Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-20 (2022) 37–42

Systematic Comparison of Supervised Learning Methods to Reduce Calibration Effort in Engine Control Development

Shreyas Sastry V M * **Prasoon Garg** * **Emilia Silvas** *,** **Frank Willems** *,**

* *Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands (e-mail: p.garg.1@tue.nl)*
** *TNO Automotive, Helmond, The Netherlands*

Abstract: Complexity of engine control systems is continuously growing due to an increased number of subsystems and the need for robust performance. For traditional map-based as well as state-of-the-art model-based approaches, this will lead to unacceptable development costs and time for future engines. Parametrization of the embedded models using supervised learning regression methods can immensely reduce the number of calibration parameters and hence the calibration effort. However, a methodology for performance assessment of different promising data-driven modelling methods for engine control development is currently lacking. In this paper, a systematic methodology that assesses model inaccuracy, and also implementation aspects such as calibration effort and computational complexity is introduced. This method is applied to assess the potential of Supervised Learning (SL) methods for parametrizing the feedforward controller of a modern diesel engine air-path controller. Using requirements analysis and the specified performance criteria, two regression methods were selected: artificial neural networks (ANN) and support vector machines (SVM). After careful data selection and model training, performance is compared with the benchmark controller, which uses a physics-based model. From simulation results, it is shown that a 97% reduction in the number of calibration parameters with both regression models can be realized. For a standard test cycle, cumulative engine out NOx emissions with regression based controllers are close to the allowable inaccuracy of 10% compared to the benchmark controller. Among the two methods, ANN shows the best performance for the studied performance criteria of inaccuracy, number of calibration parameters and computational complexity.

Keywords: supervised learning, neural networks, support vector machine, engine control

1. INTRODUCTION

To reduce pollution, the diesel engine has to adhere to strict emission norms being imposed by the regional legislation bodies. This transformation includes introducing more sensors, actuators and newer technologies e.g. air-path and the fuel-path (Eriksson and Nielsen, 2014). These technologies help to achieve the demand torque while reducing emissions and fuel consumption (Eriksson and Nielsen, 2014). For these technologies, traditional feedback and feedforward controllers are used, which increase the number of subsystems and complexity of the engine control system.

The air-path control in a diesel engine has been identified as a complex and challenging control problem (Eriksson and Nielsen (2014)). Traditionally, these feedforward controllers consisted of 2-D maps, wherein the actuator position is determined based on the engine speed and torque values. The effort to generate maps exponentially increases as the number of actuators and sensors increase in the diesel engine. To tackle this challenge a more recent trend towards model-based controllers is observed, see Criens (2014). However, with the rise in the number of actuators and complexity of the controllers, the model-based approach requires a large amount of data to develop accurate models. These steps are bound to exponentially increase the calibration effort of such controllers (Mancini et al. (2014), Willems (2017)). This will result in unaccceptable development time and costs. Therefore, the calibration effort for future engine control systems has to be dramatically reduced.

A comprehensive review of Machine Learning (ML) methods for modelling, optimisation, diagnosis and control of internal combustion engines by Aliramezani et al. (2022) suggests that the calibration effort of complex engine subsystems can be reduced using Supervised Learning (SL) methods. This is in line with Garg et al. (2021), which identified parametrization of controllers using SL methods as an important first step to reduce the calibration effort as well as automation of the calibration process using reinforcement learning. This review paper further highlights that most of the literature do not attempt to study the impact of ML-based methods on system performance, calibration effort and computational requirements. Additionally Aliramezani et al. (2022) also identified a need for a benchmark studies to compare different ML for engine control methods with real world implications such as required data size, training time, prediction accuracy and required memory size. Parametric methods have been developed in e.g. Eriksson and Nielsen (2014) using physics-based approximations to develop regression models of embedded turbocharger maps in model-based feedforward controller for air-path control system. Gaussian Process Regression model has been applied to approximate the steady-state feedforward controller for air-path control (Aran and Unel, 2018). For approximation of MPC control policy for diesel engine air-path control, multi-layer perceptron (Moriyasu et al. (2019), Peng et al. (2022)) and long short-term memory recurrent networks (LSTM-RNN) (Peng et al., 2022) are applied. For feedback control of air-fuel ratio in spark-ignition engine, recurrent neural networks (RNN) have

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2022.09.173

been applied to estimate air-fuel ratio, which can be used as a virtual sensor (Zhang et al., 2007). From these studies, however, it is difficult to determine the actual potential of data-driven methods to reduce control calibration effort. Typically, the focus is only on the model accuracy while the comparison with other methods is lacking.

