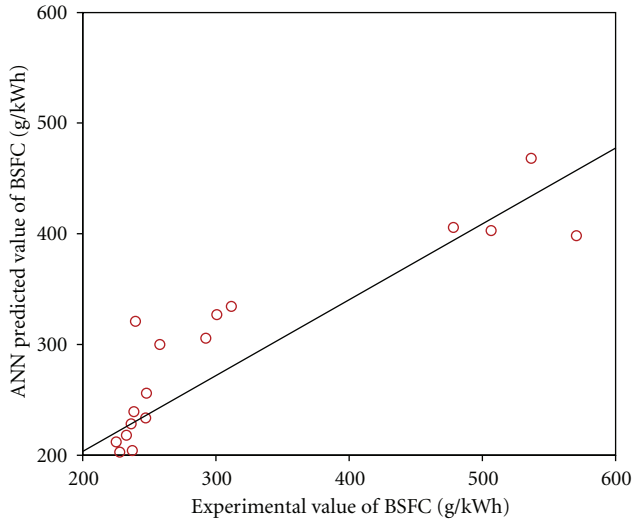
Figure 8: Comparison between ANN predicted values and the corresponding actual values for BSFC.
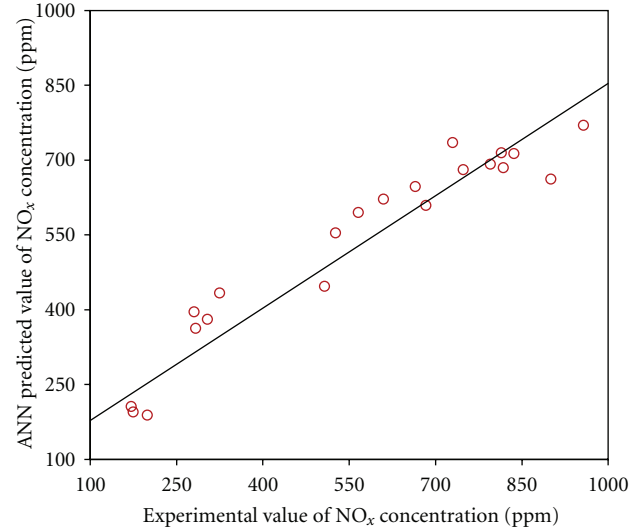

Figure 10: Comparison between ANN predicted values and the corresponding actual values for NO$_x$ concentration.
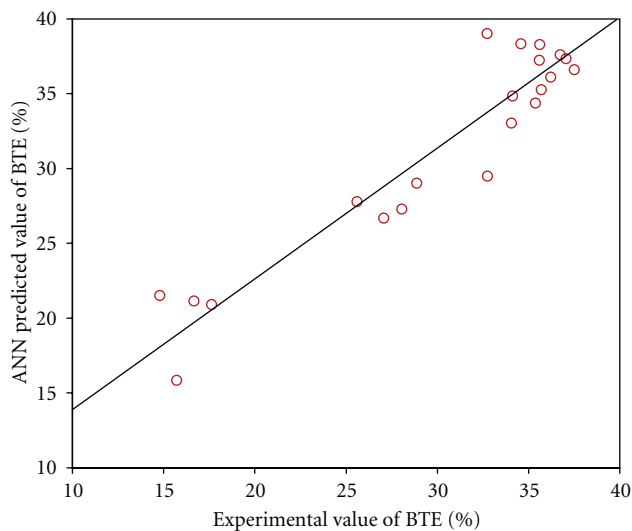

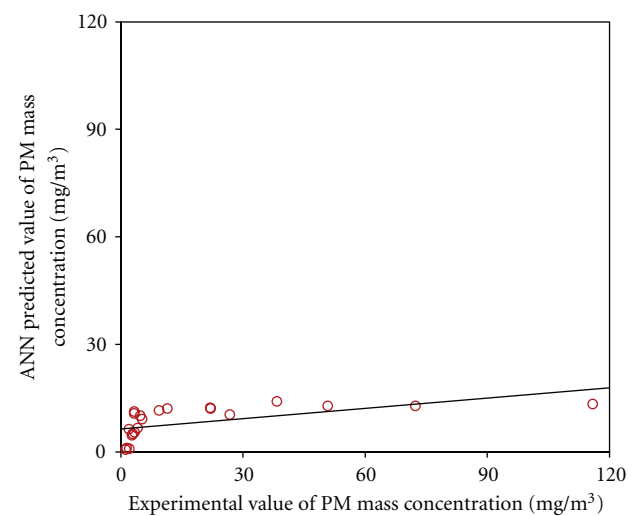Figure 9: Comparison between ANN predicted values and the corresponding actual values for BTE.

Figure 11: Comparison between ANN predicted values and the corresponding actual values for PM mass concentration.

Overall, the prediction accuracy of RVM with a small amount of training data is satisfactory.

6. Conclusions

This research is the first attempt at applying RVM to model the diesel engine performance and emission characteristics of NO$_x$ and particulate matter under the condition of rare data. Although the combustion process of the diesel engine is unknown, the RVM model has successfully demonstrated the relation between the controllable factors, which are the engine speeds, engine loads, and coolant temperature, and the output variables, including the brake-specific fuel consumption, brake thermal efficiency, NO$_x$ emission, and particulate mass concentrations. Experimental results show

that the RVM model is still acceptable even if the data sets are few. It is believed that more data sets can improve the accuracy of the model.