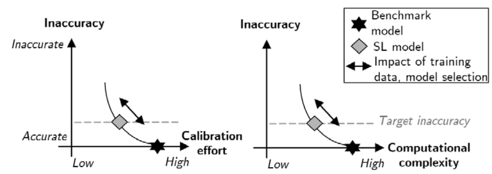

Fig. 1. Systematic method for performance comparison

To address these challenges, the main contribution of this work is the introduction of a systematic method for performance comparison of data-driven methods in engine controls development applied in steady state engine calibration, as illustrated in Figure 1. Clear performance criteria are specified, which will assist control experts to select the most promising method for their control problem. Also, it helps to understand the existing trade-offs between performance and implementation aspects. The method is applied to assess the performance of two types of supervised learning (SL) regression models compared to the benchmark physics-based model for a Diesel air path boost controller.

2. DIESEL ENGINE AIR-PATH BOOST CONTROL

2.1 Engine specification

The engine considered for this study is a Euro-5 passenger car Diesel engine. The air-path system consists of a twin-turbocharger with a high pressure exhaust gas recirculation (EGR) system. The twin-turbocharger, consisting of high-pressure (HP) and low-pressure (LP) stages, are each equipped with a wastegate valve and a compressor bypass to control the intake manifold pressure p_2. To understand if the calibration effort can be reduced using SL methods, this study focuses on HP turbocharger operation region. Additionally, an intake throttle valve is used to maintain the pressure difference across the intake and exhaust manifolds dp above a minimum threshold. The HP EGR valve controls the EGR mass flow.

2.2 Engine control

High-level control objective The engine control system aims to realize the driver's torque demand $M_{e,dem}$ with minimal fuel consumption $BSFC$ while meeting safety and tailpipe emission targets. On the engine level, this is achieved by a coordinated air-fuel path control. To demonstrate the potential of SL methods to reduce calibration effort, we focus on air-path boost control system, which requires significant control development effort. Figure 2 illustrates the total air path control scheme. The focus of this study, the feedforward controller (C_{ff}), computes the nominal control inputs $\mathbf{u_0} = [u_{0,wg} \;\; u_{0,th} \;\; u_{0,cb}]^\top$ for the wastegate, intake throttle and compressor bypass:

Air path feedforward controller The studied model-based feedforward control aims to track the specified reference for the boost pressure r_{p2}. Figure 3 shows the breakdown of the benchmark feedforward controller. The desired settings for the control inputs $\mathbf{u}^* = [u^*_{th} \;\; u^*_{wg} \;\; u^*_{cb}]^\top$ are determined based on observer $\mathbf{q}$ and measurement $\mathbf{y}$ information:

$$\mathbf{q} = [T_3 \;\; p_4 \;\; \dot{m}_c \;\; \dot{m}_{em} \;\; \Delta p_{IC} \;\; p_1 \;\; \Delta p_{wg}]^\top$$
$$\mathbf{y} = [y_{p_2}, y_{T_1}, y_{p_3}, y_{\dot{m}_{air}}]^\top$$

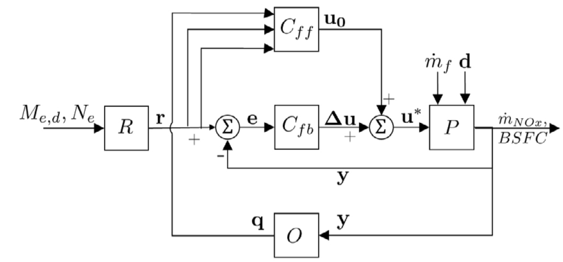

Fig. 2. Air path control scheme, where R is setpoint generator, C_{ff}, C_{fb} feedforward and feedback controllers, P engine model and O observer

where T_3 is the exhaust gas temperature, p_4 is the pressure after the turbine, $\dot{m}_c$ represents the mass flow through the HP compressor calculated based on the fress air flow rate $y_{\dot{m}_{air}}$, $\dot{m}_{em}$ is exhaust manifold gas mass flow rate computed using$\dot{m}_f$, Δp_{IC} represents the pressure drop across the intercooler and p_1 and y_{T_1} represents the pressure and temperature upstream of the HP compressor. The term Δp_{wg} denotes the pressure difference across the wastegate valve. The model-based feedforward controller is centered around a physics-based turbocharger model. As illustrated in Table 1, this approach requires large number of maps and parameters to be calibrated to ensure optimal functionality of the feedforward controller in different operating conditions and engine combustion modes. To ensure safe engine operation, rule-based logics are used in addition to these models. The calibration of such large number of parameters requires large expert effort and development times. Among the different models in the feedforward controller, the calibration of the compressor and turbine models requires the largest calibration effort. Therefore, as a starting point, these two models are parametrized using SL regression methods, which is discussed in the next section.

Table 1. List of calibration parameters in the benchmark air-path feedforward controller. Each element of look-up table is considered as one parameter.

Calibration parameter	**Type of map**	**Number of parameters**
	Compressor model	
Heat capacity of air	$c_p(T_1)$	12
Compressor efficiency map	$\eta_c(\dot{m}_c, \pi_c)$	494
Exponential function for pressure ratio calculation	$f(\gamma)$	120
Compressor acceleration parameter	$a_c(N_c)$	12
Compressor speed map	$N_c(\dot{m}_c, \pi_c)$	256
	Turbine model	
Turbine mass flow	$m_t(N_{t,r}, \pi_t)$	720
Pressure ratio	$\pi_t(P_c, N_{t,r})$	720
Turbine map high and low limits	constants	12
	Wastegate model	
Linearization	$u_{0,wg}(A_{wg}, \dot{m}_t)$	9

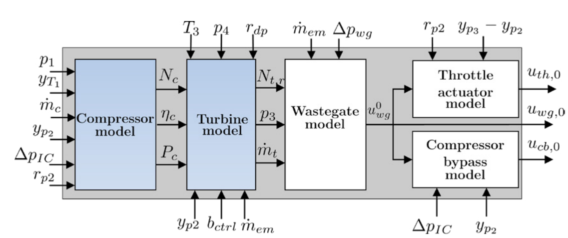

Fig. 3. Scheme of the air-path boost feedforward controller. The focus of this paper is highlighted in blue.

3. SYSTEMATIC METHODOLOGY FOR CONTROLLER PARAMETRIZATION

In this study, supervised learning (SL) regression models are applied, which have the following structure:

$$\hat{\mathbf{Y}}_{\mathbf{i}} = f(\mathbf{W}_{\mathbf{i}}, \phi_i(\mathbf{X}_{\mathbf{i}})) \tag{1}$$

where the prediction $\hat{\mathbf{Y}}_{\mathbf{i}}$ is a function of a vector of model parameters $\mathbf{W}_{\mathbf{i}}$ and ϕ_i is the non-linear basis function. The index $i = \{c, t\}$ refers to the compressor or turbine model. In the development of these data-driven models, five main steps can be distinguished:

(1) **Raw data generation** using a mean value engine simulation model. This model is developed and validated using experimental engine data;
(2) **Data pre-processing** for selection of model inputs $\mathbf{X}$ and outputs or labels $\mathbf{Y}$;
(3) **Model selection** using multiple performance criteria;
(4) **Model training and validation** to determine the model parameters $\mathbf{W}$ for the selected SL methods and evaluate model performance on a validation data set;
(5) **Controller performance testing** to compare the system level performance of engine out NOx emissions for the regression based controller with the benchmark controller on a test cycle.

3.1 Data generation

Typically, the engine control systems are calibrated for steady state performance initially. To further improve the performance of the control systems, transient factors are considered. In this study, only the first stage of steady-state engine operation is considered. Figure 4 shows the variation in the engine torque demand and engine speed as per the steady state design of experiments (DOE) generally used in automotive industry to calibrate engine controllers. The steady state approximation of worldwide harmonized light-duty vehicles test cycle (WLTC) is used as a test dataset to evaluate the performance of the controller on an independent dataset. Additionally, the engine out NO_x emissions performance of the developed regression models are compared with benchmark physics-based controller.

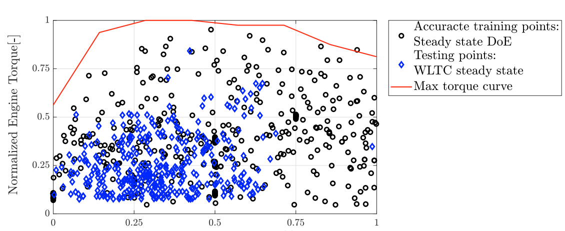

Fig. 4. Engine speed and torque distribution for model training and testing.