Furthermore, the RVM model is also compared with an ANN model. The results indicate that the average accuracy of the RVM model is higher than that of the ANN model by about 36.45%, implying that RVM is superior to ANN.

With the proposed RVM model, experimental efforts can be reduced significantly as the performance and emissions of the diesel engine can be predicted easily. By applying this RVM model as a virtual sensor on diesel vehicles, the exhaust emissions can be controlled more effectively by incorporating with some advanced control algorithms, such as model predictive control. The study of model predictive diesel emission control based on RVM model will

TABLE 5: Results of the ANN model.

Output parameters	MAPE over the training data sets (%)	MAPE over the test data sets (%)	R^2
BSFC	14.13	14.12	0.70
BTE	8.66	6.42	0.86
NO_x emission	18.15	16.64	0.85
Mass concentration	87.42	98.12	0.02
Overall average	32.09	33.83	

be considered as a future work. Since RVM can also perform online model update, the applications of RVM to online system modelling and online control will also be explored in the future.

Acknowledgments

References

[1] D. W. Dockery and C. A. Pope, "Acute respiratory effects of particulate air pollution," *Annual Review of Public Health*, vol. 15, pp. 107–132, 1994.

[2] J. E. Penner, C. C. Chuang, and K. Grant, "Climate forcing by carbonaceous and sulfate aerosols," *Climate Dynamics*, vol. 14, no. 12, pp. 839–851, 1998.

[3] M. Z. Jacobson, "Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols," *Nature*, vol. 409, no. 6821, pp. 695–697, 2001.

[4] W. Knecht, "Diesel engine development in view of reduced emission standards," *Energy*, vol. 33, no. 2, pp. 264–271, 2008.

[5] M. S. Rawlins, *Diesel engine performance modelling using neural networks*, Ph.D. thesis, Department of Mechanical Engineering, Durban Institute of Technology, Durban, South Africa, 2005.

[6] V. Çelik and E. Arcaklioğlu, "Performance maps of a diesel engine," *Applied Energy*, vol. 81, no. 3, pp. 247–259, 2005.

[7] M. Canakci, A. Erdil, and E. Arcaklioğlu, "Performance and exhaust emissions of a biodiesel engine," *Applied Energy*, vol. 83, no. 6, pp. 594–605, 2006.

[8] Shivakumar, P. Srinivasa Pai, and B. R. Shrinivasa Rao, "Artificial Neural Network based prediction of performance and emission characteristics of a variable compression ratio CI engine using WCO as a biodiesel at different injection timings," *Applied Energy*, vol. 88, no. 7, pp. 2344–2354, 2011.

[9] H. Oğuz, I. Saritas, and H. E. Baydan, "Prediction of diesel engine performance using biofuels with artificial neural network," *Expert Systems with Applications*, 2010.

[10] C. Brace, "Prediction of diesel engine exhaust emission us- ing artificial neural networks," in *Proceedings of the Neural Networks in Systems Design' IMechE Seminar S591*, Solihull, UK, 1998.

[11] M. Traver, R. Atkinson, and C. Atkinson, "Neural network-based diesel engine emissions prediction using in-cylinder combustion pressure," SAE Paper 1999-01-1532, 1999.

[12] Z. T. Liu and S. M. Fei, "Study of CNG/diesel dual fuel engine's emissions by means of RBF neural network," *Journal of Zhejiang University*, vol. 5, no. 8, pp. 960–965, 2004.

[13] M. Ghazikhani and I. Mirzaii, "Soot emission prediction of a waste-gated turbo-charged DI diesel engine using artificial neural network," *Neural Computing and Applications*, vol. 20, no. 2, pp. 303–308, 2011.

[14] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, Upper Saddle River, NJ, USA, 1999.

[15] P. K. Wong, L. M. Tam, and K. Li, "Automotive engine power performance tuning under numerical and nominal data," *Control Engineering Practice*, vol. 20, no. 3, pp. 300–314, 2011.

[16] M. E. Tipping, "Sparse Bayesian Learning and the Relevance Vector Machine," *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.

[17] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian method," in *Proceedings of 9th International Workshop on Artificial Intelligence and Statistics*, Key West, Fla, USA, January 2003.

[18] J. Yuan, K. Wang, T. Yu, and M. Fang, "Integrating relevance vector machines and genetic algorithms for optimization of seed-separating process," *Engineering Applications of Artificial Intelligence*, vol. 20, no. 7, pp. 970–979, 2007.

[19] W. Caesarendra, A. Widodo, and B. S. Yang, "Application of relevance vector machine and logistic regression for machine degradation assessment," *Mechanical Systems and Signal Processing*, vol. 24, no. 4, pp. 1161–1171, 2010.

[20] M. Han and Y. Zhao, "Dynamic control model of BOF steelmaking process based on ANFIS and robust relevance vector machine," *Expert Systems with Applications*, vol. 38, no. 12, pp. 14786–14798, 2011.

[21] D. Pyle, *Data Preparation for Data Mining*, Morgan Kaufmann Publishers, 1999.

[22] C. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, New York, NY, USA, 1995.