3.2 Data pre-processing

The selection of statistically significant inputs and outputs of the regression model is one of the key steps in the data pre-processing. First, the outputs $\mathbf{Y}$ of the compressor and turbine models are selected. With these outputs, the inputs $\mathbf{X}$ to the regression models are identified. First, available inputs are identified from the raw data. Next, statistical significance of each input for predicting the outputs are calculated using the Spearman correlation (Zheng and Casari, 2018). This is used as a criteria to determine the necessary inputs. A Spearman correlation of ± 1 indicates the input is statistically significant and a value of 0 indicates no significance of the input. In this study, the inputs with Spearman correlation greater than 0.75 are considered as statistically significant and selected. The selected inputs are ensured to be easily measured/observed for data collection.

For the compressor model, the desired outputs are compressor power, speed and efficiency $\hat{\mathbf{Y}}_{\mathbf{c}} = [P_c \; N_c \; \eta_c]^T$. The significant inputs for the compressor model are found to be $\mathbf{X}_{\mathbf{c}} = [y_{p2} \; r_{p2} \; p_1 \; T_1 \; \dot{m}_c]^T$. In a similar way, the desired outputs of the turbine model are exhaust manifold demand pressure, turbine mass flow and reduced speed $\hat{\mathbf{Y}}_{\mathbf{t}} = [y_{p3} \; \dot{m}_t \; N_{t,r}]^T$. Note that the turbocharger speed $N_{t,r}$ is added to detect for abnormal values in steady state data collected. The significant inputs for the turbine model are found to be $\mathbf{X}_{\mathbf{t}} = [r_{dp} \; y_{p2} \; P_c \; T_3 \; p_4 \; b_{ctrl} \; \dot{m}_{em} \; y_{p3} \; N_c]^T$, where b_{ctrl} denotes the switch for boost control. To improve the learning capabilities of the regression models, the inputs and outputs are scaled using a standard scalar based on the standard deviation and mean of the signal values.

3.3 Model selection

A wide range of SL regression methods can be found in literature, which can be used to predict quantitative continuous outputs (Bishop, 2006). These include methods such as, multiple linear regression, artificial neural networks (ANN) and kernel based methods. To select the most promising methods from the results established in the literature, five criteria that impact engine control development time and costs are identified:

(1) Model prediction inaccuracy $= \left(1 - R^2\right) + (|e_{max,i}|/Y_i)$
(2) Number of calibration parameters $= \theta_{sl}/\theta_b$
(3) Input dimensions $=$ length$(\mathbf{X})$
(4) Computational complexity $= \left(\frac{\text{size}(M_{sl})}{\text{size}(M_b)}\right) + \left(\frac{t_{sl}}{t_b}\right)$
(5) Tool chain availability

The prediction inaccuracy of various regression methods is assessed by also studying the results of such methods in fields of thermodynamics (Raghunatha Reddy et al., 2020; Ahmad et al., 2018) and tribology (Ikpambese and Lawrence, 2018). These sources were considered due to the scarcity of comparative studies of data-driven methods for the specific case of engines. Note that the prediction accuracy of such models are typically expressed by the the coefficient of determination R^2 in literature. In addition to the R^2 value, in this study, the relative maximum absolute deviation $(|e_{max,i}|/Y_i)$ is also considered, where $e_{max,i} = \max(\hat{Y}_i - Y_i)$. This performance metric indicates the region of the deviation in predictions. Identifying this region helps to improve, if required, the model performance in the region of deviation.

An important indication for the required calibration effort is the number of calibration parameters. For SL regression methods, only the hyper-parameters need to be tuned before training the model. The hyper-parameters θ_{sl} are tuning parameters that control the pattern learning capabilities of such regression models. These parameters are usually tuned by a human expert before running the optimisation loop to learn the model parameters $\mathbf{W}$. The total calibration effort of SL models involve data generation, data pre-processing and model training and validation. The maximum reduction in calibration effort using regression models is achieved by parameterizing the models used in the current model based approach. The input dimensions is the number of inputs that the model is capable of handling. For example, a few regression models face the curse of dimensionality with increase in the number of inputs, see Bishop (2006) for more details. Computational complexity is defined as the sum of normalized memory requirements i.e., size(M_i) and normalized time requirements i.e., t_i, of the supervised learning $(i = \{sl\})$ and benchmark $(i = \{b\})$ models, respectively. For initial selection, worst case memory and time requirement of the prediction algorithms were obtained from the studies of Pedregosa et al. (2011).

A total of 13 methods were chosen from literature. These methods were rated and two methods were identified as most suitable for this study: Artificial Neural Networks (ANN) and Support Vector Machines (SVM). From literature, it is evident that SVM and ANN models have relatively high accuracy. However, SVM models are computationally more expensive than the ANN models. Thus, a comparison between two different models of high accuracy with high and low computational complexity is found interesting for this study.

3.4 Model training

Artificial Neural Network (ANN): ANN is a parametric SL regression method capable of accurately learning complex non-linear patterns between multiple inputs and multiple outputs. The model architecture is designed by choosing the number of hidden layers and the number of neurons in each hidden layer. This choice fixes the parameters $\mathbf{W}$ to be learnt. The prediction equation of the p^{th} output for a one hidden layer feedforward ANN is,

$$\hat{y}_p = \sigma\left(\sum_{j=0}^{n} w_{p,j}^{(2)}\ h\left(\sum_{i=0}^{n} w_{j,i}^{(1)} x_i\right)\right) \tag{2}$$

where n is the number of neurons in the hidden layer, activation function $\sigma = 1$ for regression, h is the activation function, $w \in \mathbf{W}$ are the weights and biases. The superscripts (1) and (2) denote the input and the hidden layer respectively. Furthermore, i denotes the number of inputs to the model and j denotes the number of neurons in the hidden layer. In this study, the activation function of Rectified Linear Unit (ReLU) is selected. The total model parameters $\mathbf{W}$ of the ANN model are learnt by minimising the error function between the true and the predicted values as follows,

$$\min_{\mathbf{W}} \mathcal{J}(\mathbf{W}, e) = \min_{\mathbf{W}} \frac{1}{2}\beta\mathbf{W}^T\mathbf{W} + \frac{1}{n}\sum_{n=1}^{N} e_n^2 \tag{3}$$

Where, model error $e = \mathbf{Y} - \hat{\mathbf{Y}}$, $\hat{\mathbf{Y}}$ denotes the vector of all the p predicted outputs $\hat{y}_p$, and $\mathbf{W}$ denotes the vector of all weights and biases. The network represented by equation (2) is intitalized with random weights $\mathbf{W}$ and then the optimum weights are obtained by minimizing the error between the predicted value and the actual label. This optimization problem is a non-convex problem and is solved using a stochastic gradient descent algorithm with momentum, since it has the advantage of being computationally less expensive than other algorithms and helps move out of saddle points (Bishop, 2006). Finally, to prevent model over-fitting to the training data, a L2 regularisation term β is used to modify the cost function. See Bishop (2006) for more detailed information on the hyper-parameter tuning and optimization algorithms of the ANN model.

Support Vector Machines (SVM): SVM methods are kernel-based methods and learn patterns between multiple inputs and single output. These kernel-based methods do not rely on fixing a model architecture before the model training. In SVM, a subset of the training data is stored in the memory for making predictions. Such methods are referred to as non-parametric methods (Bishop, 2006). In this study, the Radial Basis Function (RBF) kernel is used based on the results of a sensitivity study among RBF, Matern 3/2 and 5/2 and exponential kernels. The RBF Kernel has low complexity and is also infinitely differentiable. This RBF kernel is given by:

$$k(\mathbf{X}, \mathbf{x}_n) = \phi(\mathbf{X})\phi(\mathbf{x}_n)^\top = e^{\left(\frac{-||\mathbf{X} - x_n||^2}{\sigma^2}\right)} \tag{4}$$

where $\mathbf{X}$ denotes the stored input support vectors and $\mathbf{x}_n$ is the input vector in query. Then, the p^{th} SVM model output is described by:

$$\hat{y}_p = \sum_{n=1}^{N} w_{p,n}\phi_p(\mathbf{x}_n) + b_p \tag{5}$$

The support vectors for the SVM model are obtained by solving the following optimization problem:

$$\min_{w,b,\xi,\hat{\xi}} \mathcal{J}(w, e) = \min_{w,b,\xi,\hat{\xi}} \frac{1}{2}w^T w + C\sum_{n=1}^{N} L_\epsilon(e_n) \tag{6}$$

where:

$$L_\epsilon(e) = \begin{cases} 0, & \text{if } \ |e| < \epsilon \\ |e| - \epsilon, & \text{otherwise} \end{cases}$$
$$e_n = \hat{y}_n(\mathbf{X}_n) - y_n \tag{7}$$

subject to

$$\hat{y}_n - y_n \le \epsilon + \xi_n$$
$$\hat{y}_n - y_n \ge \epsilon + \hat{\xi}_n$$
$$\xi_n, \hat{\xi}_n \ge 0 \text{ for all } n = 1 \ldots N$$

The loss function L_ϵ introduces sparseness and reduces memory required to store the support vectors (Bishop, 2006). Additionally, slack variables ξ are introduced, which indicate the distance to the decision boundary $\hat{y}_n - y_n$. These characteristics of the optimisation problem make it a convex optimisation problem and is solved using the sequential minimal optimisation algorithm. When this problem is converted to its dual form using the concept of Lagrange multipliers and its Karush–Kuhn–Tucker (KKT) conditions are solved, we obtain:

$$w_{p,n} = \sum_{n=1}^{N}(\lambda_{n,1} - \lambda_{n,2})\phi_p(\mathbf{X}_n) \tag{8}$$

which shows that the weights $w_{p,n}$ required for a single output p are only a function of the Lagrange multipliers λ_n and the kernel chosen and not of any model architecture (Bishop, 2006). The dependency on training data implies that the SVM method faces a challenge with memory requirement as it is directly proportional to the number of inputs. Thus with large number of inputs, the memory complexity rises significantly.

To reduce the memory complexity, a chained approach is proposed in this study where output of one model is input to the next, as shown in Figure 5. The structure for the chained model are again determined by using the statistically significant outputs for each model.The SVM models are trained using a 5-fold cross-validation to reduce the model over-fitting.

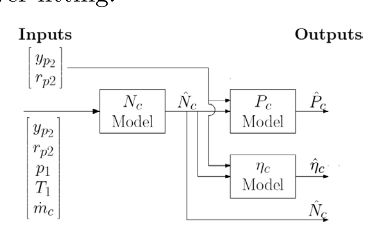

Fig. 5. Chained SVM models for the compressor outputs. A similar approach is followed for the turbine model.

Hyperparameters The 10 hyperparameters θ_{ANN} and their values for the selected ANN model are mentioned in Table 2. Similarly, the four hyperparameters θ_{SVM} for the SVM model are listed in Table 3. Since SVM model

Table 2. List of ANN model hyperparameters

Sl no.	Hyper parameter	Description	Compressor model value	Turbine model value
1	n	No. of neurons per layer	14	15
2	m	No. of hidden layers	1	1
3	h	Activation function	ReLU	ReLU
4	α	Learning rate	0.0035	0.005
5	γ	Momentum	0.9	0.9
6	s_{data}	Data Split (Training/Validation)	90/10	90/10
7	λ	Regularization factor	1e-4	1e-4
8	n_{mb}	Minibatch size	5	1
9	f_d	Drop-out factor	0.2	0.2
10	n_e	No. of epochs	48	3

Table 3. List of SVM model hyperparameters

Sl no.	Hyperparameter	Description
1	$\frac{1}{2\sigma^2}$	Characteristic length
2	C	Box-constraint
3	ϵ	Margin parameter
4	b	bias

can only predict one output at a time, a total of 24 hyper-parameters are required for the six models that are required in the controller. On the contrary, the ANN model can predict multiple outputs and hence requires 20 hyper-parameters to be tuned for the two models. The most commonly used methods to tune the hyper-parameters of a SL regression model are the grid search and the random search (Bergstra et al., 2013). However, these two methods suffer from extremely large search space and a possibility of repeating points in the region of poor performance. To avoid these drawbacks, a Bayesian optimisation based framework (Bergstra et al., 2013) is used for SVM and trial-and-error approach for the ANN model.

4. RESULTS AND DISCUSSIONS

4.1 Model validation

For the applied two SL regression methods, the performance on an independent test set is compared for three performance criteria described in section 3.3: model prediction inaccuracy, number of calibration parameters and computational complexity. The results are compared for the following models:

(1) Benchmark (physics-based model),
(2) SVM individual models,
(3) SVM chained models
(4) ANN models.

The performance of these models are calculated by combining the prediction of the compressor and turbine models at a system level on the test dataset. Overall model inaccuracy is only determined for turbine model outputs Y_t.

Impact of SL model Figure 6 shows the trade-off between the model prediction inaccuracy and the number of calibration parameters. A significant reduction of at least 97% in the number of calibration parameters compared to the benchmark model is achieved with all the three regression models. The engine control development process consists of control system design, experimentation, modelling, control calibration and validation. This significant reduction directly implies reduction in the required expert effort and development times for control calibration. A marginal reduction in calibration parameters is observed in the ANN model compared to the SVM models. The fully connected feature of the ANN model helps to learn the desired patterns better than the SVM model. Furthermore, the ANN model shows the best overall performance based on these criteria. This observation strongly recommends that the interaction between the different inputs play a crucial role in predicting the desired outputs.

Additionally, the trade-off between inaccuracy and computational complexity is shown in Figure 6. The SL models are implemented in the MATLAB Simulink environment and all computations are performed on a standard laptop with 16 GB of RAM, Intel i7-9750 processor running at 2.60 Ghz. The SVM model with best validation loss is at least 34 times computationally more expensive than the benchmark model due to the higher memory requirement of the SVM model and the non-linear RBF kernel. The significantly larger memory requirement of SVM model compared to ANN model makes it less suitable choice for real-time implementation on the Engine Control Unit (ECU). On the other hand, the ANN models are found to be 20% computationally less expensive than the benchmark model due to simplicity of the ReLU activation function and significantly small number of model parameters.

Impact of training dataset size The amount of the training dataset is directly proportional to the number of experiments. The training dataset size was varied and the impact on the three performance criteria was evaluated for the SVM model. The dataset size was reduced until the validation error begins to increase. The maximum possible reduction in dataset size was 55%, which causes a decrease of 15% in accuracy of the SVM chained model with a benefit of 3.7% computational complexity as shown in Figure 6. This trade-off between accuracy and computational complexity is acceptable with the reduced dataset because the system inaccuracy i.e., cumulative NO_x emissions over the test cycle are within the target inaccuracy of ±10%. This observation suggests that underlying patterns can be learned with the right distribution of the dataset.

4.2 Controller performance testing

Figure 7 shows the normalized cumulative engine out NOx emissions on the WLTC for the benchmark controller (in black). The overall deviation with the ANN model and SVM chained model is 9.75% and 10.15%, respectively. These deviations are close to the allowable inaccuracy of 10% compared with the benchmark controller, where the maximum deviation is due to the transients between 700-800 [s] of the cycle. The system-level accuracy is acceptable for the regression methods, despite the significant prediction inaccuracy of 0.35 in ANN and 0.44 in SVM chained model. The inaccuracies in the individual model trained models on wastegate position is marginal implying that the deviation in reference tracking performance with respect to the benchmark model is within acceptable limits.

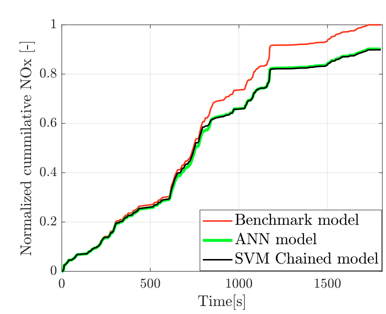

Fig. 7. Comparison of normalized WLTC cumulative engine out NOx emission of SL models and benchmark.

5. CONCLUSIONS AND FUTURE RESEARCH

In this study, a systematic method is introduced to assess and compare the performance of different SL methods for engine control development. The essential part of this method is the specification of selection criteria for not only model accuracy, but also for overall system performance, computational complexity, and calibration effort. These methods are successfully applied to the diesel engine air path control problem to assess essential criteria for real-world implementation. Based on the proposed selection criteria, two promising SL methods are selected: ANN and SVM. The results of the systematic comparison are summarised in Figure 8. Compared to the benchmark feedforward controller, which uses physics-based models, the following conclusions are drawn:

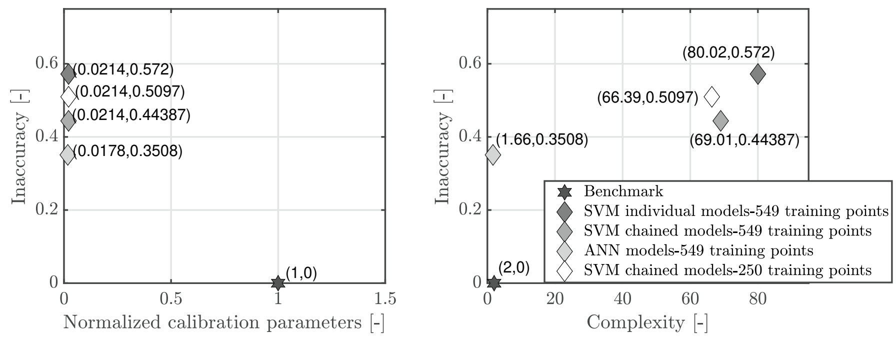

Fig. 6. Trade-offs of model inaccuracy, number of calibration parameters and computational complexity for the SL models and physics-based models from benchmark controller.

- 97% reduction in calibration parameters is achieved for both SL models, which reduces expert effort and development costs. However, a trade-off between inaccuracy and computational complexity through model selection is necessary to match the current state-of-art ECU's.;
- Model inaccuracy of both regression models is moderate, but acceptable. The model inaccuracy metrics are 0.35 for the ANN model and 0.44 for the best SVM model. These inaccuracies translate to deviations close to the allowable engine out NOx deviation of 10% on the system level;
- ANN model has lower computational and memory requirements than the benchmark. In contrast to ANN, the SVM chained model has larger computational complexity owing to the complex non-linear kernel and the memory requirement;
- Both SL models can handle models with large input dimensions. However, for the SVM method increased input dimensions go hand in hand with increased memory requirements;

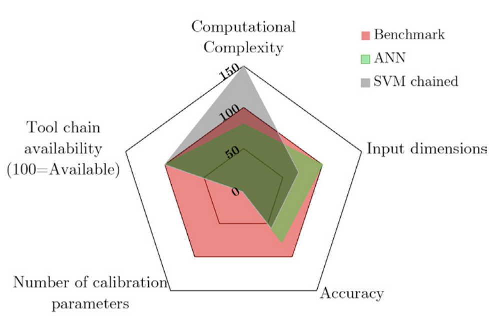

Fig. 8. Comparison of the selected ANN and SVM-based controller with the benchmark controller.

Based on this assessment, it is concluded that the (parametric) ANN model overall performs better than the (non-parametric) SVM model for the studied case. Future research concentrates on further improving the understanding of the trade-offs (inaccuracy-calibration effort-computational complexity) for the ANN model. Especially, to specify the minimal required training data set and the required model complexity. In order to assess the practical relevance of the ANN-based air path controller, its potential to efficiently cover multiple engine operating modes is interesting. Moreover, parameterized controllers are suitable to be combined with adaptive control algorithms to reduce development times by automated calibration using on-line learning or over the cloud updates to deal with system uncertainties e.g. component ageing.

ACKNOWLEDGEMENTS

The authors would like to thank DENSO for providing financial support for this study and DENSO Aachen Engineering Centre (AEC) Powertrain R&D team for the constructive discussions and also for the help with the model-based benchmark model in simulation.

REFERENCES

Ahmad, M.W., Reynolds, J., and Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *Journal of Cleaner Production*, 203, 810–821.

Aliramezani, M., Koch, C.R., and Shahbakhti, M. (2022). Modeling, diagnostics, optimization, and control of internal combustion engines via modern machine learning techniques: A review and future directions. *Progress in Energy and Combustion Science*, 88, 100967.

Aran, V. and Unel, M. (2018). Gaussian process regression feedforward controller for diesel engine airpath. *International Journal of Automotive Technology*, 19(4), 635–642.

Bergstra, J., Yamins, D., Cox, D.D., et al. (2013). Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, volume 13, 20. Citeseer.

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics).* Springer-Verlag, Berlin, Heidelberg.

Criens, C. (2014). *Air-path control of clean diesel engines : for disturbance rejection on NOx, PM and fuel efficiency.* Ph.D. thesis, Technische Universiteit Eindhoven.

Eriksson, L. and Nielsen, L. (2014). *Modeling and Control of Engines and Drivelines.* John Wiley and Sons, Ltd.

Ikpambese, K. and Lawrence, E. (2018). Comparative analysis of multiple linear regression and artificial neural network for predicting friction and wear of automotive brake pads produced from palm kernel shell. *Tribology in Industry*, 40(4), 565.

Mancini, G., Asprion, J., Cavina, N., Onder, C., and Guzzella, L. (2014). Dynamic feedforward control of a diesel engine based on optimal transient compensation maps. *Energies*, 7(8), 5400–5424.

Moriyasu, R., Nojiri, S., Matsunaga, A., Nakamura, T., and Jimbo, T. (2019). Diesel engine air path control based on neural approximation of nonlinear mpc. *Control engineering practice*, 91, 104114.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V.and Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Peng, Q., Huo, D., and Hall, C.M. (2022). Neural network-based air handling control for modern diesel engines. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering.*

Raghunatha Reddy, D.V., Bhramara, P., and Govindarajulu, K. (2020). A comparative study of multiple regression and artificial neural network models for a domestic refrigeration system with a hydrocarbon refrigerant mixtures. *Materials Today: Proceedings*, 22, 1545–1553.

Willems, F. (2017). *The self-learning powertrain : towards smart and green transport.* Eindhoven University of Technology. Inaugural lecture.

Zhang, Y., Xi, L., and Liu, J. (2007). Transient air-fuel ratio estimation in spark ignition engine using recurrent neural networks. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, 240–246. Springer.

Zheng, A. and Casari, A. (2018). *Feature engineering for machine learning: principles and techniques for data scientists.* " O'Reilly Media, Inc.".